QUANTUM PHYSICS III

2026/2027

Lecture notes.

Monday 7th September, 2026

Juan José Saborido Silva

Cibrán Santamarina Ríos

Contents

Preface.

Dear reader:

These lecture notes are aimed at students attending the 42-hour *Quantum Physics III* course at the University of Santiago de Compostela (USC). They rely heavily on various chapters and sections of the books listed in the bibliography, as well as on other sources. Most (though not all) of what is written here is a rephrasing of what you can find in those books, and one might wonder whether it would not be better to just refer you to the relevant chapters and sections of the books. We do in fact both things. We do refer you to the books and encourage you to read them, but we also make a selection of the material to cover in our course and present it here in a way that, hopefully, allows you to go through the mathematical derivations with less effort. At a time when excellent text processors are readily available, we can afford to include more intermediate algebraic steps than would typically appear in a published textbook. We do not, however, work out in detail everything. In any time-limited course, one must decide when to perform every step of a calculation and when to skip to the final result in order to cover more material. Of course, not all the calculations have the same pedagogical value: similar calculations need not be repeated, and others lie beyond the scope of this course.

It follows a short list of the books that have been more extensively used when preparing the lectures. Other references are listed in the bibliography at the end. It should be clear, in any case, that these notes are not intended to replace them in your study.

- M. Weissbluth. *Atoms and Molecules*, [16]. Chapters 15, 16, 17, 18, 19, 22 and 23. This book was published in 1978 and there has never been a second edition, but it remains an excellent source of material. The first 14 chapters offer an unusually long introduction to mathematical tools and a review of key quantum mechanics topics.

- D. J. Griffiths and D. F. Schroeter. *Introduction to Quantum Mechanics* [7]. At least chapters 8 and 10, but other chapters should be useful also. This is a clear and pedagogical textbook that you should enjoy reading. The second edition of this book, which is in the library, is also excellent.

- Bransden, B. H. and Joachain, C. J., *Physics of atoms and molecules* [2]. The second edition is a 1114 pages volume with a wide coverage of atomic and molecular physics. The first edition, at least, is in the library.

Of course the bibliography on quantum physics is very vast. A few additional textbooks that you might find useful are

- Sánchez del Río (coordinador). *Física Cuántica*, [3].

- R. Eisberg and R. Resnick. *Quantum Physics of Atoms, Molecules, Solids, Nuclei and Particles*, [4].

- M. Fox. *A Student's Guide to Atomic Physics*, [5].

- H. Haken and H.C. Wolf. *The Physics of Atoms and Quanta*, [9].
- J.J. Sakurai and J. Napolitano, *Modern Quantum Mechanics*, second edition, [13].
- S. Weinberg, *Lectures on Quantum Mechanics*, second edition, [15].

Bear in mind that this document does not necessarily cover all the material discussed in the classroom and cannot be seen as *the complete* reference for the course. It is also bound to be edited frequently, with the date appearing at the head of the pages.

One final note: since several figures and plots have been freely taken from books and other sources without asking permission (only for pedagogical reasons), these lecture notes *should not be circulated* outside the restricted domain of lecturer-student communication at the University of Santiago de Compostela.

Chapter 1

The hydrogen fine structure.

To find the relativistic corrections to hydrogen-like atomic energy levels we need to use the Dirac equation. This equation can be exactly solved for the Coulomb potential. However, the calculations are rather lengthy and since the relativistic corrections are small (provided Z is not too large), it is convenient to use perturbation theory keeping only terms up to order v^2/c^2 in the Dirac Hamiltonian.

Before getting to that, as a gentle starter, it is useful to review the very important role symmetries play in physics, and how they are implemented in quantum mechanics. In the historical development of quantum mechanics, a lot of guidance from classical physics was obtained to get explicit expressions for observables, specially Hamiltonians. Nevertheless, the recourse to the classical world is not always possible and, to a large extent, symmetry principles alone seem to provide the most solid guidance to obtain the correct expressions for observable quantities. In addition, one would say that deep understanding comes only when the fundamental equations of the theory can be derived from a few symmetry principles.

Since symmetries in the context of quantum mechanics have already been discussed in *Quantum Physics II*, the treatment here is kept to a minimum review. The material in the next section is more extensively discussed in, for example, chapter 3 of reference[1] [15], or in the first chapter of [13].

1.1 Postulates and symmetries, a brief review.

In quantum mechanics, physical states are represented as vectors in a Hilbert space, called also state vectors. The properties of a system, or observable quantities than can be measured, are represented by Hermitian operators acting on the states of the Hilbert space.

> Remember that an operator A is hermitian if $A^\dagger = A$, where $A^\dagger$ is the adjoint of A, defined such that $\langle \phi \,|\, A^\dagger \,|\, \psi \rangle = \langle \psi \,|\, A \,|\, \phi \rangle^*$. The matrix elements of $A^\dagger$ are related to those of A (denoted as a_{ij}) through $(A^\dagger)_{ij} = a^*_{ji}$. When we use the Dirac bra-ket notation to write $\langle \phi | A | \psi \rangle$, the operator A is *always* assumed to act on the sate to its right. Sometimes we could also write $\langle \phi | A | \psi \rangle = \langle \phi | A\psi \rangle$, but note carefully that, in general, $\langle \phi | A | \psi \rangle \neq \langle A\phi | \psi \rangle$, unless A is hermitian. Some authors, particularlly S. Weinberg, use the notation (ϕ, ψ) for the scalar product instead of $\langle \phi \,|\, \psi \rangle$. This notation is commonly used by mathematicians, and the operators are written in a natural way acting on the states as $(\phi, A\psi)$ or $(A\phi, \psi)$, as appropriate. For some purposes the mathematical notation is more clear than the one introduced by Dirac. For example, the adjoint $A^\dagger$ can simply be defined as the operator that satisfies $(\phi, A^\dagger \psi) = (A\phi, \psi)$. In this course we will use the Dirac notation.

[1] *Lectures on quantum mechanics*, by S. Weinberg, is an excellent (graduate level) textbook. Chapter 3, in particular, is a gem. We advise you to read it at some point in the future, especially if you plan to pursue eventually postgraduate physics studies.

When the property represented by the operator[2] A is measured on a system in a state Ψ, the result is necessarily one of the eigenvalues of A, say a. After the measurement, the state vector of the system is the eigenstate of A corresponding to the eigenvalue a, say ϕ_a. The probability of obtaining that particular value in the measurement process is given by (for discrete spectra)

$$P(a) = \frac{|\langle \phi_a \,|\, \Psi \rangle|^2}{\langle \phi_a \,|\, \phi_a \rangle \langle \Psi \,|\, \Psi \rangle}. \tag{1.1}$$

The state vector of the system evolves in time according to the Schrödinger equation,

$$i\hbar \frac{\mathrm{d}\Psi(t)}{\mathrm{d}t} = \mathcal{H}\Psi(t). \tag{1.2}$$

The hermitian operator $\mathcal{H}$ is the Hamiltonian of the system and represents its total energy.

The previous paragraphs encapsulate all principles or postulates of quantum mechanics. It remains now to decide how to choose the operators. Classical mechanics offers of course valuable guidance, but has a limited reach, so let us see how symmetry principles can help us. Symmetries in quantum mechanics are represented by unitary linear operators, $U^\dagger = U^{-1}$, which transform state vectors keeping transition probabilities invariant.

$$\langle \Psi_2 \,|\, \Psi_1 \rangle \;\longrightarrow\; \langle \Psi_2' \,|\, \Psi_1' \rangle = \langle U\Psi_2 \,|\, U\Psi_1 \rangle = \left\langle \Psi_2 \,\middle|\, U^\dagger U \Psi_1 \right\rangle = \langle \Psi_2 \,|\, \Psi_1 \rangle \,. \tag{1.3}$$

Specially important symmetries are those that can be represented by unitary linear operators that can be arbitrarily close to the identity operator, $\mathbf{1}$. In that case we may write

$$U_\epsilon = \mathbf{1} + i\epsilon T + O(\epsilon^2), \tag{1.4}$$

where ϵ is an arbitrary real infinitesimal number, and T is some operator which does not depend on ϵ. To first order in ϵ, the unitary condition is

$$U_\epsilon^\dagger U_\epsilon = \Big(\mathbf{1} - i\epsilon T^\dagger + O(\epsilon^2)\Big)\Big(\mathbf{1} + i\epsilon T + O(\epsilon^2)\Big) = \mathbf{1} + i\epsilon(T - T^\dagger) + O(\epsilon^2) = \mathbf{1} \quad\Longrightarrow\quad T = T^\dagger. \tag{1.5}$$

Therefore Hermitian operators arise naturally associated to infinitesimal symmetries. If we now take $\epsilon = \theta/n$, where θ is some n-independent finite parameter, and apply the transformation n times we have

$$\lim_{n\to\infty} \left(\mathbf{1} + \frac{i\theta T}{n}\right)^n = e^{i\theta T} = \sum_{n=0}^{\infty} \frac{(i\theta T)^n}{n!} = U(\theta). \tag{1.6}$$

The operator T is known as the *generator of the symmetry*. Many of the operators representing observables in quantum mechanics are the generators of symmetries.

Under a symmetry transformation, $\Psi' = U\Psi$, the expectation value of any observable A undergoes the following transformation

$$\langle \Psi \,|\, A\Psi \rangle \longrightarrow \langle U\Psi \,|\, AU\Psi \rangle = \left\langle \Psi \,\middle|\, U^{-1} A U \Psi \right\rangle. \tag{1.7}$$

The transformation properties of the matrix elements can then be found by transforming the observables as

$$A \longrightarrow A' = U^{-1} A U. \tag{1.8}$$

This is called a *similarity transformation*. If we take $U_\epsilon = \mathbf{1} + i\epsilon T + O(\epsilon^2)$, then the operator A is transformed as

$$A \longrightarrow A' = U_\epsilon^{-1} A U_\epsilon = (\mathbf{1} - i\epsilon T) A (\mathbf{1} + i\epsilon T) = A + i\epsilon AT - i\epsilon TA = A - i\epsilon [T, A]. \tag{1.9}$$

We see then that the effect of infinitesimal symmetry transformations on any operator can be expressed through the commutation relations of the symmetry generator with that operator. In particular, if $[T, A] = 0$ the operator A does not change under the symmetry transformation T.

[2] In order to be extra clear, in this section we denote operators with uppercase letters and eigenvalues with lowercase letters. For example, $\mathbf{P} = (P_1, P_2, P_3)$ is the momentum operator and $\mathbf{p} = (p_1, p_2, p_3)$ is an eigenvalue of that operator. In future sections we will not be so strict with the notation because, hopefully, it will be clear from the context what is a number and what is an operator.

1.1.1 Time and space translations.

It seems almost unavoidable to take time-translation invariance to be a fundamental symmetry for isolated systems: the Laws of Nature should not depend on the century when they are discovered.[3] Thus, the physical results obtained from the state vectors $\Psi(t)$ and $\Psi(t+\tau)$, where τ is an arbitrary time amount, should be physically equivalent. This means there should be some linear unitary operator $U(\tau)$ such that

$$U(\tau)\Psi(t) = \Psi(t+\tau). \tag{1.10}$$

Since τ is a continuous variable that can be arbitrarily small, it must be possible to express $U(\tau)$ as $e^{i\tau T}$. Instead of the Hermitian operator T, we may write $T = -\mathcal{H}/\hbar$, so that

$$U(\tau) = \exp\left(-\frac{i\mathcal{H}\tau}{\hbar}\right). \tag{1.11}$$

This can be taken as the *definition* of the Hamiltonian $\mathcal{H}$. By setting $t = 0$ in equation (1.10) and replacing τ with t, it follows that the time-dependence of any state vector is given by

$$\Psi(t) = \exp\left(-\frac{i\mathcal{H}t}{\hbar}\right)\Psi(0). \tag{1.12}$$

Taking the time derivative we find

$$\frac{\mathrm{d}\Psi(t)}{\mathrm{d}t} = -\frac{i\mathcal{H}}{\hbar}\exp\left(-\frac{i\mathcal{H}t}{\hbar}\right)\Psi(0) = \frac{1}{i\hbar}\mathcal{H}\Psi(t). \tag{1.13}$$

This equation is usually written as

$$\mathcal{H}\Psi(t) = i\hbar\frac{\mathrm{d}\Psi(t)}{\mathrm{d}t}. \tag{1.14}$$

This is the general version of the time-dependent Schrödinger equation. Time translation invariance leads to the energy conservation law, which applies only to isolated systems.

Another important fundamental symmetry is that of spatial translations. The laws of nature should not depend on where our laboratory is located. When performing a translation by an arbitrary three-vector $\mathbf{a}$, the expectation value of any particle coordinate $\mathbf{X}$ is transformed to $\mathbf{X}+\mathbf{a}$. (Here $\mathbf{X} = (X_1, X_2, X_3)$ is the position operator). There must exist then a unitary operator $U(\mathbf{a})$ such that

$$U^{-1}(\mathbf{a})\,\mathbf{X}\,U(\mathbf{a}) \;=\; \mathbf{X}+\mathbf{a}. \tag{1.15}$$

For $\mathbf{a}$ infinitesimal, U must take the form $U_\epsilon = \mathbf{1} + i\epsilon T + O(\epsilon^2)$. We use now the Hermitian operator $-\mathbf{P}/\hbar$ instead of T, and the three-vector parameter $\mathbf{a}$ instead of ϵ, to write

$$U(\mathbf{a}) = \mathbf{1} - \frac{i}{\hbar}\mathbf{P}\cdot\mathbf{a} + O(\mathbf{a}^2). \tag{1.16}$$

This can be taken as the *definition* of the momentum operator, $\mathbf{P} = (P_1, P_2, P_3)$. Inserting the expression for the infinitesimal translation into equation (1.15) we have, to first order in $\mathbf{a}$,

$$\left(\mathbf{1} + \frac{i}{\hbar}\mathbf{P}\cdot\mathbf{a}\right)\mathbf{X}\left(\mathbf{1} - \frac{i}{\hbar}\mathbf{P}\cdot\mathbf{a}\right) = \mathbf{X} - \frac{i}{\hbar}\mathbf{X}\cdot(\mathbf{P}\cdot\mathbf{a}) + \frac{i}{\hbar}(\mathbf{P}\cdot\mathbf{a})\cdot\mathbf{X} = \mathbf{X} + \frac{i}{\hbar}[(\mathbf{P}\cdot\mathbf{a}), \mathbf{X}]. \tag{1.17}$$

From this, we conclude that the following relation must be satisfied for any infinitesimal three-vector $\mathbf{a}$

$$\frac{i}{\hbar}[(\mathbf{P}\cdot\mathbf{a}), \mathbf{X}] = \mathbf{a}, \qquad [\mathbf{X}, (\mathbf{P}\cdot\mathbf{a})] = i\hbar\mathbf{a}. \tag{1.18}$$

[3] Do not confuse time-translation with time-reversal, which has been experimentally found not to hold in weak interactions.

Taking now the first component of the $\mathbf{X}$ operator, X_1, and a translation along the X_1-axis, we have

$$[X_1, P_1 a_1] = i\hbar a_1. \tag{1.19}$$

This relation holds for any a_1 only if we have $[X_1, P_1] = i\hbar$. Likewise, we have $[X_2, P_2] = i\hbar$ and $[X_3, P_3] = i\hbar$. In a more compact form,

$$[X_i, P_j] = i\hbar\delta_{ij}. \tag{1.20}$$

A translation by a vector $\mathbf{a}$ followed by a translation by a vector $\mathbf{b}$ gives the same result as a translation by $\mathbf{b}$ followed by a translation by $\mathbf{a}$, therefore

$$U(\mathbf{b})U(\mathbf{a}) = U(\mathbf{a})U(\mathbf{b}). \tag{1.21}$$

From this follows that

$$[P_i, P_j] = 0. \tag{1.22}$$

Because they commute, we can find a complete set of eigenvectors of all three components of momentum, so we can write for an arbitrary finite translation

$$U(\mathbf{a}) = \exp\left(-i\frac{\mathbf{P}\cdot\mathbf{a}}{\hbar}\right). \tag{1.23}$$

Let $|\Phi_0\rangle$ be a one-particle state with a definite position at the origin. This is an eigenstate of the position operator $\mathbf{X}$ with eigenvalue zero. We can form an eigenstate with definite position $\mathbf{x}$ using the space translation operator,

$$|\Phi_{\mathbf{x}}\rangle \equiv U(\mathbf{X})\,|\Phi_0\rangle\,. \tag{1.24}$$

Of course, the position operator acts on this eigenstate according to

$$\mathbf{X}\,|\Phi_{\mathbf{x}}\rangle = \mathbf{x}\,|\Phi_{\mathbf{x}}\rangle\,. \tag{1.25}$$

Remember that any state $|\psi\rangle$ can be expanded in terms of the position eigenstates as

$$|\psi\rangle = \int \mathrm{d}^3\mathbf{x}\;|\Phi_{\mathbf{x}}\rangle\,\langle\Phi_{\mathbf{x}}\,|\,\psi\rangle = \int \mathrm{d}^3\mathbf{x}\;\psi(\mathbf{x})\;|\Phi_{\mathbf{x}}\rangle\,, \tag{1.26}$$

where $\psi(\mathbf{x}) = \langle\Phi_{\mathbf{x}}\,|\,\psi\rangle$ is the state wave function in the position representation.[4] A translation operation by an amount $\mathbf{a}$ transforms the position eigenstate $|\Phi_{\mathbf{x}}\rangle$ into the new position eigenstate $U(\mathbf{a})\,|\Phi_{\mathbf{x}}\rangle = |\Phi_{\mathbf{x}+\mathbf{a}}\rangle$. For an arbitrary state $|\psi\rangle$ we have then

$$\begin{aligned} U(\mathbf{a})\,|\psi\rangle = U(\mathbf{a})\int \mathrm{d}^3\mathbf{x}\;\psi(\mathbf{x})\;|\Phi_{\mathbf{x}}\rangle &= \int \mathrm{d}^3\mathbf{x}\;\psi(\mathbf{x})\,U(\mathbf{a})\,|\Phi_{\mathbf{x}}\rangle \\ &= \int \mathrm{d}^3\mathbf{x}\;\psi(\mathbf{x})\;|\Phi_{\mathbf{x}+\mathbf{a}}\rangle = \int \mathrm{d}^3\mathbf{x}'\;\psi(\mathbf{x}'-\mathbf{a})\;|\Phi_{\mathbf{x}'}\rangle\,. \end{aligned} \tag{1.27}$$

In the last step, we have just performed the change of variable $\mathbf{x}+\mathbf{a} \to \mathbf{x}'$. Hence, the wave function transforms under a translation as[5]

$$U(\mathbf{a})\psi(\mathbf{x}) = \psi'(\mathbf{x}) = \psi(\mathbf{x}-\mathbf{a}). \tag{1.28}$$

For $|\mathbf{a}|$ infinitesimal we may Taylor expand $\psi(\mathbf{x}-\mathbf{a})$ and neglect everything beyond first order,

$$\psi(\mathbf{x}-\mathbf{a}) \approx \psi(\mathbf{x}) - \boldsymbol{\nabla}\psi(\mathbf{x})\cdot\mathbf{a}. \tag{1.29}$$

[4] The wave function of the position eigenstate $|\Phi_{\mathbf{x_0}}\rangle$ itself is $\langle\Phi_{\mathbf{x}}\,|\,\Phi_{\mathbf{x_0}}\rangle = \delta^3(\mathbf{x}-\mathbf{x_0})$. Since this function is zero everywhere except at $\mathbf{x}=\mathbf{x_0}$, we obviously have $\mathbf{X}\,|\Phi_{\mathbf{x_0}}\rangle = \mathbf{x_0}\,|\Phi_{\mathbf{x_0}}\rangle$.

[5] If the sign seems strange at first glance, think that what this equation says is that the translated function ψ' evaluated at $\mathbf{x}$ is equal to the untranslated function ψ evaluated at $\mathbf{x}-\mathbf{a}$.

On the other hand, for such an infinitesimal translation the unitary operator can be written as $U(\mathbf{a}) \approx 1 - \frac{i}{\hbar}\mathbf{P}\cdot\mathbf{a}$. Therefore

$$\langle \Phi_{\mathbf{x}} \,|\, U(\mathbf{a}) \,|\, \psi\rangle = \langle \Phi_{\mathbf{x}} \,|\, \psi\rangle - \frac{i}{\hbar}\langle \Phi_{\mathbf{x}} \,|\, \mathbf{P} \,|\, \psi\rangle \cdot \mathbf{a} = \psi(\mathbf{x}) - \frac{i}{\hbar}\langle \Phi_{\mathbf{x}} \,|\, \mathbf{P} \,|\, \psi\rangle \cdot \mathbf{a}. \tag{1.30}$$

Remember that $\mathbf{a}$ is just a *fixed* three-vector parameter. Equating (1.29) to (1.30) we get

$$\frac{i}{\hbar}\langle \Phi_{\mathbf{x}} \,|\, \mathbf{P} \,|\, \psi\rangle = \boldsymbol{\nabla}\psi(\mathbf{x}). \tag{1.31}$$

In other words, the momentum operator acts on the wavefunctions as

$$\mathbf{P}\psi(\mathbf{x}) = -i\hbar\boldsymbol{\nabla}\psi(\mathbf{x}). \tag{1.32}$$

It can be seen that space translation invariance leads to the momentum conservation law. Note that the representation of the momentum operator as $\mathbf{P} = -i\hbar\boldsymbol{\nabla}$, as well as the commutation relations $[X_i, P_j] = i\hbar\delta_{ij}$, are not postulates. They both follow from the fundamental properties of spatial translations and, in fact, are closely related in the sense that the representation $\mathbf{P} = -i\hbar\boldsymbol{\nabla}$ can also be inferred directly from the commutation relations (you are invited to check this in problem 3).

This section was just a warm-up and we will not continue now with the next obvious candidate for a review: spatial rotations. You are expected to review it yourself by reading appendix (B.5).

1.2 The Dirac equation.

Here we provide a very terse introduction to the Dirac equation, which is sufficient for our purposes. For a systematic treatment you need to attend, for example, the *Quantum Field Theory* course of the USC Physics Degree academic plan. By applying the standard quantum prescriptions we have just reviewed in the previous section

$$\mathbf{p} \to -i\hbar\boldsymbol{\nabla}, \qquad E \to i\hbar\frac{\partial}{\partial t}, \tag{1.33}$$

to the classical energy-momentum relation $\dfrac{\mathbf{p}^2}{2m} + V = E$, you get the Schrödinger equation,

$$-\frac{\hbar^2}{2m}\nabla^2\Psi + V\Psi = i\hbar\frac{\partial\Psi}{\partial t}. \tag{1.34}$$

It is easier to introduce the Dirac equation using relativistic four-vector notation, so here is a brief reminder.[6] The energy-momentum *contravariant* four-vector, or four-momentum, is $p^\mu = (E/c, p_x, p_y, p_z) = (E/c, \mathbf{p}) = (p_0, p_1, p_2, p_3)$. The corresponding *covariant* four-vector is is $p_\mu = (E/c, -\mathbf{p})$. You go from one to the other using the *metric* tensor

$$\underbrace{g^{\mu\nu} = g_{\mu\nu}}_{\mathbf{g}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad p_\nu = g_{\mu\nu}p^\mu. \tag{1.35}$$

In an expression such as $g_{\mu\nu}p^\mu$, two repeated indices, one covariant and the other contravariant, are to be summed over (Einstein summation convention). The scalar product of the four-momentum with itself is

$$\begin{aligned} p\cdot p = p^\mu g_{\mu\nu} p^\nu = p^\mu p_\mu = p_\mu p^\mu = p_0 p^0 + p_1 p^1 + p_2 p^2 + p_3 p^3 \\ = p_0 p_0 - p_1 p_1 - p_2 p_2 - p_3 p_3 = \frac{E^2}{c^2} - p_x^2 - p_y^2 - p_z^2 = m^2c^2. \end{aligned} \tag{1.36}$$

[6] Although this is a powerful notation and should not give you any fear, this is essentially the only section where we use relativistic four-vectors in this course.

The quantum prescription goes over to the relativistic regime without modifications,

$$p_\mu = (E/c, -\mathbf{p}) \to i\hbar\partial_\mu, \qquad p^\mu = (E/c, \mathbf{p}) \to i\hbar\partial^\mu. \tag{1.37}$$

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right), \qquad \partial^\mu \equiv \frac{\partial}{\partial x_\mu} = \left(\frac{1}{c}\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}\right). \tag{1.38}$$

$$\Box \equiv \partial^\mu\partial_\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} - \frac{\partial^2}{\partial z^2} = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 \tag{1.39}$$

The relativistic energy-momentum relation $E^2 - \mathbf{p}^2c^2 = m^2c^4$ can be more conveniently written in covariant[7] form as $c^2p^\mu p_\mu = m^2c^4$, or just[8] $p^2 - m^2c^2 = 0$. The free-particle Klein-Gordon equation follows upon using here the canonical quantum prescription,

$$-\hbar^2\partial^\mu\partial_\mu\Psi - m^2c^2\Psi = 0, \qquad -\frac{1}{c^2}\frac{\partial^2\Psi}{\partial t^2} + \nabla^2\Psi = \left(\frac{mc}{\hbar}\right)^2\Psi, \qquad \Box\Psi + \left(\frac{mc}{\hbar}\right)^2\Psi = 0. \tag{1.40}$$

In fact, even before Klein and Gordon, Schrödinger himself considered this equation as a relativistic starting point for his theory. However, after including electromagnetic interactions in (1.40) he did not obtain the correct hydrogen fine structure and discarded it in favor of a non-relativistic equation (which does not include any fine structure at all). Another problem with the Klein-Gordon equation is that it leads to both positive and negative probability density solutions. The inevitable conclusion is that the Klein-Gordon equation does not provide a consistent description of single-particle states for a relativistic system.[9]

The origin of the negative probability density in the Klein-Gordon equation lies in the fact that it is a second-order equation in t. In searching for a first-order equation compatible with the relativistic relation $p^\mu p_\mu - m^2c^2 = 0$, Dirac tried to factorize it. Suppose you are able to write

$$(p^\mu p_\mu - m^2c^2) = (\beta^\kappa p_\kappa + mc)(\gamma^\lambda p_\lambda - mc), \tag{1.41}$$

where a summation over repeated indices is understood and β^κ and γ^λ are eight coefficients to be determined. Multiplying out the right-hand side we have

$$p^\mu p_\mu - m^2c^2 = \beta^\kappa\gamma^\lambda p_\kappa p_\lambda - mc(\beta^\kappa - \gamma^\kappa)p_\kappa - m^2c^2. \tag{1.42}$$

For this equation to hold it is necessary to make the linear term in p_κ disappear from the right-hand side, therefore we must impose the condition $\beta^\kappa = \gamma^\kappa$. We are left with

$$p^\mu p_\mu = \gamma^\kappa\gamma^\lambda p_\kappa p_\lambda. \tag{1.43}$$

The γ^κ coefficients we are searching for cannot be just numbers. We will not spend more time on this here, but it can be shown that they should be matrices of, at least, dimension 4×4

[7] The *covariant* term is used here with a different meaning to the one in the expression *covariant vector*. In brief, in a *covariant equation* both the left-hand side and right-hand side are explicitly written as tensors of the same rank, which usually can be checked directly by simple inspection counting the number of contravariant and covariant indices, taking also into account those which are contracted (summed over). This ensures that both sides of the equation transform in the same way (they are *covariant*) when changing the frame of reference. In modern terminology, a contravariant vector is just a vector, and a covariant vector is a 1-form. Nevertheless, the scientific literature is full of expressions such as covariant vector, covariant equation and contravariant vectors, so better get used to it.

[8] Note that in this equation we are doing $p^2 = p^\mu p_\mu$, which is not the same as the square of the three-momentum, $\mathbf{p}^2$. In usual practice the context makes clear the meaning of p^2, but in any case we will very seldom use covariant notation in this course.

[9] In fact, no relativistic wave equation can provide a consistent description of single-particle states at very high energies because the number of particles is not a conserved quantity: particle-antiparticle pairs can be created. The proper interpretation of the relativistic wave equations has to be found in the Quantum Field Theory Formalism. In this formalism the Dirac wave function, for example, appears as the matrix element of a quantum field between a one-particle state and the vacuum, and not as a probability amplitude.

and should satisfy the relation $\{\gamma^\mu, \gamma^\nu\} = \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}$, where $g^{\mu\nu}$ is the metric tensor defined previously. We have then

$$(p^\mu p_\mu - m^2c^2) = (\gamma^\kappa p_\kappa + mc)(\gamma^\lambda p_\lambda - mc) = 0, \tag{1.44}$$

and the Dirac equation is conventionally chosen to arise from the relation[10] $(\gamma^\lambda p_\lambda - mc) = 0$:

$$(i\hbar\gamma^\mu\partial_\mu - mc)\Psi = 0. \tag{1.45}$$

We need an explicit representation of the γ matrices. In the non-relativistic limit the most convenient one is the so-called *standard representation*, which is

$$\gamma^0 = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix}, \quad \gamma^i = \begin{pmatrix} 0 & \sigma_i \\ -\sigma_i & 0 \end{pmatrix}, \quad i = 1,2,3 \tag{1.46}$$

where the Pauli matrices are given by

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.47}$$

Explicitly,

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \tag{1.48}$$

$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{1.49}$$

The following relations are useful,

$$[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k, \qquad \sigma_i\sigma_j = i\sigma_k, \qquad \sigma_i^2 = 1, \qquad \sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}, \qquad \sigma_2\sigma_k^* = -\sigma_k\sigma_2,$$
$$\sigma_i^\dagger = \sigma_i = \sigma_i^{-1}, \qquad (\mathbf{a}\cdot\boldsymbol{\sigma})\cdot(\mathbf{b}\cdot\boldsymbol{\sigma}) = \mathbf{a}\cdot\mathbf{b} + i\boldsymbol{\sigma}\cdot(\mathbf{a}\times\mathbf{b}), \qquad \boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3). \tag{1.50}$$

In the relation $\sigma_i\sigma_j = i\sigma_k$ the indices are understood to be an even permutation of $(1,2,3)$, but if you want a general expression showing explicitly the Levi-Civita tensor you may write $\sigma_i\sigma_j = \delta_{ij} + i\varepsilon_{ijk}\sigma_k$.

The Dirac equation, $(i\hbar\gamma^\mu\partial_\mu - mc)\Psi = 0$, is written in a manifestly covariant form, but for our purposes it will be more convenient to write it using the following matrices

$$\beta \equiv \gamma^0, \quad \beta\alpha_1 = \gamma^1, \quad \beta\alpha_2 = \gamma^2, \quad \beta\alpha_3 = \gamma^3, \qquad \alpha_k = \gamma^0\gamma^k = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix}. \tag{1.51}$$

Explicitly,

$$\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \alpha_1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \tag{1.52}$$

$$\alpha_2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}, \quad \alpha_3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}. \tag{1.53}$$

[10] Of course, it doesn't really matter which relation you choose. You could as well take $(\gamma^\kappa p_\kappa + mc) = 0$ and reach the same physical conclusions. Only the notation would be slightly changed.

And the Dirac equation becomes now

$$i\hbar\left(\beta\partial_0+\beta\alpha_1\partial_1+\beta\alpha_2\partial_2+\beta\alpha_3\partial_3-\frac{mc}{i\hbar}\right)\Psi=0. \tag{1.54}$$

Multiplying on the left by β we have (since $\beta^2=(\gamma^0)^2=1$)

$$i\hbar\left(\partial_0+\alpha_1\partial_1+\alpha_2\partial_2+\alpha_3\partial_3-\beta\frac{mc}{i\hbar}\right)\Psi=0, \tag{1.55}$$

$$i\hbar\partial_0\Psi=i\hbar\left(-\alpha_1\partial_1-\alpha_2\partial_2-\alpha_3\partial_3+\beta\frac{mc}{i\hbar}\right)\Psi. \tag{1.56}$$

But $i\hbar\left(-\alpha_1\partial_1-\alpha_2\partial_2-\alpha_3\partial_3\right)=\boldsymbol{\alpha}\mathbf{p}$ and $\partial_0=\frac{1}{c}\partial_t$, therefore

$$\frac{i\hbar}{c}\partial_t\Psi=\left(\boldsymbol{\alpha}\mathbf{p}+\beta mc\right)\Psi. \tag{1.57}$$

The operator $H_{\mathrm{D}}=c\boldsymbol{\alpha}\mathbf{p}+\beta mc^2$ is known as the Dirac Hamiltonian for a free particle. The time-independent Dirac equation for a free particle is

$$(c\boldsymbol{\alpha}\mathbf{p}+\beta mc^2)\Psi=E\Psi. \tag{1.58}$$

Of course, Ψ is a four-component (bispinor) function,

$$\Psi=\begin{pmatrix}\psi_1\\ \psi_2\\ \psi_3\\ \psi_4\end{pmatrix}. \tag{1.59}$$

In components, the free-particle Dirac equation becomes a set of four equations,

$$\begin{aligned}
c(p_x-ip_y)\psi_4+cp_z\psi_3+(mc^2-E)\psi_1 &= 0,\\
c(p_x+ip_y)\psi_3-cp_z\psi_4+(mc^2-E)\psi_2 &= 0,\\
c(p_x-ip_y)\psi_2+cp_z\psi_1-(mc^2+E)\psi_3 &= 0,\\
c(p_x+ip_y)\psi_1-cp_z\psi_2-(mc^2+E)\psi_4 &= 0.
\end{aligned} \tag{1.60}$$

Using the Pauli matrices we can write these equations in two-component form for the spinors ψ_u and ψ_v defined as

$$\psi_u=\begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix},\qquad \psi_v=\begin{pmatrix}\psi_3\\ \psi_4\end{pmatrix}. \tag{1.61}$$

$$c\boldsymbol{\sigma}\mathbf{p}\psi_v+(mc^2-E)\psi_u = 0, \tag{1.62}$$

$$c\boldsymbol{\sigma}\mathbf{p}\psi_u-(mc^2+E)\psi_v = 0. \tag{1.63}$$

From (1.62) we have

$$\psi_u=\frac{c\boldsymbol{\sigma}\mathbf{p}}{E-mc^2}\psi_v. \tag{1.64}$$

Let's make an estimation of the relative magnitude of ψ_u and ψ_v in the non-relativistic approximation. From the general relation $(\mathbf{a}\cdot\boldsymbol{\sigma})\cdot(\mathbf{b}\cdot\boldsymbol{\sigma})=\mathbf{a}\cdot\mathbf{b}+i\boldsymbol{\sigma}\cdot(\mathbf{a}\times\mathbf{b})$ we can write $(\boldsymbol{\sigma}\mathbf{p})^2=p^2$, and so $\boldsymbol{\sigma}\mathbf{p}=p\approx mv$. Furthermore[11],

$$\begin{aligned}
E-mc^2&=\sqrt{(mc^2)^2+p^2c^2}-mc^2=mc^2\sqrt{1+\left(\frac{p}{mc}\right)^2}-mc^2\\
&=mc^2\left[1+\frac{1}{2}\left(\frac{p}{mc}\right)^2-\frac{1}{8}\left(\frac{p}{mc}\right)^4+\ldots\right]-mc^2=\frac{p^2}{2m}-\frac{p^4}{8m^3c^2}+\ldots\approx\frac{1}{2}mv^2
\end{aligned} \tag{1.65}$$

[11] Using $\sqrt{1+x^2}=1+\frac{1}{2}x^2-\frac{1}{8}x^4+\ldots$.

Thus

$$\frac{\psi_u}{\psi_v} \approx \frac{cmv}{\frac{1}{2}mv^2} = 2\frac{c}{v}, \tag{1.66}$$

and for $v/c \ll 1$ only the ψ_u component (also known as large component) needs to be considered. Inserting $\psi_v = \psi_u(E - mc^2)/(c\boldsymbol{\sigma}\mathbf{p})$ in (1.63) we get

$$c\boldsymbol{\sigma}\mathbf{p}\psi_u - \frac{(mc^2 + E)(E - mc^2)}{c\boldsymbol{\sigma}\mathbf{p}}\psi_u = 0, \qquad c^2(\boldsymbol{\sigma}\mathbf{p})^2\psi_u = (E + mc^2)(E - mc^2)\psi_u. \tag{1.67}$$

At non-relativistic energies we have $E' \equiv E - mc^2 \approx \frac{1}{2}mv^2 \ll mc^2$, so

$$(E - mc^2)(E + mc^2) = E'(E' + 2mc^2) = 2mc^2E'\left(1 + \frac{E'}{2mc^2}\right) \approx 2mc^2E'. \tag{1.68}$$

Using this, and taking into account that $(\boldsymbol{\sigma}\mathbf{p})^2 = p^2$, equation (1.67) becomes

$$E'\psi_u = \frac{p^2}{2m}\psi_u, \qquad E'\psi_u = -\frac{\hbar^2}{2m}\nabla^2\psi_u. \tag{1.69}$$

This is still a two-component equation, but since both ψ_u components satisfy the relation, we may as well drop the subscript and write generically

$$E'\psi = -\frac{\hbar^2}{2m}\nabla^2\psi, \tag{1.70}$$

which is the time-independent Schrödinger equation for a free particle. Our next job will be to find the non-relativistic approximation of the Dirac equation with an electromagnetic coupling.

1.3 The Electromagnetic Coupling in the Dirac equation.

Let's consider an electromagnetic field given by the vector and scalar potentials $\mathbf{A}$ and ϕ. Remember that the magnetic and electric fields are (in Gaussian units)

$$\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}, \qquad \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} - \boldsymbol{\nabla}\phi. \tag{1.71}$$

At present, the best understanding we have on how to describe the electromagnetic interaction is by requiring the Dirac Lagrangian to be invariant under a certain type of local gauge transformations. It can be seen[12] that this *local gauge invariance principle* leads us from the free-particle Dirac Hamiltonian $H_0 = c\boldsymbol{\alpha}\mathbf{p} + \beta mc^2$ to $H = \boldsymbol{\alpha}(c\mathbf{p} - q\mathbf{A}) + \beta mc^2 + q\phi$, which takes into account the interaction with the electromagnetic field. The Dirac equation for a particle in the presence of an electromagnetic field is then

$$\left(\boldsymbol{\alpha}(c\mathbf{p} - q\mathbf{A}) + \beta mc^2 + q\phi\right)\psi = E\psi. \tag{1.72}$$

In two-component form,

$$\boldsymbol{\sigma}(c\mathbf{p} - q\mathbf{A})\psi_v + (mc^2 + q\phi)\psi_u = E\psi_u, \tag{1.73}$$

$$\boldsymbol{\sigma}(c\mathbf{p} - q\mathbf{A})\psi_u - (mc^2 - q\phi)\psi_v = E\psi_v. \tag{1.74}$$

Let's examine more closely the *large* component. From (1.74),

$$\psi_v(E + mc^2 - q\phi) = \boldsymbol{\sigma}(c\mathbf{p} - q\mathbf{A})\psi_u, \qquad \psi_v = \frac{\boldsymbol{\sigma}(c\mathbf{p} - q\mathbf{A})}{E + mc^2 - q\phi}\psi_u. \tag{1.75}$$

[12] We'll explore a bit this mechanism in problem (1), but with the Schrödinger equation.

Insert now this in (1.73), denoting $E' = E - mc^2$ and $\boldsymbol{\pi} = \mathbf{p} - \frac{q}{c}\mathbf{A}$,

$$\boldsymbol{\sigma}(c\mathbf{p} - q\mathbf{A})\frac{\boldsymbol{\sigma}(c\mathbf{p} - q\mathbf{A})}{E + mc^2 - q\phi}\psi_u + (mc^2 + q\phi)\psi_u = E\psi_u, \tag{1.76}$$

$$(\boldsymbol{\sigma}\boldsymbol{\pi})\frac{c^2}{E' + 2mc^2 - q\phi}(\boldsymbol{\sigma}\boldsymbol{\pi})\,\psi_u = (E - mc^2 - q\phi)\psi_u = (E' - q\phi)\psi_u, \tag{1.77}$$

$$\frac{1}{2m}(\boldsymbol{\sigma}\boldsymbol{\pi})\frac{2mc^2}{E' + 2mc^2 - q\phi}(\boldsymbol{\sigma}\boldsymbol{\pi})\,\psi_u = (E' - q\phi)\psi_u, \tag{1.78}$$

$$(E' - q\phi)\psi_u = \frac{1}{2m}(\boldsymbol{\sigma}\boldsymbol{\pi})K(\boldsymbol{\sigma}\boldsymbol{\pi})\,\psi_u, \tag{1.79}$$

where

$$K = \frac{2mc^2}{E' + 2mc^2 - q\phi} = \frac{1}{1 + \dfrac{E' - q\phi}{2mc^2}}. \tag{1.80}$$

When the kinetic energy E' is small or, to be precise, the combination $E' - q\phi$ is much smaller than $2mc^2$, we can approximate $K \approx 1$ and obtain

$$(E' - q\phi)\psi_u = \frac{1}{2m}(\boldsymbol{\sigma}\boldsymbol{\pi})^2\,\psi_u. \tag{1.81}$$

In problem (6) we will prove the relation

$$(\boldsymbol{\sigma}\boldsymbol{\pi})^2 = \left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)^2 - \frac{q\hbar}{c}\,\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A})\,. \tag{1.82}$$

Inserting this into equation (1.81) we arrive at

$$(E' - q\phi)\,\psi_u = \left[\frac{1}{2m}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)^2 - \frac{q\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A})\right]\,\psi_u. \tag{1.83}$$

$$E'\,\psi_u = \left[\frac{1}{2m}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)^2 - \frac{q\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A}) + q\phi\right]\,\psi_u. \tag{1.84}$$

This is the Dirac equation to order v/c. It is known as the *Pauli equation* because Pauli himself had studied it before, following a different reasoning. It predicts an interaction term between the magnetic field, $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$, and the spin operator of a spin 1/2 particle, $\mathbf{S} = \hbar\boldsymbol{\sigma}/2$.

We need to go to the next higher approximation, to order v^2/c^2, which comes from doing,[13]

$$K = \frac{1}{1 + \dfrac{E' - q\phi}{2mc^2}} \approx 1 - \frac{E' - q\phi}{2mc^2}. \tag{1.85}$$

The procedure to follow is to insert this approximation for K into equation (1.79) and retain only terms up to order v^2/c^2. We skip here the rather tedious derivation (see appendix B.9) which would lead us to

$$\begin{aligned}(E' - q\phi)\,\psi \;=\; \Bigg[&\frac{1}{2m}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)^2 - \frac{q\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A}) \\ &- \frac{p^4}{8m^3c^2} + \frac{\hbar^2 q}{8m^2c^2}\boldsymbol{\nabla}(\boldsymbol{\nabla}\phi) + \frac{\hbar q}{4m^2c^2}\boldsymbol{\sigma}\left[(\boldsymbol{\nabla}\phi)\times\mathbf{p}\right]\Bigg]\,\psi.\end{aligned} \tag{1.86}$$

[13] Remember that $1/(1+x) \approx 1 - x$ for small x.

Now we set $q = -e$, where e is the absolute value of the electron charge,

$$(E' + e\phi)\,\psi = \left[\frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 + \frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A}) - \frac{p^4}{8m^3c^2} - \frac{\hbar^2 e}{8m^2c^2}\boldsymbol{\nabla}^2\phi - \frac{\hbar e}{4m^2c^2}\boldsymbol{\sigma}\left[(\boldsymbol{\nabla}\phi)\times\mathbf{p}\right]\right]\psi. \quad (1.87)$$

This is the Dirac equation for an electron to order v^2/c^2, and it will be the basis to compute the first-order relativistic corrections to the Hydrogen spectrum. Higher-order corrections, however, are not correctly given by relativistic wave equations, and a full quantum field theory treatment would be needed.[14]

Before describing the relative importance of the different terms which appear in this equation, a few words about units. The primary laboratory data in atomic spectroscopy are the wavelengths of electromagnetic radiation emitted by atoms. For a given transition between two states, one does not measure directly the energy of the emitted photon, but its wavelength. This is the reason why it is commong practice to work with inverse of centimeters instead of energy units.[15] From the relation between energy and wavelength for electromagnetic radiation, $1/\lambda = E/hc$, we find, for $E = 1\,\mathrm{eV}$,

$$\frac{1}{\lambda} = \frac{E}{2\pi\hbar c} = \frac{1\,\mathrm{eV}}{2\pi \times 197.3269788(12) \times 10^6\,\mathrm{eV} \times 10^{-13}\,\mathrm{cm}} = 8065.544005\ \mathrm{cm}^{-1}. \quad (1.88)$$

The relative importance of the different terms in equation (1.87) is the following:

- The electric scalar potential energy, $e\phi$, with a size of about $10^5\,\mathrm{cm}^{-1}$, or $\sim 12\,\mathrm{eV}$.
- The term $\frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2$ contains the kinetic energy and the interaction with the vector potential $\mathbf{A}$. It has also a size of about $10^5\mathrm{cm}^{-1}$ and is responsible of, or contributes to, several physical processes such as the absorption, emission and scattering of electromagnetic waves, diamagnetism and the Zeeman effect.
- The interaction of the electron spin magnetic moment with a magnetic field, $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$, is given by $\frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A})$, with an energy of about $1\mathrm{cm}^{-1}$ ($1.2\times10^{-4}\,\mathrm{eV}$).
- The relativistic correction to the kinetic energy, $\frac{p^4}{8m^3c^2}$, with a size of about $0.1\mathrm{cm}^{-1}$ ($1.2\times10^{-5}\,\mathrm{eV}$), comes from the expansion of the relativistic kinetic energy

$$T = E - mc^2 = \sqrt{(mc^2)^2 + p^2c^2} - mc^2 = mc^2\sqrt{1 + \left(\frac{p}{mc}\right)^2} - mc^2 = mc^2\left[1 + \frac{1}{2}\left(\frac{p}{mc}\right)^2 - \frac{1}{8}\left(\frac{p}{mc}\right)^4 + \ldots\right] - mc^2 = \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \ldots$$

[14] As it was pointed out before, relativistic wave equations do not provide a consistent framework to deal with high-energy particle physics problems. It is true that the non-relativistic limit of the Dirac equation in an external electromagnetic field reduces to the Schrödinger equation with a Hamiltonian which contains an expansion in powers of the electron velocity, i.e. relativistic corrections. *A posteriori*, the field theoretical treatment shows that the first-order corrections produced by the Dirac wave equation are correct, in the sense that they coincide with the field theory result. In the language of Feynman diagrams, we would say that the relativistic wave equations reproduce the tree-level diagram result. For more details see section 3.6 of M. Maggiore's textbook *A Modern Introduction to Quantum Field Theory*, Oxford University Press, 2013.

[15] Be careful not to mix the spectroscopic use of cm^{-1} with its use in the natural system of units. For more details see equation (B.6) in appendix (B.1).

- The Darwin term, $\dfrac{\hbar^2 e}{8m^2c^2}\boldsymbol{\nabla}^2\phi$, has no classic analog and produces an energy shift in s-states by less than 0.1cm^{-1}.
- The spin-orbit interaction term is $-\dfrac{\hbar e}{4m^2c^2}\boldsymbol{\sigma}\left[(\boldsymbol{\nabla}\phi)\times\mathbf{p}\right]$. In Hydrogen, this represents an energy correction of about $10^{-5}\,\text{eV}$, but for heavy atoms it can be considerably larger, reaching a a size from 10 to $10^3\,\text{cm}^{-1}$, (0.0012 to 0.12 eV). If ϕ depends only on r the gradient is just $\boldsymbol{\nabla}\phi = \dfrac{d\phi}{dr}\dfrac{\mathbf{r}}{r}$. Using $\mathbf{L} = \mathbf{r}\times\mathbf{p}$ and $\mathbf{S} = \dfrac{\hbar}{2}\boldsymbol{\sigma}$, we obtain the more familiar expression

$$-\frac{\hbar e}{4m^2c^2}\boldsymbol{\sigma}\left[(\boldsymbol{\nabla}\phi)\times\mathbf{p}\right] = -\frac{e}{2m^2c^2}\frac{d\phi}{dr}\frac{1}{r}\mathbf{S}(\mathbf{r}\times\mathbf{p}) = -\frac{e}{2m^2c^2}\frac{d\phi}{dr}\frac{1}{r}\mathbf{LS} \tag{1.89}$$

Finally, a remark should be made concerning the approximation we have used for the K factor,

$$K = \frac{2mc^2}{E' + 2mc^2 - q\phi} \approx 1 - \frac{E' - q\phi}{2mc^2}.$$

When the electric scalar potential ϕ becomes singular, as it happens with the Coulomb potential for $r = 0$, this approximation is not valid and a different approach is needed. We will see how to deal with this problem when computing the hyperfine splitting in chapter 2.

1.4 The hydrogen atom without higher-order corrections.

What follows is a brief account on how to solve the Schrödinger equation for the Hydrogen system because we assume you have already seen this in detail in previous courses. If this is not the case you are advised to read, for example, chapter four of Griffiths's book [7], and work through exercises 9 and 10 at the end of this chapter.

For an electron in a static field without higher-order relativistic corrections equation (1.87) reduces to

$$(E' + e\phi)\,\psi \;=\; \frac{p^2}{2m}\,\psi. \tag{1.90}$$

At non-relativistic energies we have $E' \approx mv^2/2$. Let's write now this energy directly as E instead of E', to simplify notation. Next we replace $\mathbf{p}$ by $-i\hbar\boldsymbol{\nabla}$ and write the electric scalar potential as $e\phi = -V$ to get

$$\mathcal{H}\psi = \left(-\frac{\hbar^2}{2m}\nabla^2 + V\right)\,\psi \;=\; E\psi. \tag{1.91}$$

In spherical coordinates,

$$r^2\left(\frac{\partial^2}{\partial r^2} + \frac{2}{r}\frac{\partial}{\partial r}\right)\psi + \frac{2mr^2}{\hbar^2}(E - V)\psi \;=\; \frac{1}{\hbar^2}L^2\psi, \tag{1.92}$$

where

$$L^2 = -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]. \tag{1.93}$$

Assuming that the potential has only a radial dependence, $V = V(r)$, and that the eigenfunctions can be written as

$$\psi(r,\theta,\varphi) = \frac{1}{r}u(r)Y(\theta,\varphi) = R(r)Y(\theta,\varphi), \tag{1.94}$$

so the Schrödinger equation can be separated into

$$\frac{d^2u(r)}{dr^2} + \frac{2m}{\hbar^2}\big[E - V(r)\big]u(r) \;=\; \frac{\lambda}{r^2}u(r), \tag{1.95}$$

$$L^2Y(\theta,\varphi) \;=\; \hbar^2\lambda Y(\theta,\varphi), \tag{1.96}$$

where λ is a constant. Next, we need to apply some appropriate boundary conditions. It seems sensible to assume that ψ and its first derivatives are everywhere continuous, single-valued and finite. The consequences of imposing these conditions are that

$$\lambda = \ell(\ell+1), \qquad (\ell = 0, 1, 2\ldots), \tag{1.97}$$

and that the functions $Y(\theta,\varphi)$ are the spherical harmonics, $Y_\ell^m(\theta,\varphi)$, with $m = -\ell, -\ell+1, \ldots, +\ell$. The first few of them, for $\ell = 1, 2, 3$, are listed in table (1.1). The equation for $u(r)$ can be written now as

$$\left[\frac{d^2}{dr^2} + \frac{2m}{\hbar^2}[E - V(r)] - \frac{\ell(\ell+1)}{r^2}\right] u(r) = 0. \tag{1.98}$$

Solutions to this equation are searched for taking into account that $V(r) \to 0$ as $r \to \infty$. Asymptotically, as $r \to \infty$, the previous equation reduces to

$$\frac{d^2u(r)}{dr^2} + \frac{2m}{\hbar^2}Eu(r) = 0, \tag{1.99}$$

which has solutions of the form

$$u(r) = e^{\pm ar}, \quad \text{with} \quad a = \sqrt{-\frac{2m}{\hbar^2}E}. \tag{1.100}$$

For $E < 0$ the solution e^{+ar} *is not* acceptable because it grows without limit as $r \to \infty$. On the contrary, the solution e^{-ar} *is* acceptable because goes to zero as r goes to infinity. For $E > 0$ either sign in the exponent would satisfy the boundary conditions, nevertheless we will focus here on the bound states of the atom, that is, solutions with $E < 0$. In this case, according to the asymptotic behavior, we should look for solutions of the form $u(r) = e^{-ar}f(r)$, where $f(r)$ is a function to be determined by the radial equation (1.98) and the boundary conditions.

Consider now an electron of mass m interacting with a nucleus of charge Ze, so that[16] $V(r) = -Ze^2/r$. The radial equation is now

$$\left[\frac{d^2}{dr^2} + \frac{2m}{\hbar^2}\left(E + \frac{Ze^2}{r}\right) - \frac{\ell(\ell+1)}{r^2}\right] u(r) = 0. \tag{1.101}$$

In Rydberg units, where the unit of length is taken to be the Bohr radius, a_0, and the unit of energy the magnitude of the ground state bounding energy of a Hydrogen with infinite nucleus mass, we have the relations[17]

$$r_R = \frac{r}{a_0}, \qquad E_R = \frac{E}{R_\infty}, \qquad \nabla_R^2 = a_0^2\nabla^2, \qquad a_0 = \frac{\hbar^2}{me^2} = \frac{\hbar}{\alpha mc}, \tag{1.102}$$

$$R_\infty = \frac{1}{2}\left(\frac{e^2}{\hbar c}\right)^2 mc^2 = \frac{1}{2}\alpha^2 mc^2 = \frac{\hbar^2}{2ma_0^2} = \frac{e^2}{2a_0}, \qquad \alpha = \frac{e^2}{\hbar c}, \tag{1.103}$$

which, for example, let us do this transformation,

$$\left(-\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{r}\right)\psi = E\psi \quad \longrightarrow \quad \left(-\nabla_R^2 - \frac{2Z}{r_R}\right)\psi = E_R\psi. \tag{1.104}$$

[16] In the SI system $V(r) = -\frac{Ze^2}{4\pi\epsilon_0 r}$. In the Gaussian system of units, the constants are absorbed into a redefinition of the electric charge unit, so we have $V(r) = -\frac{Ze^2}{r}$. To go from one system to the other just make the replacement $\epsilon_0 \to \frac{1}{4\pi}$. Have a look at the appendix B.1 for more information about various systems of units.

[17] In the SI system the fine structure constant is $\alpha = \dfrac{e^2}{4\pi\epsilon_0\hbar c}$, and the Bohr radius reads $a_0 = \dfrac{4\pi\epsilon_0\hbar^2}{e^2 m}$. The expression $a_0 = \dfrac{\hbar}{mc\alpha}$ is valid in any system of units. R_∞ is the Rydberg constant for infinite nuclear mass.

We have also

$$\left[\frac{1}{a_0^2}\frac{d^2}{dr_R^2}+\frac{1}{a_0^2}E_R+\frac{1}{a_0^2}\frac{2Z}{r_R}-\frac{1}{a_0^2}\frac{\ell(\ell+1)}{r_R^2}\right]u(r_R)=0. \tag{1.105}$$

Since in Rydberg units $a_0=1$ by definition, we get (writing r instead of r_R etc. from now on, to shorten the notation)

$$\frac{d^2u(r)}{dr^2}+\left[E+\frac{2Z}{r}-\frac{\ell(\ell+1)}{r^2}\right]u(r)=0. \tag{1.106}$$

Now we substitute $u(r)=e^{-ar}f(r)$ and get

$$\frac{d^2f}{dr^2}-2a\frac{df}{dr}+\left[a^2+E+\frac{2Z}{r}-\frac{\ell(\ell+1)}{r^2}\right]f=0, \tag{1.107}$$

But $a^2=-2mE/\hbar^2=-2mE_RR_\infty/\hbar^2=-E_R/a_0^2$, so when E is expressed in Rydberg units (where $a_0=1$) we have $a^2=-E$ and the previous expression reduces to

$$\frac{d^2f}{dr^2}-2a\frac{df}{dr}+\left[\frac{2Z}{r}-\frac{\ell(\ell+1)}{r^2}\right]f=0. \tag{1.108}$$

Next, we try a power series solution,

$$f=r^s\left(A_0+A_1r+A_2r^2+\ldots\right), \tag{1.109}$$

with $s>0$ to ensure that f remains finite as $r\to 0$. After inserting this power series in the differential equation for $f(r)$ it is found that $s=\ell+1>0$ and that the following recursion relation for the coefficients must hold,[18]

$$\frac{A_k}{A_{k-1}}=2\,\frac{(\ell+k)a-Z}{k^2+(2\ell+1)k}=2\,\frac{\frac{\ell a}{k}+a-\frac{Z}{k}}{k+2\ell+1}. \tag{1.110}$$

We see that, for $k\to\infty$, the recursion relation behaves as $\dfrac{A_k}{A_{k-1}}\to\dfrac{2a}{k}$. This is also what happens to the coefficients of the e^{2ar} power series expansion,

$$e^{2ar}=\sum_{k=0}^{\infty}\frac{(2a)^k}{k!}r^k,\qquad \frac{(2a)^k/k!}{(2a)^{k-1}/(k-1)!}=\frac{2a}{k}. \tag{1.111}$$

This tells us that the function $f(r)$ behaves like e^{2ar} for large k and, therefore, $u(r)=e^{-ar}f(r)\approx e^{-ar}e^{2ar}=e^{ar}$ for large k if $f(r)$ is given by an infinite series expansion. Since this implies that $u(r)\to\infty$ for $r\to\infty$, which is not an acceptable solution, we must demand the series (1.109) to terminate at some finite value of k. This can only be done by setting to zero the numerator of the recursion relation (1.110). Recalling that $a^2=-E$ in Rydberg units,

$$(\ell+k)a-Z=0,\qquad a=\frac{Z}{\ell+k},\qquad E=-\frac{Z^2}{(\ell+k)^2}\equiv-\frac{Z^2}{n^2}. \tag{1.112}$$

In SI units,

$$E_n=-\left[\frac{m}{2\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0}\right)^2\right]\frac{Z^2}{n^2}=-\frac{1}{2}mc^2\left(\frac{e^2}{4\pi\epsilon_0\hbar c}\right)^2\frac{Z^2}{n^2}=-\frac{1}{2}\alpha^2mc^2\frac{Z^2}{n^2}=\frac{E_1}{n^2}, \tag{1.113}$$

[18] To see clearly how a recursion relation like this arises, work carefully through problem number 9 at the end of the chapter.

where E_1 is the binding energy of the hydrogen ground state. Note that the boundary conditions imposed on the differential equations have forced the bound states to have a *discrete* energy spectrum. The acceptable solutions for $u(r)$ can be written in terms of the Laguerre polynomials,

$$L_{n+\ell}^{2\ell+1}(x) = -\left[(n+\ell)!\right]^2 \sum_{k=0}^{n-\ell-1} \frac{(-x)^k}{k!(n-\ell-k-1)!(2\ell+k+1)!} = B_0 + B_1 x + B_2 x^2 + \ldots + B_{n-\ell-1} x^{n-\ell-1}. \quad (1.114)$$

$$B_k = -B_{k-1}\frac{n-\ell-k}{k(2\ell+k+1)}, \qquad B_{n-\ell-1} = (-1)^{n-\ell}\frac{(n+\ell)!}{(n-\ell-1)!} \quad (1.115)$$

To keep the order of the Laguerre polynomial finite, the acceptable values for the integer n are those that make the relation $n-\ell-1$ to be equal to zero or to some positive integer, therefore we have

$$n = \ell+1, \ell+2, \ell+3, \ldots \quad (1.116)$$

The full solution for the hydrogen bound states is,

$$\boxed{|n\ell m\rangle \equiv \psi_{n\ell m}(r,\theta,\varphi) = R_{n\ell}(r)Y_\ell^m(\theta,\varphi) = \frac{1}{r}u_{n\ell}(r)Y_\ell^m(\theta,\varphi)} \quad (1.117)$$

where,

$$u_{n\ell}(r) = \sqrt{\frac{(n-\ell-1)!Z}{n^2[(n+\ell)!]^3 a_0}}\left(\frac{2Zr}{na_0}\right)^{\ell+1} e^{-Zr/(na_0)}\, L_{n+\ell}^{2\ell+1}\left(\frac{2Zr}{na_0}\right), \quad (1.118)$$

and

$$\int_0^\infty u_{n\ell}(r)u_{n'\ell}(r)\,\mathrm{d}r = \int_0^\infty R_{n\ell}(r)R_{n'\ell}(r)\,r^2\,\mathrm{d}r = \delta_{nn'}. \quad (1.119)$$

The orthogonality properties of $u_{n\ell}(r)$ and $Y_\ell^m(\theta,\varphi)$ lead to the following orthogonality relations for the Hydrogen eigenfunctions.

$$\langle n\ell m|n'\ell'm'\rangle = \int \psi^*_{n\ell m}(r,\theta,\varphi)\psi_{n'\ell'm'}(r,\theta,\varphi)\,\mathrm{d}\tau = \delta_{nn'}\delta_{\ell\ell'}\delta_{mm'}. \quad (1.120)$$

We have three quantum numbers, n, ℓ and m, where n is the principal quantum number taking the values $n = 1, 2, 3, \ldots$. The orbital angular momentum quantum number takes the values $\ell = 0, 1, 2, \ldots (n-1)$. The magnetic (or projection) quantum number takes the values $m = -\ell, \ell+1, \ldots, 0, \ldots \ell-1, +\ell$. In spectroscopic notation we have $s, p, d, f, g \ldots$ instead of $\ell = 0, 1, 2, 3, 4, \ldots$ Note that the degeneracy of the Hydrogen eigenstates[19] is n^2,

$$\sum_{\ell=0}^{n-1}(2\ell+1) = n^2 \quad (1.121)$$

Table (1.2) shows the hydrogen-like atom wave functions (all written with a positive phase factor) for states with $n = 1$ and $n = 2$, together with the radial functions $u_{n\ell}(r) = rR_{n\ell}(r)$.

It is interesting to see some plots of $R_{n\ell}(r) = \frac{1}{r}u_{n\ell}(r)$. The higher the value of n, the greater is the number of nodes in $R_{n\ell}(r)$ for a given value of ℓ. The probability of finding an electron in an element of volume $\mathrm{d}\tau$ is

$$\psi^*\psi\,\mathrm{d}\tau = \frac{1}{r^2}u_{n\ell}^2(r)Y_\ell^{m*}(\theta,\varphi)Y_\ell^m(\theta,\varphi)\,r^2\sin\theta\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\varphi. \quad (1.122)$$

[19] For a Hamiltonian invariant under three-dimensional rotations, as it is the case when $V = V(r)$, one in general expects eigenfunctions of the form $R(r)Y_\ell^m(\theta,\varphi)$ with a degeneracy of $(2\ell+1)$. The fact that for Hydrogen we have a degeneracy of n^2 is due to the special property of the Coulomb field. This degeneracy is often called accidental. For multi-electron atoms, the energy of the orbitals depends on both n and ℓ because $V(r)$ is no longer the Coulomb potential.

ℓ, m	$Y_\ell^m(\theta,\varphi)$	ℓ, m	$Y_\ell^m(\theta,\varphi)$
$0, 0$	$\left(\frac{1}{4\pi}\right)^{1/2}$	$2, \pm 2$	$\left(\frac{15}{32\pi}\right)^{1/2} \sin^2\theta\, e^{\pm i2\varphi}$
$1, 0$	$\left(\frac{3}{4\pi}\right)^{1/2} \cos\theta$	$3, 0$	$\left(\frac{7}{16\pi}\right)^{1/2} (5\cos^3\theta - 3\cos\theta)$
$1, \pm 1$	$\mp\left(\frac{3}{8\pi}\right)^{1/2} \sin\theta\, e^{\pm i\varphi}$	$3, \pm 1$	$\mp\left(\frac{21}{64\pi}\right)^{1/2} \sin\theta\,(5\cos^2\theta - 1)\, e^{\pm i\varphi}$
$2, 0$	$\left(\frac{5}{16\pi}\right)^{1/2} (3\cos^2\theta - 1)$	$3, \pm 2$	$\left(\frac{105}{32\pi}\right)^{1/2} \sin^2\theta\, \cos\theta\, e^{\pm i2\varphi}$
$2, \pm 1$	$\mp\left(\frac{15}{8\pi}\right)^{1/2} \sin\theta\cos\theta\, e^{\pm i\varphi}$	$3, \pm 3$	$\mp\left(\frac{35}{64\pi}\right)^{1/2} \sin^3\theta\, e^{\pm i3\varphi}$

Table 1.1: Spherical harmonics for $\ell = 1, 2, 3$.

n, ℓ, m	$\psi_{n\ell m}(r,\theta,\varphi)$	$u_{n\ell}(r)$
$1, 0, 0$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{1}{\sqrt{\pi}}\, e^{-Zr/a_0}$	$\left(\frac{Z}{a_0}\right)^{3/2} 2r\, e^{-Zr/a_0}$
$2, 0, 0$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{1}{\sqrt{32\pi}} \left(2 - \frac{Zr}{a_0}\right) e^{-Zr/2a_0}$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{r}{\sqrt{2}} \left(1 - \frac{Zr}{2a_0}\right) e^{-Zr/2a_0}$
$2, 1, 0$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{1}{\sqrt{32\pi}} \frac{Zr}{a_0} e^{-Zr/2a_0} \cos\theta$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{1}{2\sqrt{6}} \frac{Zr^2}{a_0} e^{-Zr/2a_0}$
$2, 1, \pm 1$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{1}{\sqrt{64\pi}} \frac{Zr}{a_0} e^{-Zr/2a_0} \sin\theta\, e^{\pm i\varphi}$	
$3, 0, 0$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{1}{81\sqrt{3\pi}} \left[27 - 18\frac{Zr}{a_0} + 2\left(\frac{Zr}{a_0}\right)^2\right] e^{-Zr/3a_0}$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{2r}{3\sqrt{3}} \left[1 - \frac{2Zr}{3a_0} + \frac{2}{27}\left(\frac{Zr}{a_0}\right)^2\right] e^{-Zr/3a_0}$
$3, 1, 0$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{1}{81}\sqrt{\frac{2}{\pi}} \left[6\frac{Zr}{a_0} - \left(\frac{Zr}{a_0}\right)^2\right] e^{-Zr/3a_0} \cos\theta$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{8}{27\sqrt{6}} \frac{Zr^2}{a_0} \left(1 - \frac{Zr}{6a_0}\right) e^{-Zr/3a_0}$
$3, 1, \pm 1$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{1}{81\sqrt{\pi}} \left[6\frac{Zr}{a_0} - \left(\frac{Zr}{a_0}\right)^2\right] e^{-Zr/3a_0} \sin\theta\, e^{\pm i\varphi}$	
$3, 2, 0$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{1}{81\sqrt{6\pi}} \left(\frac{Zr}{a_0}\right)^2 e^{-Zr/3a_0} (3\cos^2\theta - 1)$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{4}{81\sqrt{30}} \frac{Z^2r^3}{a_0^2}\, e^{-Zr/3a_0}$
$3, 2, \pm 1$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{1}{81\sqrt{\pi}} \left(\frac{Zr}{a_0}\right)^2 e^{-Zr/3a_0} \sin\theta\, \cos\theta\, e^{\pm i\varphi}$	
$3, 2, \pm 2$	$\left(\frac{Z}{a_0}\right)^{3/2} \frac{1}{162\sqrt{\pi}} \left(\frac{Zr}{a_0}\right)^2 e^{-Zr/3a_0} \sin^2\theta\, e^{\pm i2\varphi}$	

Table 1.2: Hydrogen-like wave functions (all written with a positive phase factor) and radial wave functions for states with $n = 1, 2, 3$.

If $\psi^*\psi\, \mathrm{d}\tau$ is integrated over the surface of a sphere you get the probability of finding the electron in a shell between two spheres of radii r and $r + \mathrm{d}r$. Since the spherical harmonics are normalized to unity the result is just $r^2 R_{n\ell}^2\, \mathrm{d}r = u_{n\ell}^2(r)\, \mathrm{d}r$. Therefore $u_{n\ell}^2(r)$ is a radial charge density in the same sense that $\psi^*\psi$ is a charge density. The probability of finding the electron between θ and $\theta + \mathrm{d}\theta$ is proportional to

$$Y_\ell^{m*}(\theta,\varphi) Y_\ell^m(\theta,\varphi) \sin\theta\, \mathrm{d}\theta = \left[P_\ell^{|m|}(\cos\theta)\right]^2 \sin\theta\, \mathrm{d}\theta. \tag{1.123}$$

The probability of finding the electron between φ and $\varphi + \mathrm{d}\varphi$ is just proportional to $\mathrm{d}\varphi$.

The orbital angular momentum operator, $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, commutes with the Laplace operator, ∇^2, therefore the Hamiltonian $\mathcal{H} = -(\hbar^2/2m)\nabla^2 + V(r)$ also commutes with $\mathbf{L}$ and L^2. The components of $\mathbf{L}$ do not commute with one another, but any one of them commutes with L^2 and $\mathcal{H}$. It is conventional to write the Hydrogen bound states as eigenstates of the operators L_z, L^2 and $\mathcal{H}$, which satisfy

$$[L^2, \mathcal{H}] = [L_z, \mathcal{H}] = [L^2, L_z] = 0. \tag{1.124}$$

Since our Hamiltonian does not contain spin operators, the following relations also hold,

$$[S^2, \mathcal{H}] = [S_z, \mathcal{H}] = 0. \tag{1.125}$$

Going back to equation (1.100), we see that if $E > 0$ the parameter $a = \sqrt{-(2m/\hbar^2)E}$ becomes a pure imaginary number, which may be written as $a = \pm i\kappa$. The radial function $u(r) = e^{\pm i\kappa r}$ remains now finite for $r \to \infty$ and it is not necessary (in fact it is not possible) to terminate the series in (1.109). As a consequence, the energy is not quantized and can take any value from zero to infinity, giving rise to a continuum set of eigenstates. Each positive eigenvalue is infinitely degenerate: to each value of E there corresponds an infinite number of states, with ℓ taking all integral values from 0 to ∞ (with all possible values of m for a given ℓ). The hydrogen eigenstates for $E > 0$ can still be written as the product of a radial function and the spherical harmonics,

$$\psi_{k\ell m}(r, \theta, \varphi) = R_{k\ell}(r) Y_\ell^m(\theta, \varphi), \tag{1.126}$$

where $k = \sqrt{2mE/\hbar^2}$ is a continuous parameter (wave number) that may take any positive value. The radial function $R_{k\ell}(r)$ can be written as the product of an exponential function and a confluent hypergeometric function. We don't need to provide here an explicit expression,[20] but they satisfy the following conditions of normalization and orthogonality

$$\int_0^\infty r^2 R_{k'\ell} R_{k\ell} \, \mathrm{d}r \;=\; \delta\left(\frac{k'-k}{2\pi}\right) \;=\; 2\pi\, \delta(k'-k). \tag{1.127}$$

If $R_{n\ell}(r)$ is a radial function of the discrete spectrum and $R_{k\ell}(r)$ a radial function of the continuum spectrum, the following orthogonality relation holds[21]

$$\int_0^\infty r^2 R_{n\ell} R_{k\ell} \, \mathrm{d}r \;=\; 0. \tag{1.128}$$

The orthogonality of the full eigenstates of the continuum spectrum for different values of ℓ and m is ensured by the spherical harmonics,

$$\langle k\ell m | k'\ell' m' \rangle = \int \psi^*_{k\ell m}(r, \theta, \varphi) \psi_{k'\ell' m'}(r, \theta, \varphi) \, \mathrm{d}\tau = 2\pi\, \delta_{\ell\ell'}\, \delta_{mm'}\, \delta(k'-k). \tag{1.129}$$

It should be emphasized that, if we wish to expand a function in terms of the Hydrogen atom solutions, it is necessary to include both the continuum states and the bound discrete states. The bound states form an infinite discrete set, but not a complete one.

1.4.1 Expectation values.

The expectation value of a given power of r is defined as

$$\left\langle r^k \right\rangle = \int r^k \psi^* \psi \, \mathrm{d}\tau = \int_0^\infty r^k R_{n\ell}^2(r)\, r^2 \, \mathrm{d}r = \int_0^\infty r^k u_{n\ell}^2(r) \, \mathrm{d}r. \tag{1.130}$$

The expectation value of r itself is found to be

$$\langle r \rangle_{n\ell} = \frac{a_0}{2Z} \left[3n^2 - \ell(\ell+1)\right]. \tag{1.131}$$

This may be interpreted as the *size* of the atom and is inversely proportional to Z.[22] For $\ell = 0$ states the value of $\langle r \rangle_{n\ell}$ is proportional to n^2, the deviations from this proportionality being

[20] The interested reader may consult [11] (chapter 5) or [1] (page 21), for example.

[21] Remember the theorem which states that eigenfunctions corresponding to different eigenvalues are orthogonal.

[22] Of course, this is only true for hydrogenic atoms with nuclear charge Ze. The radii of multi-electron atoms with high Z do not decrease as $1/Z$ because the charge of the nucleus is greatly screened by the inner electrons, so the outer electrons do not see the full Ze charge. Heavy atoms are slightly larger than light ones because the electrons in the outer orbits have larger values of the principal quantum number n, but for the same n the atomic radius decreases as Z grows. In some contexts, the size of the atom is defined as one-half the distance between the nuclei of identical atoms that are bonded together. For the H_2 molecule this is 0.74×10^{-10} m, which would translate to a Hydrogen radius of about 0.37×10^{-10} m (the Bohr radius is $a_0 = 0.529 \times 10^{-10}$ m).

small for states with $\ell \neq 0$. Note that for the ground state of the Hydrogen atom we have $\langle r \rangle = 3a_0/2$, while the maximum value of the probability density function occurs at a value of r given by $r_{\max} = a_0$. For states with $\ell = n-1$ the probability density function has a single maximum located at $r_{\max} = a_0 n^2/Z$. The expectation value of r^2 is

$$\langle r^2 \rangle_{n\ell} = \frac{a_0^2}{Z^2}\frac{n^2}{2}\left[5n^2 + 1 - 3\ell(\ell+1)\right]. \tag{1.132}$$

We collect below the expressions for the expectation values of a few inverse powers of r that will be used later to compute several corrections.

$$\left\langle \frac{1}{r} \right\rangle_{n\ell} = \int_0^\infty \frac{1}{r} u_{n\ell}^2(r)\, dr = \frac{Z}{a_0}\frac{1}{n^2}, \qquad \left\langle \frac{1}{r^2} \right\rangle_{n\ell} = \frac{Z^2}{a_0^2}\frac{1}{n^3(\ell+\frac{1}{2})}. \tag{1.133}$$

$$\left\langle \frac{1}{r^3} \right\rangle_{n\ell} = \frac{Z^3}{a_0^3}\frac{1}{n^3(\ell+1)(\ell+\frac{1}{2})\ell}. \tag{1.134}$$

$$\left\langle \frac{1}{r^4} \right\rangle_{n\ell} = \frac{Z^4}{a_0^4}\frac{3n^2 - \ell(\ell+1)}{2n^5(\ell+\frac{3}{2})(\ell+1)(\ell+\frac{1}{2})\ell(\ell-\frac{1}{2})}. \tag{1.135}$$

1.4.2 One-electron wave functions.

We have just seen that for the Hamiltonian $\mathcal{H} = \dfrac{p^2}{2m} - \dfrac{Ze^2}{r}$ the complete solution to the time-independent Schrödinger equation $\mathcal{H}\psi = E\psi$ for the bound states consists of the orbitals $\psi_{nlm}(r,\theta,\varphi)$ together with the energy eigenvalues E_n. The Dirac equation to order v^2/c^2 (1.87) contains additional terms which modify both the ψ_{nlm} eigenfunctions and the E_n eigenvalues. We will evaluate later the corrections to E_n using time-independent perturbation theory with the ψ_{nlm} solutions as a zero-order basis set.

The Hamiltonian of the free-particle Schrödinger equation, $\mathcal{H} = p^2/2m$, commutes with the orbital angular momentum, $\mathbf{L}$, but the Hamiltonian of the free-particle Dirac equation does not:

$$[\mathcal{H}_{\rm D}, \mathbf{L}] \neq 0, \qquad \mathcal{H}_{\rm D} = c\boldsymbol{\alpha}\mathbf{p} + \beta mc^2, \tag{1.136}$$

where $\beta = \gamma^0$ and $\alpha_k = \gamma^0\gamma^k$, $(k = 1,2,3)$. It does commute, however, with the total angular momentum operator $\mathbf{J} = \mathbf{L} + \mathbf{S}$, where the spin operator is defined as[23]

$$\mathbf{S} = \frac{\hbar}{2}\begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \equiv \frac{\hbar}{2}\boldsymbol{\Sigma}, \qquad S_i = \frac{\hbar}{2}\begin{pmatrix} \sigma_i & 0 \\ 0 & \sigma_i \end{pmatrix}, \quad i = 1,2,3 \tag{1.137}$$

$$\mathbf{S}^2 = \frac{\hbar^2}{4}\begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_1^2 \end{pmatrix} + \frac{\hbar^2}{4}\begin{pmatrix} \sigma_2^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} + \frac{\hbar^2}{4}\begin{pmatrix} \sigma_3^2 & 0 \\ 0 & \sigma_3^2 \end{pmatrix} = \frac{3\hbar^2}{4}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{1.138}$$

It is easy to verify that the components of $\mathbf{S}$ satisfy the same commutation relations as the orbital angular momentum, $[S_i, S_j] = i\hbar S_k$, with (i,j,k) an even permutation of $(1,2,3)$. A particle at rest ($\mathbf{L} = 0$) satisfying the Dirac equation is described by a spinor which verifies the eigenvalue equation

$$\mathbf{S}^2\chi = \hbar^2 S(S+1)\chi = \hbar^2\frac{3}{4}\chi,$$

and thus has intrinsic angular momentum $S = 1/2$. The ability to *automatically* incorporate in such a way the fact that the electron has spin 1/2 was historically one of the great accomplishments of the Dirac equation.

[23] This is the spin operator which acts on the *bispinor* states (4 components) which are solutions of the Dirac equation. The spin operator which acts on the spinor states (2 components) is just $\hbar\boldsymbol{\sigma}/2$.

We may adopt now the notation $\phi(\boldsymbol{r}) \equiv \phi_{n\ell m_\ell}(r,\theta,\varphi)$ for a one-electron spatial orbital characterized by the quantum numbers n, ℓ and m_ℓ, instead of the previous notation $\psi_{n\ell m}(r,\theta,\varphi)$. For the one-electron spin function we may write,

$$\chi_{m_s} \equiv \chi(m_s) = \begin{cases} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = |\tfrac{1}{2}, +\tfrac{1}{2}\rangle\,, & \text{for } m_s = +\tfrac{1}{2} \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |\tfrac{1}{2}, -\tfrac{1}{2}\rangle\,, & \text{for } m_s = -\tfrac{1}{2} \end{cases} \tag{1.139}$$

The complete one-electron orbital would be

$$\psi(\lambda) = \psi(r,\theta,\varphi,m_s) = \phi(\boldsymbol{r})\chi(m_s). \tag{1.140}$$

Whenever we write an integral involving spin orbitals, a spatial integration, as well as a summation over spin coordinates, is to be understood. It will often be the case that the radial part of the wave function remains fixed, then the principal quantum number (n) can be suppressed without loss of information. For many purposes, the Dirac notation turns out to be the most convenient one. The notation $|\ell s m_\ell m_s\rangle$ stands for the angular and spin parts of the wave function. The subscripts in m_ℓ and m_s are used to distinguish the projection quantum numbers associated to ℓ and to s. The spin orbitals $\psi(\lambda) = |n\ell s m_\ell m_s\rangle$ are still eigenfunctions of $\mathcal{H} = \dfrac{p^2}{2m} - \dfrac{Ze^2}{r}$, with the same energy eigenvalues associated to $\phi_{n\ell m_\ell}(r,\theta,\varphi)$. However, the degeneracy has been increased from n^2 to $2n^2$. We have the following relations,

$$\begin{aligned} L^2\,|\ell s m_\ell m_s\rangle &= \ell(\ell+1)\hbar^2\,|\ell s m_\ell m_s\rangle\,, \\ S^2\,|\ell s m_\ell m_s\rangle &= s(s+1)\hbar^2\,|\ell s m_\ell m_s\rangle = \frac{3}{4}\hbar^2\,|\ell s m_\ell m_s\rangle\,, \\ L_z\,|\ell s m_\ell m_s\rangle &= m_\ell\hbar\,|\ell s m_\ell m_s\rangle\,, \\ S_z\,|\ell s m_\ell m_s\rangle &= m_s\hbar\,|\ell s m_\ell m_s\rangle = \pm\frac{1}{2}\hbar\,|\ell s m_\ell m_s\rangle\,. \end{aligned} \tag{1.141}$$

We are free to make any linear combination of degenerate eigenfunctions to build other sets of degenerate eigenfunctions. A specially important example is

$$|\ell s j m\rangle = \sum_{m_\ell m_s} |\ell s m_\ell m_s\rangle\,\langle \ell s m_\ell m_s|\ell s j m\rangle\,, \tag{1.142}$$

where $\langle \ell s m_\ell m_s|\ell s j m\rangle$ are the Clebsh-Gordan coefficients for the coupling of $\mathbf{L}$ and $\mathbf{S}$ to obtain the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$. The $|\ell s j m\rangle$ eigenstates satisfy

$$\begin{aligned} L^2\,|\ell s j m\rangle &= \ell(\ell+1)\hbar^2\,|\ell s j m\rangle\,, \\ S^2\,|\ell s j m\rangle &= s(s+1)\hbar^2\,|\ell s j m\rangle = \frac{3}{4}\hbar^2\,|\ell s j m\rangle\,, \\ J^2\,|\ell s j m\rangle &= j(j+1)\hbar^2\,|\ell s j m\rangle\,, \\ J_z\,|\ell s j m\rangle &= m\hbar\,|\ell s j m\rangle\,. \end{aligned} \tag{1.143}$$

And,

$$j = \ell - \frac{1}{2}, \ell + \frac{1}{2}, \qquad m = m_\ell + m_s \tag{1.144}$$

The $|\ell s j m\rangle$ states are not eigenfunctions of L_z or S_z. They are said to be in the coupled basis, in contrast to the $|\ell s m_\ell m_s\rangle$ states, which are in the uncoupled basis. In the spectroscopic notation for the orbitals, the small case letters s, p, d, f, ... stand for $\ell = 0, 1, 2, 3, \ldots$ in the $|\ell s m_\ell m_s\rangle$ basis. In the coupled basis $|\ell s j m\rangle$ (also called terms) we use upper case letters, S, P, D, F, ...

to stand for $\ell = 0, 1, 2, 3, \ldots$, and the spin s is indicated by writing $2s+1$ as a superscript to the left. For example, the expression

$$^{2s+1}\mathrm{L}_j \;=\; {}^2\mathrm{D}_{3/2} \tag{1.145}$$

means $\ell = 2$ and $2s+1 = 2$, so $s = 1/2$.[24] The value of the total angular momentum, $j = 3/2$, is written as a subscript, and m may be indicated separately as follows

$$|\ell s j m\rangle = \left|2\,\frac{1}{2}\,\frac{3}{2}\,-\frac{1}{2}\right\rangle = \left|{}^2\mathrm{D}_{3/2}\,-\frac{1}{2}\right\rangle. \tag{1.146}$$

Of course, the choice of basis for the states does not affect the physics results.

In the coupled basis the states cannot be written as a product of a single space-wavefunction and a constant spinor, like we did in equation (1.140) for the uncoupled basis. This is because the angular part of the wavefunction turns out to be different for the two components of the spinor. Indeed, equation (1.142) tells us that the state $|\ell s j m\rangle$ is built as a linear combination of $|\ell s m_\ell m_s\rangle$ states with different values of m_ℓ, hence different spherical harmonics.

It is instructive to see this difference by finding the explicit form of the one-electron wavefunctions which are simultaneously eigenstates of L^2, S^2, J^2 and J_z. To be very clear, we can write the total angular momentum as

$$\mathbf{J} = \mathbf{L}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\hbar}{2}\boldsymbol{\sigma}. \tag{1.147}$$

The J_z operator is then

$$J_z = -i\hbar\begin{pmatrix} \frac{\partial}{\partial\varphi} & 0 \\ 0 & \frac{\partial}{\partial\varphi} \end{pmatrix} + \frac{\hbar}{2}\sigma_3 \;=\; \hbar\begin{pmatrix} -i\frac{\partial}{\partial\varphi} + \frac{1}{2} & 0 \\ 0 & -i\frac{\partial}{\partial\varphi} - \frac{1}{2} \end{pmatrix}. \tag{1.148}$$

From this expression, it is easy to see that the eigenfunction of J_z corresponding to the eigenvalue $m\hbar$ can be written as

$$\Psi = \begin{pmatrix} R_1(r)F_1(\theta)\, e^{i(m-\frac{1}{2})\varphi} \\ R_2(r)F_2(\theta)\, e^{i(m+\frac{1}{2})\varphi} \end{pmatrix}, \tag{1.149}$$

where R_1, R_2, F_1, F_2, are radial and angular functions to be determined. For Ψ to be an eigenfunction of L^2 we need to require that $L^2\Psi = \ell(\ell+1)\hbar^2\Psi$. In full detail,

$$\begin{aligned} L^2 R_1(r)F_1(\theta)\, e^{i(m-\frac{1}{2})\varphi} &= \ell(\ell+1)\hbar^2 R_1(r)F_1(\theta)\, e^{i(m-\frac{1}{2})\varphi}, \\ L^2 R_2(r)F_2(\theta)\, e^{i(m+\frac{1}{2})\varphi} &= \ell(\ell+1)\hbar^2 R_2(r)F_2(\theta)\, e^{i(m+\frac{1}{2})\varphi}. \end{aligned} \tag{1.150}$$

Since we know that $L^2 Y_\ell^m = \ell(\ell+1)\hbar^2\, Y_\ell^m$, we conclude

$$F_1(\theta)\, e^{i(m-\frac{1}{2})\varphi} = Y_\ell^{m-\frac{1}{2}}(\theta,\varphi), \qquad F_2(\theta)\, e^{i(m+\frac{1}{2})\varphi} = Y_\ell^{m+\frac{1}{2}}(\theta,\varphi),$$

Therefore

$$\Psi = \begin{pmatrix} R_1(r)\, Y_\ell^{m-\frac{1}{2}}(\theta,\varphi) \\ R_2(r)\, Y_\ell^{m+\frac{1}{2}}(\theta,\varphi) \end{pmatrix}. \tag{1.151}$$

Finally, we need to add the requirement that Ψ is an eigenfunction of $J^2 = L^2 + S^2 + 2\mathbf{LS}$. When acting on Ψ, the eigenvalues of L^2 and S^2 are $\ell(\ell+1)\hbar^2$ and $3\hbar^2/4$, respectively. Therefore we can write

$$J^2 = \ell(\ell+1)\hbar^2 + \frac{3}{4}\hbar^2 + \hbar\mathbf{L}\boldsymbol{\sigma} = \hbar^2\begin{pmatrix} \ell(\ell+1) + \frac{3}{4} + \frac{1}{\hbar}L_z & \frac{1}{\hbar}L_- \\ \frac{1}{\hbar}L_+ & \ell(\ell+1) + \frac{3}{4} - \frac{1}{\hbar}L_z \end{pmatrix}, \tag{1.152}$$

[24] For one-electron systems we always have $s = 1/2$ and $2s+1 = 2$, but this notation will be very convenient when applied to multi-electron systems.

where

$$L_\pm Y_\ell^m = (L_x \pm iL_y)Y_\ell^m = \hbar\sqrt{\ell(\ell+1) - m(m\pm 1)}\, Y_\ell^{m\pm 1}. \tag{1.153}$$

Imposing now the condition $J^2\Psi = j(j+1)\hbar^2\,\Psi$, leads to a system of two equations which can be solved to obtain the following normalized eigenfunctions (see problem 8),

$$\begin{aligned}
\Psi_{n,\ell,j=\ell+\frac{1}{2},m} &= \frac{R_{n,j-\frac{1}{2}}(r)}{\sqrt{2\ell+1}} \begin{pmatrix} \sqrt{\ell+m+\frac{1}{2}}\, Y_\ell^{m-\frac{1}{2}} \\ \sqrt{\ell-m+\frac{1}{2}}\, Y_\ell^{m+\frac{1}{2}} \end{pmatrix}, \qquad \ell = j-\frac{1}{2}, \quad \ell \neq 0 \\
\Psi_{n,\ell,j=\ell-\frac{1}{2},m} &= \frac{R_{n,j+\frac{1}{2}}(r)}{\sqrt{2\ell+1}} \begin{pmatrix} \sqrt{\ell-m+\frac{1}{2}}\, Y_\ell^{m-\frac{1}{2}} \\ -\sqrt{\ell+m+\frac{1}{2}}\, Y_\ell^{m+\frac{1}{2}} \end{pmatrix}, \qquad \ell = j+\frac{1}{2}, \quad \ell \neq 0
\end{aligned} \tag{1.154}$$

Of course, the radial functions in the expressions above are the usual hydrogen-like radial functions. Also, be careful to note that the m quantum number is the magnetic quantum number of the total angular momentum, $m \equiv m_j = m_\ell + m_s$. Thus, the spherical harmonic index in the upper components of the spinors tells us that

$$m_\ell = m_j - \frac{1}{2} = m_\ell + m_s - \frac{1}{2} \quad \Longrightarrow \quad m_s = \frac{1}{2}.$$

While for the lower components we have

$$m_\ell = m_j + \frac{1}{2} = m_\ell + m_s + \frac{1}{2} \quad \Longrightarrow \quad m_s = -\frac{1}{2}.$$

The case $\ell = 0$ must be treated separately. If $\ell = 0$ we only have the spin part and the J_z operator becomes

$$J_z = \frac{\hbar}{2}\sigma_3 \;=\; \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The eigenfunctions of L^2, S^2, J^2 and J_z can then be written for $\ell = 0$ as

$$\Psi_{n,\ell=0,j=\frac{1}{2},m} = R_{n,0}(r)\, Y_0^0(\theta,\varphi) \begin{pmatrix} m+\frac{1}{2} \\ \frac{1}{2}-m \end{pmatrix}, \qquad j = \frac{1}{2}, \quad \ell = 0. \tag{1.155}$$

Note that in this case there is no angular dependence inside the spinor and the state is written as a separate product of a space-wavefunction and a constant spinor. The states for the two possible values of the magnetic quantum number m are

$$\begin{aligned}
\Psi_{n,\ell=0,j=\frac{1}{2},m=\frac{1}{2}} &= R_{n,0}(r)\, Y_0^0(\theta,\varphi) \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\
\Psi_{n,\ell=0,j=\frac{1}{2},m=-\frac{1}{2}} &= R_{n,0}(r)\, Y_0^0(\theta,\varphi) \begin{pmatrix} 0 \\ 1 \end{pmatrix}.
\end{aligned} \tag{1.156}$$

Equations (1.154) can also be written in terms of abstract Hilbert states without reference to a particular representation (space or momentum) as follows

$$\begin{aligned}
\left|n,\ell,j=\ell+\frac{1}{2},m\right\rangle &= \frac{\sqrt{\ell+m+\frac{1}{2}}}{\sqrt{2\ell+1}} \left|n,\ell,m_\ell = m-\frac{1}{2}, m_s = \frac{1}{2}\right\rangle \\
&\quad + \frac{\sqrt{\ell-m+\frac{1}{2}}}{\sqrt{2\ell+1}} \left|n,\ell,m_\ell = m+\frac{1}{2}, m_s = -\frac{1}{2}\right\rangle, \\
\left|n,\ell,j=\ell-\frac{1}{2},m\right\rangle &= \frac{\sqrt{\ell-m+\frac{1}{2}}}{\sqrt{2\ell+1}} \left|n,\ell,m_\ell = m-\frac{1}{2}, m_s = \frac{1}{2}\right\rangle \\
&\quad - \frac{\sqrt{\ell+m+\frac{1}{2}}}{\sqrt{2\ell+1}} \left|n,\ell,m_\ell = m+\frac{1}{2}, m_s = -\frac{1}{2}\right\rangle.
\end{aligned} \tag{1.157}$$

1.5 Spin-Orbit coupling.

In the rest frame of the electron the moving body is the nucleus. This moving nuclear electric charge creates a magnetic field at the electron position which interacts with the intrinsic magnetic moment of the electron (originating from its spin). From a classical point of view, this would give rise to a torque that tends to align the electron spin along the direction of the field. Alternatively, you may prefer to analyze the problem in the rest frame of the nucleus. In this case, of course, there is no magnetic field due to the movement of the nucleus and you could argue that the electron only *sees* the nucleus Coulombic electric field. This is correct (ignoring the nucleus intrinsic magnetic moment), but electrodynamics tells us that a moving magnetic dipole (the electron) acquires an *electric* dipole moment, and in the nucleus rest frame the spin-orbit coupling is due to the interaction of the *electric* field of the nucleus with the *electric* dipole moment of the moving electron.

If the electrostatic potential depends only on r, from equation (1.89) we read off the following spin-orbit Hamiltonian,

$$\mathcal{H}_{\rm so} = -\frac{e}{2m^2c^2}\frac{1}{r}\frac{d\phi}{dr}\mathbf{LS} = \xi(r)\mathbf{LS}. \tag{1.158}$$

For hydrogen-like atoms $\phi = Ze/r$, therefore

$$\xi(r) = \frac{Ze^2}{2m^2c^2r^3}. \tag{1.159}$$

The one-electron Hamiltonian with spin-orbit coupling becomes then $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\rm so}$, where

$$\mathcal{H}_0 = -\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{r}, \qquad \mathcal{H}_{\rm so} = \xi(r)\mathbf{LS}. \tag{1.160}$$

We will treat the term $\mathcal{H}_{\rm so}$ as a small perturbation.[25] If the hydrogenic states $|n\ell sjm\rangle$ were non-degenerate, the first-order correction to the energy obtained from time-independent perturbation theory would be

$$E_n^{(1)} = \left\langle \psi_n^{(0)} \middle| \mathcal{H}_{\rm so} \middle| \psi_n^{(0)} \right\rangle. \tag{1.161}$$

However, one has to be careful because $\mathcal{H}_0$ has degenerate eigenvalues. There are a number of eigenstates $\psi_{ni}^{(0)}$ with the same eigenvalue, so in this case the first-order corrections to the energy are obtained from the solutions of the secular equation,

$$\left| H_{ij}^{\rm so} - E_n^{(1)}\delta_{ij} \right| = 0, \quad \text{where} \quad H_{ij}^{\rm so} = \left\langle \psi_{ni}^{(0)} \middle| \mathcal{H}_{\rm so} \middle| \psi_{nj}^{(0)} \right\rangle, \tag{1.162}$$

where $\psi_{ni}^{(0)}, \psi_{nj}^{(0)}$ are degenerate eigenfunctions of $\mathcal{H}_0$. To decide in which basis it is more convenient to compute the matrix elements we note that

$$\mathbf{J} = \mathbf{L} + \mathbf{S}, \qquad J^2 = L^2 + S^2 + 2\mathbf{LS}, \qquad \mathbf{LS} = \frac{1}{2}(J^2 - L^2 - S^2), \tag{1.163}$$

$$\left[\mathcal{H}_0, L^2\right] = \left[\mathcal{H}_0, S^2\right] = \left[\mathcal{H}_0, J^2\right] = [\mathcal{H}_0, L_\mu] = [\mathcal{H}_0, S_\mu] = [\mathcal{H}_0, J_\mu] = 0, \tag{1.164}$$

where L_μ, S_μ and J_μ stand for any component (rectangular or spherical) of $\mathbf{L}$, $\mathbf{S}$ and $\mathbf{J}$, respectively. We have also (see problem (12)),

$$\left[\mathbf{LS}, L^2\right] = \left[\mathbf{LS}, S^2\right] = \left[\mathbf{LS}, J^2\right] = [\mathbf{LS}, J_z] = 0, \tag{1.165}$$

but $[\mathbf{LS}, L_z] \neq 0$ and $[\mathbf{LS}, S_z] \neq 0$. Both $|\ell s m_\ell m_s\rangle$ and $|\ell s j m\rangle$ are eigenfunctions of $\mathcal{H}_0$, however $\mathbf{LS}$ does not commute with L_z and S_z, which means that $|\ell s m_\ell m_s\rangle$ will not necessarily be an eigenfunction of $\mathbf{LS}$. The $|\ell s j m\rangle$ states are simultaneous eigenfunctions of L^2, S^2, J^2, J_z and

[25] It would be a good idea to read appendix B.2 before you proceed.

LS, therefore the coupled basis is the one we should use because only diagonal matrix elements will appear in the secular determinant. We have,

$$\begin{aligned}\langle \ell' s\, j' m' |\, \mathbf{LS}\, | \ell s j m\rangle &= \frac{1}{2}\langle \ell' s\, j' m' |\, (J^2 - L^2 - S^2)\, | \ell s j m\rangle \\ &= \frac{\hbar^2}{2}\Big[j(j+1) - \ell(\ell+1) - s(s+1)\Big]\delta_{j'j}\delta_{\ell'\ell}\delta_{m'm}. \end{aligned} \tag{1.166}$$

The expectation value of the radial function $\xi(r)$ is,

$$\begin{aligned}\langle \xi(r)\rangle = \xi_{n\ell} = \langle n\ell |\, \xi(r)\, | n\ell\rangle &= \int_0^\infty \xi(r) R_{n\ell}^2(r) r^2\, dr = \int_0^\infty \xi(r) u_{n\ell}^2(r)\, dr \\ = \frac{Ze^2}{2m^2c^2}\int_0^\infty \frac{1}{r^3} u_{n\ell}^2(r)\, dr &= \frac{Ze^2}{2m^2c^2}\left\langle n\ell \left| \frac{1}{r^3} \right| n\ell \right\rangle = \frac{Ze^2}{2m^2c^2}\,\frac{Z^3}{a_0^3}\frac{1}{n^3(\ell+1)(\ell+\frac{1}{2})\ell}, \end{aligned} \tag{1.167}$$

where we have used the value of $\langle 1/r^3\rangle$ given by equation (1.134). Combining everything together,

$$\langle \ell s j m\,|\, \xi(r)\mathbf{LS}\,|\, \ell s j m\rangle = \frac{Z^4 e^2 \hbar^2}{4m^2c^2a_0^3}\,\frac{j(j+1) - \ell(\ell+1) - s(s+1)}{n^3(\ell+1)(\ell+\frac{1}{2})\ell}. \tag{1.168}$$

To simplify this take into account that

$$\alpha = \frac{e^2}{\hbar c}, \qquad a_0 = \frac{\hbar^2}{me^2} = \frac{\hbar}{\alpha mc}, \qquad \frac{e^2}{2a_0} = \frac{1}{2}\alpha^2 mc^2 = -E_1^{(0)} = 1\,\text{Ry}. \tag{1.169}$$

So

$$\frac{e^2\hbar^2}{4m^2c^2a_0^3} = \frac{e^2\hbar^2}{4m^2c^2}\,\frac{\alpha^3 m^3 c^3}{\hbar^3} = \frac{e^2}{\hbar c}\,\frac{\alpha^3 mc^2}{4} = \frac{\alpha^4 mc^2}{4} = \left(\frac{1}{2}\alpha^2 mc^2\right)\frac{\alpha^2}{2} = \frac{\alpha^2}{2}\,\text{Ry}. \tag{1.170}$$

Therefore in Rydberg units

$$E_{\text{so}} = \frac{Z^4\alpha^2}{n^3}\,\frac{j(j+1) - \ell(\ell+1) - s(s+1)}{2\ell(\ell+\frac{1}{2})(\ell+1)}\ \text{Ry}, \qquad (\text{for } \ell \neq 0). \tag{1.171}$$

To get an expression valid in any system of units note that

$$E_n^{(0)} = -(Z^2/n^2)\,\text{Ry}, \qquad E_n^{(0)} = -\frac{1}{2}\alpha^2 mc^2\frac{Z^2}{n^2}, \qquad (Z\alpha)^2 = -2n^2\frac{E_n^{(0)}}{mc^2}, \tag{1.172}$$

$$\frac{Z^4\alpha^2}{2n^3} = \frac{1}{2n}\frac{Z^2}{n^2}(Z\alpha)^2 = \frac{1}{2n}E_n^{(0)}\frac{2n^2E_n^{(0)}}{mc^2} = \frac{n}{mc^2}\left(E_n^{(0)}\right)^2, \tag{1.173}$$

so we can rewrite E_{so} as

$$E_{\text{so}} = \frac{\left(E_n^{(0)}\right)^2}{mc^2}\,\frac{n[j(j+1) - \ell(\ell+1) - s(s+1)]}{\ell(\ell+\frac{1}{2})(\ell+1)}. \tag{1.174}$$

If we want to make explicit the fine-structure constant,

$$E_{\text{so}} = -E_n^{(0)}\,(Z\alpha)^2\,\frac{j(j+1) - \ell(\ell+1) - s(s+1)}{2n\ell(\ell+\frac{1}{2})(\ell+1)}. \tag{1.175}$$

Note that this formula for E_{so} diverges when $\ell = 0$. This is due to the fact that we have computed the value of $\langle 1/r^3\rangle$ with the eigenfunctions of the Hamiltonian in (1.87), which has been obtained in the non-relativistic regime where $e\phi \ll mc^2$. However, the Coulomb energy Ze^2/r becomes infinite at $r = 0$ (where the wave function for $\ell = 0$ is still finite) and therefore

the approximation $e\phi \ll mc^2$ becomes invalid. In any case, atomic nuclei are not point-like electric charges generating an exact Coulomb potential, so a more careful study should be done at short distances. A proper calculation with an improved approximation[26] reveals that $\langle 1/r^3 \rangle$ is finite and we can safely use the fact that the matrix element of **LS** vanishes for $\ell = 0$ to conclude that $E_{\rm so} = 0$ for s states.

For one-electron systems $s(s+1) = 3/4$, therefore

$$j(j+1) - \ell(\ell+1) - s(s+1) = \begin{cases} \ell & \text{for } j = \ell + \frac{1}{2} \\ -(\ell+1) & \text{for } j = \ell - \frac{1}{2} \end{cases} \tag{1.176}$$

We can then rewrite $E_{\rm so}$ in Rydberg units as

$$E_{\ell+\frac{1}{2}} = \frac{Z^4\alpha^2}{n^3} \frac{\ell}{2\ell(\ell+\frac{1}{2})(\ell+1)} \text{ Ry}, \qquad E_{\ell-\frac{1}{2}} = -\frac{Z^4\alpha^2}{n^3} \frac{(\ell+1)}{2\ell(\ell+\frac{1}{2})(\ell+1)} \text{ Ry} \tag{1.177}$$

$$\Delta E_{\rm so} = E_{\ell+\frac{1}{2}} - E_{\ell-\frac{1}{2}} = \frac{Z^4\alpha^2}{n^3} \frac{1}{\ell(\ell+1)} \text{ Ry}. \tag{1.178}$$

And in arbitrary units,

$$E_{\ell+\frac{1}{2}} = \frac{\left(E_n^{(0)}\right)^2}{mc^2} \frac{n\ell}{\ell(\ell+\frac{1}{2})(\ell+1)}, \qquad E_{\ell-\frac{1}{2}} = -\frac{\left(E_n^{(0)}\right)^2}{mc^2} \frac{n(\ell+1)}{\ell(\ell+\frac{1}{2})(\ell+1)} \tag{1.179}$$

$$\Delta E_{\rm so} = E_{\ell+\frac{1}{2}} - E_{\ell-\frac{1}{2}} = \frac{\left(E_n^{(0)}\right)^2}{mc^2} \frac{2n}{\ell(\ell+1)}. \tag{1.180}$$

The size of the spin-orbit energy shift in the Hydrogen atom is of the order of 10^{-5} eV, or $0.1\,\text{cm}^{-1}$. Since the energies between states with different values of the principal quantum number, n, are about $10^4\,\text{cm}^{-1}$, we see that the spin-orbit energy shifts are much smaller and the first-order perturbation theory seems appropriate. Nevertheless, one has to take into account that the spin-orbit energy increases rapidly with the atomic number due to the Z^4 dependence.

The conditions on j, ℓ and m for the spin-orbit interaction matrix elements to be nonzero can be read directly from equation (1.166). They are

$$\Delta j = \Delta\ell = \Delta m = \Delta(m_\ell + m_s) = 0. \tag{1.181}$$

To see if there are any restrictions on the individual values of m_ℓ and m_s we should compute the matrix elements of **LS** in the uncoupled basis, $|\ell s m_\ell m_s\rangle$. To do this it is convenient to express **L** and **S** in spherical components,

$$L_{+1} = -\frac{1}{\sqrt{2}}(L_x + iL_y), \qquad L_{-1} = \frac{1}{\sqrt{2}}(L_x - iL_y), \qquad L_0 = L_z \tag{1.182}$$

$$S_{+1} = -\frac{1}{\sqrt{2}}(S_x + iS_y), \qquad S_{-1} = \frac{1}{\sqrt{2}}(S_x - iS_y), \qquad S_0 = S_z \tag{1.183}$$

$$L_x = -\frac{1}{\sqrt{2}}(L_{+1} - L_{-1}), \qquad L_y = \frac{i}{\sqrt{2}}(L_{+1} + L_{-1}), \qquad L_z = L_0 \tag{1.184}$$

$$S_x = -\frac{1}{\sqrt{2}}(S_{+1} - S_{-1}), \qquad S_y = \frac{i}{\sqrt{2}}(S_{+1} + S_{-1}), \qquad S_z = S_0 \tag{1.185}$$

This leads to $\mathbf{LS} = -L_{+1}S_{-1} + L_0S_0 - L_{-1}S_{+1}$, and therefore the matrix element

$$\left\langle \ell s m'_\ell m'_s \right| -L_{+1}S_{-1} + L_0S_0 - L_{-1}S_{+1} \left| \ell s m_\ell m_s \right\rangle \tag{1.186}$$

[26] We do not perform this particular calculation here, but a similar one will be done when studying hyperfine interactions.

	$\left\vert 1\frac{1}{2}\right\rangle$	$\left\vert 1-\frac{1}{2}\right\rangle$	$\left\vert 0\frac{1}{2}\right\rangle$	$\left\vert 0-\frac{1}{2}\right\rangle$	$\left\vert -1\frac{1}{2}\right\rangle$	$\left\vert -1-\frac{1}{2}\right\rangle$
$\left\langle 1\frac{1}{2}\right\vert$	$\frac{1}{2}$					
$\left\langle 1-\frac{1}{2}\right\vert$		$-\frac{1}{2}$	$\frac{1}{\sqrt{2}}$			
$\left\langle 0\frac{1}{2}\right\vert$		$\frac{1}{\sqrt{2}}$	0			
$\left\langle 0-\frac{1}{2}\right\vert$				0	$\frac{1}{\sqrt{2}}$	
$\left\langle -1\frac{1}{2}\right\vert$				$\frac{1}{\sqrt{2}}$	$-\frac{1}{2}$	
$\left\langle -1-\frac{1}{2}\right\vert$						$\frac{1}{2}$

Table 1.3: $\left\langle 1\frac{1}{2}m_\ell m_s \,\middle|\, \xi(r)\mathbf{LS} \,\middle|\, 1\frac{1}{2}m_\ell m_s\right\rangle$ matrix elements for p states in units of $\hbar^2\xi_{n\ell} = \hbar^2\langle n\ell|\xi(r)|n\ell\rangle$. The notation has been shortened from $|\ell s m_\ell m_s\rangle$ to $|m_\ell m_s\rangle$ because $\ell = 1$ and $s = 1/2$ for all the states in the table.

vanishes unless $\Delta m_\ell = 0, \pm 1$ and $\Delta m_s = 0, \pm 1$. A further requirement is that, if m_ℓ increases by one unit, m_s must decrease by one unit and vice-versa. In summary, the matrix element is zero unless

$$\Delta m_\ell = \Delta m_s = 0, \qquad \text{or} \qquad \Delta m_\ell = -\Delta m_s = \pm 1. \tag{1.187}$$

As a reference, the different values of $\left\langle 1\frac{1}{2}m_\ell m_s \,\middle|\, \xi(r)\mathbf{LS} \,\middle|\, 1\frac{1}{2}m_\ell m_s\right\rangle$ are shown in table (1.3) in units of $\hbar^2\xi_{n\ell} = \hbar^2\langle n\ell|\xi(r)|n\ell\rangle$.

1.6 Relativistic correction to the kinetic energy.

From (1.87) we see that the relativistic correction to the kinetic energy is given by

$$\mathcal{H}_r = -\frac{p^4}{8m^3c^2} = -\frac{1}{2mc^2}\left(\frac{p^2}{2m}\right)^2. \tag{1.188}$$

Using $\mathcal{H}_0 = \dfrac{p^2}{2m} - \dfrac{Ze^2}{r}$ we have,

$$\mathcal{H}_r = -\frac{1}{2mc^2}\left(\mathcal{H}_0 + \frac{Ze^2}{r}\right)^2 = -\frac{1}{2mc^2}\left[\mathcal{H}_0^2 + Ze^2\left(\mathcal{H}_0\frac{1}{r} + \frac{1}{r}\mathcal{H}_0\right) + \frac{(Ze^2)^2}{r^2}\right]. \tag{1.189}$$

We want to compute the effects produced by $\mathcal{H}_r$ within a multiplet of states specified by particular values of n, ℓ, s and j, as for example the ${}^2S_{1/2}$ or ${}^2P_{3/2}$ eigenstates of $\mathcal{H}_0$ belonging to a particular value of n. Therefore, treating $\mathcal{H}_r$ as a perturbation, the basis set is $|n\ell sjm\rangle$ and the relevant matrix elements are $\langle n\ell' sj'm'|\mathcal{H}_r|n\ell sjm\rangle$. Let $E_n^{(0)}$ be an eigenvalue of $\mathcal{H}_0$, then

$$\left\langle n\,\ell' sj'm' \,\middle|\, \mathcal{H}_0^2 \,\middle|\, n\ell sjm\right\rangle = \left(E_n^{(0)}\right)^2 \delta_{\ell\ell'}\delta_{jj'}\delta_{mm'}. \tag{1.190}$$

Since $[\mathbf{p}, \mathbf{r}] \neq 0$, $\mathcal{H}_0$ and $1/r$ do not commute. However $\mathcal{H}_0$ is Hermitian,[27] therefore we find the following two matrix elements to be equal

$$\begin{aligned}\left\langle n\ell' sj'm' \,\middle|\, \mathcal{H}_0\frac{1}{r} \,\middle|\, n\ell sjm\right\rangle &= \left\langle n\ell' sj'm' \,\middle|\, \frac{1}{r}\mathcal{H}_0 \,\middle|\, n\ell sjm\right\rangle \\ &= E_n^{(0)}\left\langle n\ell' sj'm' \,\middle|\, \frac{1}{r} \,\middle|\, n\ell sjm\right\rangle = E_n^{(0)}\left\langle \frac{1}{r}\right\rangle \delta_{\ell\ell'}\delta_{jj'}\delta_{mm'}.\end{aligned} \tag{1.191}$$

[27] Remember that a Hermitian operator satisfies $\langle\psi \,|\, A^\dagger\phi\rangle = \langle A\psi \,|\, \phi\rangle = \langle\phi \,|\, A\psi\rangle^*$. So

$$\begin{aligned}\left\langle n\ell sjm \,\middle|\, \mathcal{H}_0\frac{1}{r} \,\middle|\, n\ell sjm\right\rangle = \left\langle \psi_{n\ell sjm} \,\middle|\, \mathcal{H}_0\frac{1}{r} \,\middle|\, \psi_{n\ell sjm}\right\rangle &= \left\langle \mathcal{H}_0^\dagger\psi_{n\ell sjm} \,\middle|\, \frac{1}{r} \,\middle|\, \psi_{n\ell sjm}\right\rangle \\ &= \left\langle \mathcal{H}_0\psi_{n\ell sjm} \,\middle|\, \frac{1}{r} \,\middle|\, \psi_{n\ell sjm}\right\rangle = E_n^{(0)}\left\langle \psi_{n\ell sjm} \,\middle|\, \frac{1}{r} \,\middle|\, \psi_{n\ell sjm}\right\rangle\end{aligned}$$

In the last step, we have taken into account that the value of $\langle 1/r \rangle$ only depends on the radial part of the wave function and is independent of ℓ, j and m. Similarly

$$\left\langle n\ell' sj'm' \left| \frac{1}{r^2} \right| n\ell sjm \right\rangle = \left\langle \frac{1}{r^2} \right\rangle \delta_{\ell\ell'}\delta_{jj'}\delta_{mm'}. \tag{1.192}$$

Thus $\mathcal{H}_r$ has only diagonal matrix elements and its expectation value is

$$\langle \mathcal{H}_r \rangle = -\frac{E_n^{(0)}}{2mc^2}\left[E_n^{(0)} + 2Ze^2\left\langle \frac{1}{r} \right\rangle + \frac{(Ze^2)^2}{E_n^{(0)}}\left\langle \frac{1}{r^2} \right\rangle\right]. \tag{1.193}$$

For Hydrogen we have

$$E_n^{(0)} = -\frac{Z^2}{n^2}\frac{e^2}{2a_0} = -\frac{Z^2}{n^2}\frac{me^4}{2\hbar^2} = -\frac{Z^2}{n^2}\,\mathrm{Ry}, \tag{1.194}$$

and

$$\left\langle \frac{1}{r} \right\rangle = \frac{Z}{a_0}\frac{1}{n^2}, \qquad \left\langle \frac{1}{r^2} \right\rangle = \frac{Z^2}{a_0^2}\frac{1}{n^3(\ell+\frac{1}{2})}. \tag{1.195}$$

$$\begin{aligned}\langle \mathcal{H}_r \rangle &= -\frac{E_n^{(0)}}{2mc^2}\left[-\frac{Z^2}{n^2}\frac{e^2}{2a_0} + 2Ze^2\frac{Z}{a_0}\frac{1}{n^2} - \frac{n^2}{Z^2}\frac{2a_0}{e^2}(Ze^2)^2\frac{Z^2}{a_0^2}\frac{1}{n^3(\ell+\frac{1}{2})}\right] \\ &= -\frac{E_n^{(0)}Z^2e^2}{mc^2a_0}\left[-\frac{1}{4n^2} + \frac{1}{n^2} - \frac{1}{n(\ell+\frac{1}{2})}\right] = E_n^{(0)}Z^2\alpha^2\left[-\frac{3}{4n^2} + \frac{1}{n(\ell+\frac{1}{2})}\right],\end{aligned} \tag{1.196}$$

where

$$\alpha^2 = \left(\frac{e^2}{\hbar c}\right)^2 = \frac{e^2}{mc^2a_0}, \qquad a_0 = \frac{\hbar^2}{me^2}. \tag{1.197}$$

We have obtained a formula for $\langle \mathcal{H}_r \rangle$ which is valid for all values of ℓ (including $\ell = 0$), namely.

$$E_{\mathrm{r}} = \langle \mathcal{H}_r \rangle = E_n^{(0)}\frac{(Z\alpha)^2}{n}\left[-\frac{3}{4n} + \frac{1}{(\ell+\frac{1}{2})}\right]. \tag{1.198}$$

To write this in Rydberg units note that $E_n^{(0)} = -(Z^2/n^2)$ Ry, therefore

$$E_{\mathrm{r}} = -\frac{Z^4\alpha^2}{n^3}\left[-\frac{3}{4n} + \frac{1}{(\ell+\frac{1}{2})}\right]\ \mathrm{Ry} \tag{1.199}$$

Using the relations

$$E_n^{(0)} = -\frac{1}{2}\alpha^2mc^2\frac{Z^2}{n^2}, \qquad \frac{Z^2}{n^2} = -\frac{2E_n^{(0)}}{\alpha^2mc^2}, \tag{1.200}$$

we can also write the relativistic correction in this useful form

$$E_{\mathrm{r}} = \frac{E_n^{(0)}Z^2\alpha^2}{4n^2}\left(\frac{4n}{\ell+\frac{1}{2}} - 3\right) = -\frac{2\left(E_n^{(0)}\right)^2\alpha^2}{4\alpha^2mc^2}\left(\frac{4n}{\ell+\frac{1}{2}} - 3\right) \tag{1.201}$$

$$E_{\mathrm{r}} = -\frac{\left(E_n^{(0)}\right)^2}{2mc^2}\left(\frac{4n}{\ell+\frac{1}{2}} - 3\right) \tag{1.202}$$

1.7 The Darwin term.

For the Darwin term let's consider $\mathbf{E} = -\boldsymbol{\nabla}\phi$, with $\phi = Ze/r$. Then

$$\mathcal{H}_{\mathrm{D}} = \frac{e\hbar^2}{8m^2c^2}\boldsymbol{\nabla}\cdot\mathbf{E} = -\frac{Ze^2\hbar^2}{8m^2c^2}\nabla^2\left(\frac{1}{r}\right). \tag{1.203}$$

The direct computation of $\nabla^2(1/r)$ would give zero everywhere except, possibly, at $r = 0$. This also happens with the divergence of $\hat{\mathbf{r}}/r^2$,

$$\boldsymbol{\nabla}\left(\frac{\hat{\mathbf{r}}}{r^2}\right) = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{1}{r^2}\right) = 0, \quad \text{for} \quad r \neq 0. \tag{1.204}$$

On the other hand, using Gauss's theorem we can compute the volume integral of $\boldsymbol{\nabla}(\hat{\mathbf{r}}/r^2)$ over the interior of a sphere of radius R, centered at the origin, as the integral of $\hat{\mathbf{r}}/r^2$ over the surface of the same sphere. The result is

$$\int \boldsymbol{\nabla}\left(\frac{\hat{\mathbf{r}}}{r^2}\right)\mathrm{d}^3\mathbf{r} = \oint\left(\frac{\hat{\mathbf{r}}}{r^2}\right)\mathrm{d}\boldsymbol{\sigma} = \int\left(\frac{\hat{\mathbf{r}}}{R^2}\right)R^2\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi\,\hat{\mathbf{r}} = 4\pi. \tag{1.205}$$

Taken together, equations (1.204) and (1.205) tell us that $\boldsymbol{\nabla}(\hat{\mathbf{r}}/r^2) = 4\pi\delta(\mathbf{r})$. Now, since $\boldsymbol{\nabla}(1/r) = -\hat{\mathbf{r}}/r^2$, we have $\nabla^2(1/r) = -\boldsymbol{\nabla}(\hat{\mathbf{r}}/r^2) = -4\pi\delta(\mathbf{r})$. Therefore the Hamiltonian of the Darwin term is

$$\mathcal{H}_{\mathrm{D}} = \frac{4\pi Ze^2\hbar^2}{8m^2c^2}\delta(\mathbf{r}). \tag{1.206}$$

It is clear that we only need to consider the radial part of the wave function. The matrix elements of the delta function are just the values of the wave functions at the origin, $\langle n\ell|\delta(\mathbf{r})|n\ell\rangle = |\psi(0)|^2$, therefore

$$\langle\mathcal{H}_{\mathrm{D}}\rangle = \frac{\pi Ze^2\hbar^2}{2m^2c^2}|\psi(0)|^2 = E_{\mathrm{D}}. \tag{1.207}$$

This formula implies that $\langle\mathcal{H}_{\mathrm{D}}\rangle$ is nonzero only for s states. In Hydrogen

$$|\psi(0)|^2 = \frac{Z^3}{\pi n^3 a_0^3}, \qquad E_{\mathrm{D}} = \frac{\pi Ze^2\hbar^2}{2m^2c^2}\frac{Z^3}{\pi n^3 a_0^3} = \frac{Z^4e^2\hbar^2}{2n^3m^2c^2a_0^3}. \tag{1.208}$$

In Rydberg units,

$$E_{\mathrm{D}} = \frac{Z^4\alpha^2}{n^3}\ \mathrm{Ry}, \qquad (\ell = 0 \ \text{ only}) \tag{1.209}$$

Since $E_n^{(0)} = -(Z^2/n^2)$ Ry, we can also write

$$E_{\mathrm{D}} = -E_n^{(0)}\frac{(Z\alpha)^2}{n}, \qquad (\ell = 0 \ \text{ only}). \tag{1.210}$$

To make the comparison with mc^2 explicit we can trade $(Z\alpha)^2$ for another factor of $E_n^{(0)}$ using the relation $E_n^{(0)} = -\frac{1}{2}\alpha^2mc^2\frac{Z^2}{n^2}$. Therefore

$$E_{\mathrm{D}} = 2n\frac{\left(E_n^{(0)}\right)^2}{mc^2}. \tag{1.211}$$

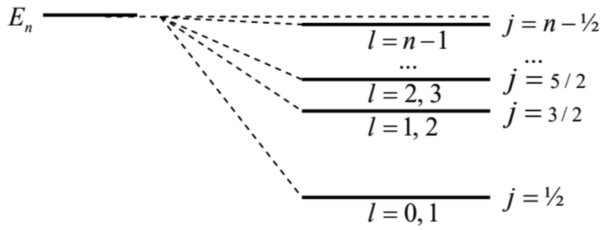


Figure 1.1: Fine structure energy shifts on the n state of a hydrogen atom (not to scale).

1.8 Combining all corrections.

To recapitulate, the corrections due to the spin-orbit effect and the Darwin term are

$$E_{\rm so} = \frac{Z^4\alpha^2}{n^3}\,\frac{j(j+1)-\ell(\ell+1)-s(s+1)}{2\ell(\ell+\frac{1}{2})(\ell+1)}\ {\rm Ry}, \qquad E_{\rm D} = \frac{Z^4\alpha^2}{n^3}\ {\rm Ry}. \tag{1.212}$$

The expression for $E_{\rm so}$ is only valid for $\ell \neq 0$, although we know that $E_{\rm so} = 0$ for states with $\ell = 0$ anyhow. On the contrary, the Darwin correction applies only to states with $\ell = 0$. The relativistic correction is valid for all values of ℓ,

$$E_{\rm r} = -\frac{Z^4\alpha^2}{n^3}\left[\frac{1}{(\ell+\frac{1}{2})} - \frac{3}{4n}\right]\ {\rm Ry}. \tag{1.213}$$

Taking all this into account, it is easy to show (problem 14) that by adding the three corrections we obtain a formula that depends only on n and j, but not on ℓ or m.

$$E_{\rm so} + E_{\rm r} + E_{\rm D} = -\frac{Z^4\alpha^2}{n^3}\left[\frac{1}{j+\frac{1}{2}} - \frac{3}{4n}\right]\ {\rm Ry}. \tag{1.214}$$

The zero-order Hydrogen energy levels are given by $E_n^{(0)} = -(Z^2/n^2)$ Ry. By including all corrections we have computed we arrive then at

$$E_{nj} = -\frac{Z^2}{n^2}\left[1 + \frac{Z^2\alpha^2}{n}\left(\frac{1}{j+\frac{1}{2}} - \frac{3}{4n}\right)\right]\ {\rm Ry}. \tag{1.215}$$

In arbitrary units,

$$E_{nj} = E_n^{(0)}\left[1 + \frac{Z^2\alpha^2}{n}\left(\frac{1}{j+\frac{1}{2}} - \frac{3}{4n}\right)\right]. \tag{1.216}$$

The splitting that arises within a manifold of states belonging to the same value of the principal quantum number, n, is known as *fine structure*.

To summarize: without relativistic corrections, the energy E_n depends only on the principal quantum number n. After relativistic corrections, the E_n energy level splits into n different levels, one for each value $j = 1/2, 3/2, 5/2 \ldots, n-(1/2)$ of the total angular momentum quantum number j. This splitting is called *fine structure splitting*, and the n levels $j = 1/2, \ldots, n-(1/2)$ are said to form a *fine structure multiplet* (see Fig. 1.1). The dimensionless constant that controls the scale of the splitting, $\alpha \approx 1/137$, receives the name of *fine structure constant*.

1.9 The Lamb shift.

It is interesting to note that equation (1.215) predicts the energies of the $^2{\rm S}_{1/2}$ and $^2{\rm P}_{1/2}$ levels (notation $^{2s+1}{\rm L}_j$) to be the same for a given n (see Fig. 1.2). This is a result that holds even for

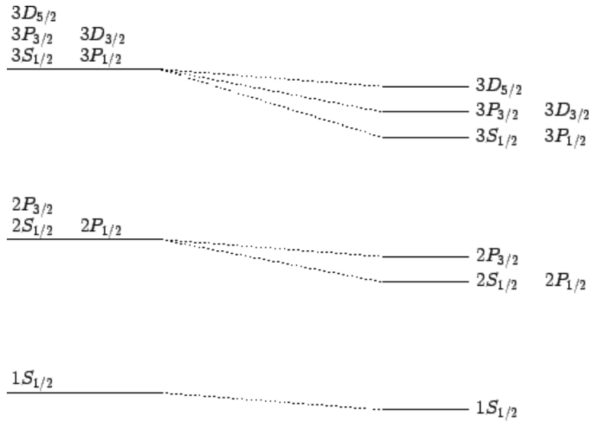


Figure 1.2: Fine structure energy shifts on the $n = 1, 2, 3$ states of a hydrogen atom (not to scale).

the exact solution of the Dirac equation for Hydrogen, which predicts that states with the same value of the quantum number j, but different values of ℓ, have the same energy. However, already in the late 1930's several small discrepancies were observed between the measured spectra and theoretical expectations. The $^2\mathrm{S}_{1/2}$ and $^2\mathrm{P}_{1/2}$ levels, for example, did not coincide because there existed a slight upward shift of the $^2\mathrm{S}_{1/2}$ level. Nevertheless, not all observations around that time agreed with each other. The broadening of spectral lines, due mainly to the Doppler effect, made high-precision spectroscopy measurements very difficult. In 1947 W. E. Lamb and R. C. Retherford performed a brilliant experiment using microwave techniques to stimulate a direct radio-frequency transition between the $^2\mathrm{S}_{1/2}$ and $^2\mathrm{P}_{1/2}$ levels. Since the Doppler effect is proportional to the frequency of the waves, working in the microwave regime guarantees negligible line-broadening due to this effect.

The need to explain the so-called *Lamb shift* stimulated the revision of fundamental physical concepts such as the renormalization of mass, and contributed to the final formulation of the theory of quantum electrodynamics. The calculation of the Lamb shift can only be done with the methods of quantum field theory and is beyond the scope of this course. However, a qualitative explanation may be provided in the following way. A quantized radiation field in its lowest energy state is not characterized by a vanishing amplitude of electromagnetic fields. There exist zero-point oscillations similar to those which appear in the study of the quantum-mechanical harmonic oscillator. This means that even in the vacuum there are fluctuations in this zero-point radiation field which can act on the electron, causing it to execute rapid oscillatory motions so that its charge is *smeared out* and the point electron effectively looks like a spherical charge of a certain radius. If the electron is bound by a non-uniform electric field, as in atomic systems, these oscillations will make it experience a potential that is slightly different from that corresponding to its mean position. In particular, the electron in a one-electron atom would not be so strongly attracted to the nucleus at short distances. As a result, s-states (which are most sensitive to short distance modifications because $|\psi(0)|^2 \neq 0$ for these states) are raised in energy with respect to other states for which the corresponding modifications are much smaller.

The present measured value for the energy difference between the $^2\mathrm{S}_{1/2}$ and $^2\mathrm{P}_{1/2}$ levels of Hydrogen is $3.36 \times 10^{-5}\,\mathrm{eV}$ (we do not quote the full available accuracy here) and there is good agreement between theory and experiment. The proper calculation of the Lamb shift has to take into account higher-order radiative corrections in addition to the vacuum polarization effect just described, and also the proton size.[28] Figure (1.3) shows some of the loop diagrams contributing

[28] In fact, vacuum polarization contributes only a small part of the radiative corrections in ordinary atoms. However, in muonic atoms (where an electron is replaced by a muon), the vacuum polarization dominates the

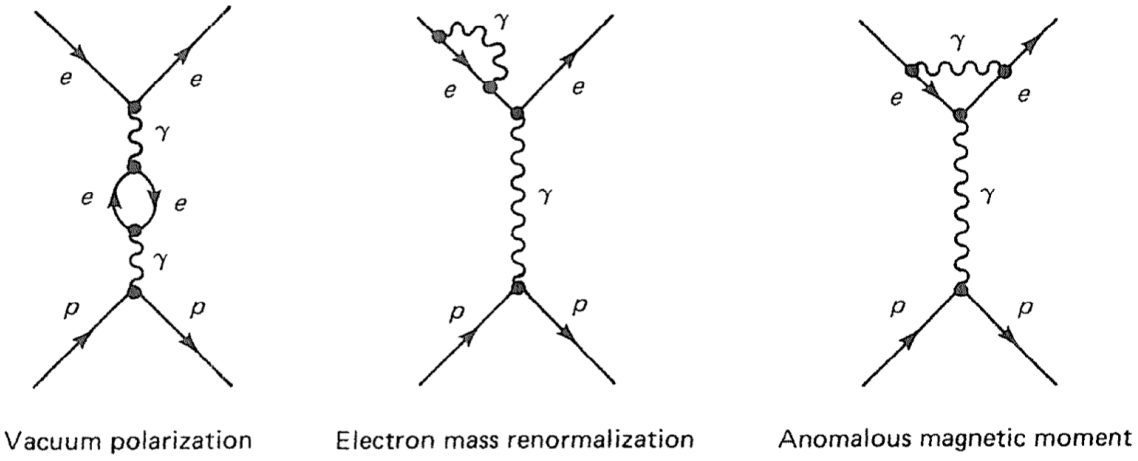


Figure 1.3: Some loop diagrams contributing to the Lamb shift. Taken from Griffiths [8].

to the Lamb shift. At present, the greatest uncertainty in the theoretical calculations comes from the root-mean-square charge radius of the proton, which has to be taken from experiment.

For $\ell = 0$ the Lamb shift has been computed to be

$$\Delta E_{\text{Lamb}} = \alpha^5 mc^2 \frac{1}{4n^3} k(n, 0), \tag{1.217}$$

where $k(n, 0)$ is a numerical factor that varies slightly with n, from 12.7 for $n = 1$ to 13.2 for $n \to \infty$. For $\ell \neq 0$ the Lamb shift is

$$\Delta E_{\text{Lamb}} = \alpha^5 mc^2 \frac{1}{4n^3} \left[k(n, \ell) \pm \frac{1}{\pi(j + \frac{1}{2})(\ell + \frac{1}{2})} \right] \quad \text{for} \quad j = \ell \pm \frac{1}{2}, \tag{1.218}$$

where $k(n, \ell)$ is a very small number (less than 0.05) that varies slightly with n and ℓ. The Lamb shift is very small, except for s states where is about 1/10 of the size of the fine structure.

The hierarchy of corrections to the (Bohr) Hydrogen energy levels is the following

Bohr energies:	of order	$\alpha^2 mc^2$
Fine structure:	of order	$\alpha^4 mc^2$
Lamb shift:	of order	$\alpha^5 mc^2$
Hyperfine splitting:	of order	$(m/m_p)\alpha^4 mc^2$

The hyperfine splitting, which will be computed in chapter 2 (see equation (2.172), for example), is smaller than the Lamb shift by an order of magnitude due to the small value of the electron/proton mass ratio.

1.10 Special hydrogen-like atomic systems.

1.10.1 Antihydrogen.

The antihydrogen, a $\overline{\text{p}}\text{e}^+$ bound system, is specially interesting to study possible fundamental differences between matter and antimatter. To be sure, we already know there are differences between matter and antimatter. Perhaps the most conspicuous one is that antimatter is essentially absent in the visible Universe. Still, we expect the antihydrogen to behave exactly as

radiative corrections.

hydrogen, since this is an electromagnetic system where we expect the weak interaction to play a negligible role.

Antihydrogen atoms were first produced at CERN in 1996. Several experiments take nowadays place in the so-called CERN's Antiproton Decelerator (AD), a complex that produces low-energy antiprotons for studies of antimatter, and creates antiatoms. In one recent publication the ALPHA collaboration has reported the observation of the 1S-2S transition in trapped antihydrogen atoms. They are magnetically trapped using a very sophisticated superconducting trapping device. The vacuum region where they are confined is a cylinder of 28 centimeters length and a diameter of 4.4 cm. Special windows let a laser beam go through the chamber. One of the goals is to measure the frequency of the 1S-2S transition in antihydrogen and compare it with the very precise measured valued in hydrogen, 2466061413187035(10) Hz. A significant deviation from this value would mean a breakdown of CPT symmetry, but no anomaly has been observed so far.

1.10.2 Positronium and muonium.

The positronium is a e^+e^- bound system. Essentially all formulas we have studied so far can adapted to this (and others) exotic hydrogenic atom by changing the mass of the electron by the so-called reduced mass. For a system of two particles with masses m_1 and m_2, the reduced mass is defined as

$$\mu = \frac{m_1 m_2}{m_1 + m_2}. \tag{1.219}$$

For positronium $\mu = m_e/2$, and the unperturbed Bohr energies are given by half of the formula we derived for hydrogen,

$$E_n^{\rm pos} = \frac{1}{2} E_n = -\alpha^2 mc^2 \frac{1}{4n^2}. \tag{1.220}$$

The Bohr radius for this system is a factor of 2 larger, $a_0^{\rm pos} = 2a_0 = 1.06 \times 10^{-8}\,$cm. The perturbative corrections get similar factors, except the hyperfine splitting, which is of the same order than the fine structure because we don't have anymore the mass suppression ratio m_e/m_p. The positronium is an interesting test system for QED. It has been produced in the laboratory and there are several measurements that could be discussed, but this is not the place to do it. Let's finish this section by remarking that this is not a stable system. It decays to two photons with a lifetime of

$$\tau = \frac{2\hbar}{\alpha^5 mc^2} = 1.25 \times 10^{-10}\ \text{sec}. \tag{1.221}$$

The muonium is a μ^+e^- bound system. The mass of the muon is about 207 times the mass of the electron, and all formulas have to be adapted with a new value of the reduced mass. The lifetime of the muonium is 2.2×10^{-6} seconds, which is the lifetime of the muon itself.

1.10.3 Rydberg atoms.

An atom with an electron in a very high-energy state, with principal quantum number even as high as $n = 100$, is called a Rydberg atom. The size of such an atom is truly big compared to atomic standards. With electron orbital radii of the order of 10^{-7} m, it is almost as big as a bacteria. Of course, the ionization energy for hydrogen atoms with $n = 100$ is very small, about 1.36×10^{-3} eV. Another important feature is that the energy separation for states around $n = 100$ is also very small, because we are approaching the region were the spectrum becomes continuum.

The study of Rydberg atoms is important for several areas of astrophysics, plasma physics and quantum optics. The highly excited electron of a Rydberg can be removed very easily, even in thermal collisions. However, in the interstellar vacuum space a Rydberg atom can in principle have a long lifetime and emit radio waves which can be detected on Earth by radio telescopes.

1.10.4 Pionium.

The pionium is a $\pi^+\pi^-$ bound system, called usually $A_{2\pi}$. The reduced mass is now $m_\pi/2$ ($m_\pi = 139.57\,\mathrm{MeV}/c^2$), leading to a Bohr radius of $378\,\mathrm{fm} = 3.78 \times 10^{-13}\,\mathrm{m}$. This *exotic atom* is very unstable, it decays mainly from the $1s$ state to two neutral pions, $A_{2\pi} \to \pi^0 + \pi^0$, with a lifetime of the order of 10^{-15} seconds. The measurement of the $A_{2\pi}$ decay width to two neutral pions, $\Gamma_{2\pi^0}$, is an important input to understand quantum chromodynamics (QCD) at low energy.[29] We mention this hydrogen-like atom because, as it happens, the University of Santiago de Compostela was a founding member of the DIRAC (Dimeson Relativistic Atom Complex) collaboration at CERN, which had as its primary goal to measure $\Gamma_{2\pi^0}$. To produce pionium atoms, protons accelerated up to $24\,\mathrm{GeV}/c$ by the CERN PS accelerator were directed into thin target foils of less than 100 μm thickness. The proton collision with the target produces many particles, most of them pions, which can form $A_{2\pi}$ atoms. These atoms, produced at relativistic energies with a large boost, travel inside the target foil interacting with the ordinary atoms of the target. This may lead to the break up (ionization) of the $A_{2\pi}$ atom before it has had time to decay to $2\pi^0$. By measuring the momentum spectra of pairs of oppositely charged pions leaving the target foil it is possible to infer the ionization probability and the $\Gamma_{2\pi^0}$ decay width.

1.10.5 Quarkonium.

In the quark model all mesons are bound states of a quark and an antiquark, $\mathrm{q_1\bar{q}_2}$. The light-quark mesons, like for example the π^0, which is a combination of $\mathrm{u\bar{u}}$ and $\mathrm{d\bar{d}}$ states, are intrinsically relativistic and a simple analysis is not possible. Heavy-quark mesons, however, like the J/ψ ($\mathrm{c\bar{c}}$) or the η_b ($\mathrm{b\bar{b}}$), could be possible candidates. Even in these systems the interaction energy, E, is so large that the different energy levels are treated as *different particles*, with masses given by

$$M = m_1 + m_2 + \frac{E}{c^2} \tag{1.222}$$

A very important difference with respect to the preceding systems, where the binding is due to the electromagnetic interaction, is that the binding is now due to the strong force. The potential to use, instead of the Coulomb one, could in principle be derived from quantum chromodynamics (QCD), but nobody has been able to do it yet. For reasons that we don't go into here, it is known that the strong potential at short distances (but not very short) goes like $1/r$, as in the Coulomb case. For large distances, the potential may be taken to be linear with r, so a possible parametrization could be

$$V(r) = -\frac{4}{3}\frac{\alpha_s \hbar c}{r} + F_0 r, \tag{1.223}$$

where α_s is the chromodynamic analog to the fine structure constant. If the constant F_0 is chosen to fit the experimental data, the result is about 900 MeV fm^{-1}. To have a feeling of how strong two quarks attract each other, we may translate those 900 MeV fm^{-1} to 14×10^4 J/m and boldly say that it would be sufficient to lift a weight of about 14 tons. However, quarks cannot be found as free particles outside hadrons. If enough energy is put into a $\mathrm{q\bar{q}}$ system, a new quark-antiquark pair is formed instead of tearing apart the old one.

[29] Although the $A_{2\pi}$ is an electromagnetic bound system, the decay to $2\pi^0$ is due to the strong interaction.

1.11 Problems.

1. By definition, a vector is a three-element tensor V_i, $i = 1, 2, 3$, that verifies the following commutation relations with the angular momentum operators

$$[J_i, V_j] = i\hbar\epsilon_{ijk}V_k,$$

where Einstein's convention of summing over repeated indexes applies. The spherical components of a vector are constructed as follows

$$V_{+1} = -\frac{1}{\sqrt{2}}(V_1 + iV_2), \qquad V_{-1} = \frac{1}{\sqrt{2}}(V_1 - iV_2), \qquad V_0 = V_3.$$

Also by definition, an irreducible tensor of rank j is a set of $2j + 1$ operators, O_j^m, with $m = -j, -j + 1, \ldots, j - 1, j$, such that their commutators with the angular momentum operators are

$$[J_3, O_j^m] = \hbar m\, O_j^m,$$
$$[J_1 \pm iJ_2, O_j^m] = \hbar\sqrt{j(j+1) - m(m \pm 1)}\, O_j^{m\pm1}.$$

Demonstrate that the spherical components of any vector constitute an irreducible tensor of rank $j = 1$.

2. The purpose of this problem is to help you understand the Wigner-Eckart theorem by providing simple examples of how to use it. Consider the cartesian tensor of rank 2 defined as $T_{ij} = x_i x_j$, where x_i denotes the i-th Cartesian component of the position vector.

   a) Factorize T_{ij} using the irreducible tensor of rank two $O^m_{k=2} = \sqrt{\frac{32\pi}{15}} r^2 Y_2^m(\theta, \varphi)$ and the irreducible tensor of rank zero $O^0_{k=0} = \sqrt{4\pi} r^2 Y_0^0(\theta, \varphi)$, where $Y_\ell^m(\theta, \varphi)$ are spherical harmonics, in this case for $\ell = 0, 2$. Why contributions of an irreducible tensor of rank 1 are not expected in this case?

   Hint1: write explicitly the Y_2^m and Y_0^0 as a function of x_i and r and identify the T_{ij} components.

   Hint2: You may also need to consider:

$$Y_0^0(\theta, \varphi) = \sqrt{\frac{1}{4\pi}} = \sqrt{\frac{1}{4\pi}}\frac{(x^2 + y^2 + z^2)}{r^2}.$$

   b) If $|n\ell m\rangle$ denotes an eigenfunction of the Hydrogen atom (without considering its spin), and $\chi \equiv \langle 322|xy|300\rangle$. Compute the matrix elements $\langle 32m|T_{ij}|300\rangle$ as a function of the matrix element χ. This exercise shows that, thanks to the Wigner-Eckart theorem, all $\langle 32m|T_{ij}|300\rangle$ matrix elements are equal to $\langle 322|xy|300\rangle$ multiplied by a combination of CG coefficients. Thus, if needed, we would only compute the χ matrix element and then look up the relevant CG coefficients in a table to find the values of the remaining ones.

3. a) Working in one dimension to make the notation simpler, follow the next steps to prove that the position representation of the momentum operator can be

derived from the canonical commutation relation $[X, P] = i\hbar$. Take two eigenstates of the position operator, $X\,|\Phi_{x_0}\rangle = x_0\,|\Phi_{x_0}\rangle$ and $X\,|\Phi_{x_1}\rangle = x_1\,|\Phi_{x_1}\rangle$. Using the commutation relation it is clear that

$$\langle\Phi_{x_1}\,|\,[X,P]\,|\,\Phi_{x_0}\rangle = i\hbar\,\delta(x_1 - x_0).$$

Taking into account that X is a Hermitian operator compute separately $\langle\Phi_{x_1}\,|\,XP\,|\,\Phi_{x_0}\rangle$ and $\langle\Phi_{x_1}\,|\,PX\,|\,\Phi_{x_0}\rangle$, expressing the results in terms of $\langle\Phi_{x_1}\,|\,P\,|\,\Phi_{x_0}\rangle$. From this you should be able to get a new expression for $\langle\Phi_{x_1}\,|\,[X,P]\,|\,\Phi_{x_0}\rangle$, which you can compare to the previous $i\hbar\,\delta(x_1 - x_0)$ result and conclude that $(x_1 - x_0)\,\langle\Phi_{x_1}\,|\,P\,|\,\Phi_{x_0}\rangle = i\hbar\,\delta(x_1 - x_0)$. Now remember that the Dirac delta distribution verifies the relation $x\delta'(x) = -\delta(x)$ and use it to infer that the position representation of the momentum operator is $P = -i\hbar\,\mathrm{d}/\mathrm{d}x$.

b) Using the commutation relation $[P, X] = -i\hbar$, and the general commutator property $[A, BC] = B[A, C] + [A, B]C$, prove by induction that

$$[P, X^n] = -i\hbar\,nX^{n-1}.$$

This tells us that the commutator $[P, X^n]$ acts like a derivative. In fact, for any function $f(X)$ expressible as a power series in X we can write compactly

$$[P, f(X)] = -i\hbar\frac{\mathrm{d}f}{\mathrm{d}X}.$$

This is the operator identity that underlines the coordinate representation of the momentum operator, $P = -i\hbar\,\mathrm{d}/\mathrm{d}x$. In an analoguous way, prove that the representation of the position operator in momentum space is $X = i\hbar\,\mathrm{d}/\mathrm{d}p$.

4. A quantum system with only three linearly independent states is described by the following Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 + \delta\mathcal{H} = V_0\begin{pmatrix}(1-\varepsilon) & 0 & 0\\ 0 & 1 & \varepsilon\\ 0 & \varepsilon & 2\end{pmatrix}, \qquad \delta\mathcal{H} = \varepsilon V_0\begin{pmatrix}-1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix},$$

where V_0 is a constant and $\varepsilon \ll 1$ is a very small parameter.

- Write down the eigenvectors and eigenvalues of the unperturbed Hamiltonian, $\mathcal{H}_0$.
- Find the *exact* eigenvalues of $\mathcal{H}$ and expand them as a power series in ε, up to second order.
- Use first- and second-order non-degenerate perturbation theory to find the approximate eigenvalue for the perturbed state arising from the non-degenerate eigenvector of $\mathcal{H}_0$.
- Use degenerate perturbation theory to find the first-order correction to the two initially degenerate eigenvalues. Compare with the exact results.

5. Using the $\boldsymbol{\alpha}$ and β matrices show that the four components of the free-particle Dirac equation are given by the expressions in equation (1.60). Check also that using the Pauli matrices you are able to write the Dirac equation in the two-component form shown in equations (1.62) and (1.63).

6. Prove the relations $\mathbf{pA} - \mathbf{Ap} = -i\hbar\boldsymbol{\nabla}\mathbf{A}$ and $\mathbf{p}\times\mathbf{A} + \mathbf{A}\times\mathbf{p} = -i\hbar\,(\boldsymbol{\nabla}\times\mathbf{A})$. Use the last one to prove $(\boldsymbol{\sigma}\boldsymbol{\pi})^2 = \left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)^2 - \frac{q\hbar}{c}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A})$, where $\boldsymbol{\pi} = \mathbf{p} - q\mathbf{A}$.
Hint: Use the identity $(\mathbf{A}\cdot\boldsymbol{\sigma})\cdot(\mathbf{B}\cdot\boldsymbol{\sigma}) = \mathbf{A}\cdot\mathbf{B} + i\boldsymbol{\sigma}\cdot(\mathbf{A}\times\mathbf{B})$, with $\mathbf{A} = \mathbf{B} = \boldsymbol{\pi}$.

7. The free-particle Dirac Hamiltonian can be written as $\mathcal{H}_{\rm D} = c\boldsymbol{\alpha}\mathbf{p} + \beta mc^2$, where $\beta = \gamma^0$ and $\alpha_k = \gamma^0\gamma^k$, $(k = 1, 2, 3)$. Prove that $[\mathcal{H}_{\rm D}, \mathbf{L}] = -i\hbar c(\boldsymbol{\alpha}\times\mathbf{p})$. This shows that, for a particle satisfying the Dirac equation the orbital angular momentum operator does not commute with the Dirac Hamiltonian, and therefore does not correspond to a conserved quantity. Consider now the spin operator defined as

$$\mathbf{S} = \frac{\hbar}{2}\begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \equiv \frac{\hbar}{2}\boldsymbol{\Sigma},$$

and prove that $[\mathcal{H}_{\rm D}, \mathbf{S}] = i\hbar c(\boldsymbol{\alpha}\times\mathbf{p})$. This, combined with $[\mathcal{H}_{\rm D}, \mathbf{L}] = -i\hbar c(\boldsymbol{\alpha}\times\mathbf{p})$, means that $\mathcal{H}_{\rm D}$ commutes with the total angular momentum operator $\mathbf{J} = \mathbf{L} + \mathbf{S}$, and thus a particle satisfying the Dirac equation has intrinsic angular momentum $S = 1/2$. This is a profound result, since it shows that spin can be considered to emerge as a direct consequence of requiring the wavefunction to satisfy the Dirac equation.

8. Complete the missing steps in the lecture notes to find the wavefunctions of an electron in a central potential which are simultaneously eigenstates of L^2, S^2, J^2 and J_z.

Hint: from the eigenvalue condition $J^2\Psi = j(j+1)\hbar^2\,\Psi$ you can obtain a system of two coupled equations for the radial functions (R_1 and R_2 in the lecture notes notation). The existence of a non-trivial solution requires that the determinant of the coffecients of R_1 and R_2 vanish. From this condition you should be able to obtain that the only possible values of j are $j = \ell \pm 1/2$, which determine the possible values of the ratio R_1/R_2. Finally, normalize the obtained eigenfunctions.

9. Consider the Hydrogen atom with a *modified* electrostatic potential given by

$$V(r) = -\frac{Ze^2}{r}\left(1 - \frac{\beta}{r}\right).$$

- Although $V(r)$ has spherical symmetry, it is not a Coulomb potential, so there is no reason to expect that the spectrum will show the accidental degeneracy characteristic of Hydrogen. Show that the energy levels for this potential are dependent on both n and ℓ according to the following expression (in Ry)

$$E_{n\ell} = -\frac{Z^2}{\left(n - \left(\ell + \frac{1}{2}\right) + \sqrt{\left(\ell + \frac{1}{2}\right)^2 + \gamma}\right)^2},$$

where $\gamma = \frac{2mZe^2\beta}{\hbar^2} = \frac{2Z\beta}{a_0}$ is a dimensionless parameter. For $\gamma = 0$ we recover the Hydrogen spectrum.

- Assume now that β is small, treat $Ze^2\beta/r^2$ as a perturbation and compute the energy levels using first-order perturbation theory. Show your result as a function of γ.
- Discuss the compatibility of the exact and first-order perturbation results.

Hints: You may consider reproducing the line of thought going from equation (1.106) to equation (1.112) with the required modifications. For the second part you may want to use that for the hydrogen eigenstates the expected value of $1/r^2$ is

$$\left\langle \frac{1}{r^2} \right\rangle = \frac{Z^2}{n^3\left(\ell + \frac{1}{2}\right)}$$

10. In his historical 1925 Christmas break in Arosa (Switzerland), aside from formulating his non-relativistic wave equation, Erwin Schrödinger studied a covariant equation for spinless particles,

$$(cp^\mu - eA^\mu)(cp_\mu - eA_\mu)\,\Psi = m^2c^4\Psi,$$

which we now know as the Klein-Gordon equation with electromagnetic interaction, after its rediscovery by Oskar Klein and Walter Gordon in 1926 (Schrödinger never published his finding). For the Hydrogen atom with a pure timelike Coulomb four-potential, $A^\mu = (\frac{e}{r}, 0, 0, 0)$, this equation becomes

$$\left[\left(i\hbar\frac{\partial}{\partial t} - \frac{e^2}{r}\right)^2 + (c\hbar)^2\nabla^2\right]\Psi = \left(mc^2\right)^2\Psi.$$

Use the separation technique and spherical coordinates

$$\Psi = T(t)R(r)Y_{\ell m}(\theta,\varphi) = T(t)\frac{u(r)}{r}Y_{\ell m}(\theta,\varphi),$$

where $T(t) = e^{-\frac{iEt}{\hbar}}$ and $Y_{\ell m}(\theta,\varphi)$ are the spherical harmonics, to solve this equation and find that its energy spectrum is

$$E_{n\ell} = mc^2\left\{\left[1 + \left(\frac{\alpha}{n - \left(\ell+\frac{1}{2}\right) + \sqrt{\left(\ell+\frac{1}{2}\right)^2 - \alpha^2}}\right)^2\right]^{-\frac{1}{2}} - 1\right\},$$

where α is the fine structure constant. Discuss the discrepancies with the equation shown in problem 16. These discrepancies lead to a disagreement with laboratory measurements of the Hydrogen fine structure, therefore Erwin Schrödinger decided to discard the Klein-Gordon equation and publish instead the non-relativistic equation named after him.

Hint: You may consider reproducing the line of thought going from equation (1.91) to equation (1.112) with the required modifications.

11. For the Hydrogen atom calculate the energy splitting originated in the spin-orbit coupling if the electrostatic potential were given by

$$\phi(r) = \frac{Ze}{r}\left(1 - \frac{\beta}{r}\right).$$

This is the same potential considered in problem 9. To use the unmodified Hydrogen atom wave functions we need to assume that β is small. Compare the magnitude of the spin-orbit splitting with the shift in the energy levels at first-perturbative order obtained in that problem.

Hint: You might need the following Hydrogen atom relations in a.u.,

$$\left\langle n\ell \left| \frac{1}{r^3} \right| n\ell \right\rangle = \frac{Z^3}{n^3(\ell+1)\left(\ell+\frac{1}{2}\right)\ell}$$

$$\left\langle n\ell \left| \frac{1}{r^4} \right| n\ell \right\rangle = \frac{Z^4 \frac{1}{2}\left[3n^2 - \ell(\ell+1)\right]}{n^5\left(\ell+\frac{3}{2}\right)(\ell+1)\left(\ell+\frac{1}{2}\right)\ell\left(\ell-\frac{1}{2}\right)}.$$

12. Evaluate the commutators $[\mathbf{LS}, \mathbf{L}]$, $[\mathbf{LS}, \mathbf{S}]$, $[\mathbf{LS}, \mathbf{J}]$, $[\mathbf{LS}, L^2]$, $[\mathbf{LS}, S^2]$ and $[\mathbf{LS}, J^2]$.

13. Compute the matrix elements shown in table (1.3).

14. Check that by adding the three corrections $E_{\rm so}$, $E_{\rm r}$ and $E_{\rm D}$, given by equations (1.171), (1.199) and (1.209), you obtain formula (1.214), which depends only on n and j, but not on ℓ or m.

15. How many protons should the nucleus have (atomic number) to make the relativistic correction to the kinetic energy of a $1s$-electron to be comparable to 1% and 10% of the electron rest mass energy? To do this calculation ignore the effect of the remaining electrons in the atom.

16. The Dirac equation can be solved exactly (without using perturbation theory) to obtain the Hydrogen fine-structure levels. The result is

$$E_{nj} = mc^2 \left\{ \left[1 + \left(\frac{\alpha}{n - \left(j+\frac{1}{2}\right) + \sqrt{\left(j+\frac{1}{2}\right)^2 - \alpha^2}} \right)^2 \right]^{-1/2} - 1 \right\}. \qquad (1.224)$$

Expand this to order α^4 and recover equation (1.215)

17. Compute the value of the energy splitting due to the spin-orbit coupling in the $2p$, $3p$ and $4p$ states of the Hydrogen atom. Express your answers in Ry, eV and cm^{-1}.

18. In your *quantum physics laboratory course* you have seen a very conspicuous red Balmer line in the visible region of the Hydrogen spectrum, coming from the transition $n = 3$ to $n = 2$. Determine the wavelength and frequency of this line using the

unperturbed Hydrogen eigenvalues. Once you take into account the fine structure this line splits into several closely spaced lines. Determine how many and how close are they. Draw an energy level diagram showing all possible transitions from $n = 3$ to $n = 2$, together with the respective energies and frequencies.

19. Consider a state of a multi-electron atom with total electronic spin $S = 2$ and total electronic orbital angular momentum $L = 2$. In analogy with the hydrogenic case, assume that the spin-orbit interaction is described by the Hamiltonian $\mathcal{H}_{\rm so} = A\,\mathbf{LS}$. Show the splitting of this state due to the spin-orbit interaction on an energy level diagram in units of A (assume A to be positive). Indicate the value of J and the energy of each level, ${}^{2S+1}L_J$, of the multiplet. Compute the shift of the multiplet *center of gravity* (mean average energy) with respect to the unperturbed energy of the term (don't forget that for each value of J you have a $2J + 1$ degeneracy).

20. The fluctuations in the electromagnetic field associated with the QED vacuum perturbs the electrical potential that an electron *sees* inside an atom. This perturbation causes fluctuations in the position of the electron. Assuming that the fluctuations are isotropic, the effective potential energy of the electron can be written as the following average

$$\langle V(\mathbf{r}+\delta\mathbf{r})\rangle \approx V(\mathbf{r}) + \frac{1}{6}\left\langle(\delta\mathbf{r})^2\right\rangle \nabla^2 V(\mathbf{r}) = V(\mathbf{r}) + \frac{1}{6}r_0^2\,\nabla^2 V(\mathbf{r}),$$

where $r_0^2 \equiv \left\langle(\delta\mathbf{r})^2\right\rangle$ is the spherical-average quadratic size of the fluctuations, and $V(\mathbf{r}) = -e^2/r$ is the Coulomb energy. Treating the r_0^2 term as a very small perturbation compared to $V(\mathbf{r}) = -e^2/r$, compute the Lamb shifts for the ${}^2\mathrm{S}_{1/2}$ and ${}^2\mathrm{P}_{1/2}$ levels of the hydrogen atom. Leave the result in terms of r_0 and fundamental constants. Compute the value of r_0 taking the Lamb shift for the ${}^2\mathrm{S}_{1/2}$ hydrogen level to be $\Delta E_{\rm Lamb} = 3.36 \times 10^{-5}$ eV.

The unperturbed hydrogen wave functions you should use are

$$\psi_{1s} = \psi_{100}(r,\theta,\varphi) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0}, \qquad \psi_{2pm} = \psi_{21m}(r,\theta,\varphi) = \frac{r}{\sqrt{24 a_0^5}}\, e^{-r/2a_0}\, Y_1^m.$$

21. Compute the wavelength of the photon emitted in the following He^+ transition taking into account the fine structure

$$2p\ {}^2P_{3/2} \longrightarrow 1s\ {}^2S_{1/2}.$$

The measured value of this wavelength is 30.378040982 nm . If your result is not compatible with this value try to explain the origin of the discrepancy. (Note: be sure to use enough significant digits for the value of the Rydbgerg constant).

22. Consider a two-particle system with the Yukawa potential

$$V(r) = -V_0 \frac{e^{-r/a}}{r/a}, \qquad V_0, a > 0.$$

The system has reduced mass m. Use the variational principle and the family of

functions

$$\psi_\lambda(r) = N_\lambda e^{-\lambda r/(2a)}, \qquad \lambda > 0,$$

to address the following items.

a) Show that

$$N_\lambda = \sqrt{\frac{\lambda^3}{8\pi a^3}}.$$

b) Show that for the considered functions the energy mean value is

$$E(\lambda) = \frac{\hbar^2}{2ma^2}\left(\frac{\lambda^2}{4} - \frac{(\kappa/2)\lambda^3}{(1+\lambda)^2}\right),$$

where $\kappa = \frac{2ma^2V_0}{\hbar^2}$.

c) By calculating the derivative of the previous expression demonstrate that the minimal energy condition, achieved at $\lambda = \lambda_0$, is equivalent to

$$(1+\lambda_0)^3 = \kappa\lambda_0(3+\lambda_0).$$

Use this expression to find that the minimum energy estimate is

$$E_0 = \frac{\hbar^2\lambda_0^2}{8ma^2}\left(\frac{1-\lambda_0}{3+\lambda_0}\right).$$

From this, it should be easy to prove that this family of functions predicts a bound state if $\kappa > 2$.

d) If the energy minimum is at $\lambda_0 = 2$ show that

$$E_0 = -\frac{1}{10}\frac{\hbar^2}{ma^2}, \qquad V_0 = \frac{27}{20}\frac{\hbar^2}{ma^2}.$$

Let us suppose that the system under study is the nucleus of deuterium, the hydrogen isotope constituted by a proton and a neutron. The radius of deuterium is: $a = 2.126$ fm. What is E_0 for this system? Compare your result with the experimental value of $E_0 = -2.22$ MeV.

e) Find for deuterium in its estimated ground state the energy shift due to the relativistic kinetic energy correction:

$$\mathcal{H}_r = -\frac{1}{2mc^2}\left(\frac{p^2}{2m}\right)^2.$$

Hint: The next expression (which is not difficult to prove) might be useful:

$$\int_0^\infty r^n e^{-\alpha r} dr = \frac{n!}{\alpha^{n+1}}.$$

23. The Coulomb problem in two dimensions is given by the Hamiltonian

$$\mathcal{H} = \frac{\mathbf{p}^2}{2m} - \frac{Ze^2}{r},$$

where $r = \sqrt{x^2+y^2}$ and $\mathbf{p} = -i\hbar(\partial_x, \partial_y)$. The corresponding Schrödinger equation

in polar plane coordinates is

$$-\frac{\hbar^2}{2m}\left[\frac{\partial^2\psi}{\partial r^2}+\frac{1}{r}\frac{\partial\psi}{\partial r}+\frac{1}{r^2}\frac{\partial^2\psi}{\partial\phi^2}\right]-\frac{Ze^2}{r}\psi=E\psi.$$

Use separation of variables:

$$\psi(r,\phi)=R(r)\frac{e^{im_\ell\phi}}{\sqrt{2\pi}},\qquad m_\ell=0,\pm1,\pm2,...$$

and the ansatz

$$R(r)=\frac{u(r)}{\sqrt{r}},$$

to find the eigenstates of this system and show that the eigenvalues are

$$E_n=-\frac{Z^2e^2}{2a_0}\frac{1}{(n+1/2)^2}=-\frac{Z^2}{(n+1/2)^2}\,R_\infty,$$

where $a_0=\hbar^2/me^2$ is the Bohr radius, R_∞ is the Rydeberg constant and $n=0,1,2,3,....$

Chapter 2

Atoms in static fields and hyperfine interactions.

2.1 Magnetic fields.

Let us consider a hydrogen atom surrounded by a constant magnetic field $\mathbf{B}$. The vector potential can be taken to be[1]

$$\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r}. \tag{2.1}$$

We have seen that the Schrödinger equation for an electron interacting with fields given by the potentials $\mathbf{A}$ and ϕ is

$$\begin{aligned}(E' + e\phi)\,\psi \;=\; \Bigg[\frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 + \frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A}) \\ - \frac{p^4}{8m^3c^2} - \frac{\hbar^2 e}{8m^2c^2}\nabla^2\phi - \frac{\hbar e}{4m^2c^2}\boldsymbol{\sigma}\left[(\boldsymbol{\nabla}\phi)\times\mathbf{p}\right]\Bigg]\,\psi.\end{aligned} \tag{2.2}$$

The interaction terms that depend on the vector potential are

$$\frac{e}{2mc}(\mathbf{pA} + \mathbf{Ap}) + \frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A}) + \frac{e^2}{2mc^2}A^2. \tag{2.3}$$

With $\mathbf{A} = \frac{1}{2}\mathbf{B}\times\mathbf{r}$ we have $\boldsymbol{\nabla}\mathbf{A} = 0$, so[2] $\mathbf{pA} = \mathbf{Ap}$ and the Hamiltonian which describes the interaction with the magnetic field can be written as

$$\mathcal{H}_{\rm m} = \frac{e}{mc}\mathbf{Ap} + \frac{e\hbar}{2mc}\boldsymbol{\sigma}\mathbf{B} + \frac{e^2}{2mc^2}A^2. \tag{2.4}$$

2.1.1 The linear term. The Zeeman and Paschen-Back effects.

Limiting ourselves for the moment only to effects linear in $\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$, the magnetic Hamiltonian can be approximated by

$$\mathcal{H}_{\rm m} \approx \frac{e}{mc}\mathbf{Ap} + \frac{e\hbar}{2mc}\boldsymbol{\sigma}\mathbf{B}. \tag{2.5}$$

The orbital angular momentum operator is $\mathbf{L} = \mathbf{r}\times\mathbf{p}$, and $\mathbf{S} = \dfrac{\hbar}{2}\boldsymbol{\sigma}$, so[3]

$$\mathbf{Ap} = \frac{1}{2}(\mathbf{B}\times\mathbf{r})\mathbf{p} = \frac{1}{2}\mathbf{B}(\mathbf{r}\times\mathbf{p}) = \frac{1}{2}\mathbf{BL} \tag{2.6}$$

[1] If $\mathbf{B}$ is uniform, check that with this choice of $\mathbf{A}$ the relations $\boldsymbol{\nabla}\times\mathbf{A} = \mathbf{B}$ and $\boldsymbol{\nabla}\mathbf{A} = 0$ are verified.

[2] Using a dummy function it is easy to see that

$$(\mathbf{pA} - \mathbf{Ap})\psi = -i\hbar(\boldsymbol{\nabla}A)\psi - i\hbar\mathbf{A}\boldsymbol{\nabla}\psi + i\hbar\mathbf{A}\boldsymbol{\nabla}\psi = -i\hbar(\boldsymbol{\nabla}A)\psi.$$

[3] Remember the properties of the scalar triple product: $(\mathbf{a}\times\mathbf{b})\mathbf{c} = (\mathbf{b}\times\mathbf{c})\mathbf{a} = (\mathbf{c}\times\mathbf{a})\mathbf{b}$.

$$\mathcal{H}_{\rm m} = \frac{e}{2mc}\mathbf{BL} + \frac{e}{mc}\mathbf{SB} = \frac{e}{2mc}\mathbf{B}(\mathbf{L}+2\mathbf{S}) = \frac{1}{\hbar}\mu_{\rm B}\,\mathbf{B}(\mathbf{L}+2\mathbf{S}), \tag{2.7}$$

where the positive constant

$$\mu_{\rm B} = \frac{e\hbar}{2mc} = 9.273\times10^{-21}\,\text{erg/G}, \tag{2.8}$$

is known as the *Bohr magneton*. In the SI system of units, the Bohr magneton is written as[4]

$$\mu_{\rm B} = \frac{e\hbar}{2m} = 9.273\times10^{-24}\,\text{J/T} \;=\; 5.788\times10^{-11}\,\text{MeV/T}. \tag{2.9}$$

Defining (with a minus sign due to the charge of the electron)

$$\boldsymbol{\mu}_{\rm L} = -\frac{\mu_{\rm B}}{\hbar}\mathbf{L}, \qquad \boldsymbol{\mu}_{\rm S} = -\frac{2\mu_{\rm B}}{\hbar}\mathbf{S}, \tag{2.10}$$

we can write

$$\mathcal{H}_{\rm m} = -\boldsymbol{\mu}_{\rm L}\mathbf{B} - \boldsymbol{\mu}_{\rm S}\mathbf{B}. \tag{2.11}$$

This expression is similar to that of the classical energy[5] of a magnetic dipole in a magnetic field, suggesting that $\boldsymbol{\mu}_{\rm L}$ and $\boldsymbol{\mu}_{\rm S}$ should be interpreted as magnetic moment operators associated to $\mathbf{L}$ and $\mathbf{S}$, respectively. The factor 2 which appears in the relation $\boldsymbol{\mu}_{\rm S} = -2\mu_{\rm B}\mathbf{S}/\hbar$ is not expected on classical grounds. On the contrary, the expression $\boldsymbol{\mu}_{\rm L} = -\mu_{\rm B}\mathbf{L}/\hbar$ is exactly what classical electrodynamics would predict for an *orbiting point charge*. It was one of the early triumphs of the Dirac equation to be able to *predict* the factor 2. Higher-order corrections computed in quantum electrodynamics show, however, that this number is not exactly 2, but rather we have

$$\boldsymbol{\mu}_{\rm S} = -\frac{g_e\mu_{\rm B}}{\hbar}\mathbf{S}, \qquad g_e \approx 2.0023, \tag{2.12}$$

where g_e is the gyromagnetic ratio for the electron.[6]

The quantities $\boldsymbol{\mu}_{\rm L}$ and $\boldsymbol{\mu}_{\rm S}$ defined previously are magnetic moment operators. They should not be confused with the corresponding *magnetic moment* values, which we define now. The orbital magnetic moment, $\mu_{\rm L}$, is

$$\mu_{\rm L} = \left\langle \ell, m_\ell = \ell \,\middle|\, \mu_z^L \,\middle|\, \ell, m_\ell = \ell \right\rangle, \tag{2.13}$$

where μ_z^L is the z-component of the operator $\boldsymbol{\mu}_{\rm L}$. This is the matrix element of the third component of $\boldsymbol{\mu}_{\rm L}$ in a state with the maximum value of the third component ($m_\ell = \ell$) of the orbital angular momentum. Since $\boldsymbol{\mu}_{\rm L} = -\mu_{\rm B}\mathbf{L}/\hbar$ and $\mu_z^L = -\mu_{\rm B}L_z/\hbar$, we have

$$\mu_{\rm L} = -\frac{\mu_{\rm B}}{\hbar}\left\langle \ell, m_\ell = \ell \,\middle|\, L_z \,\middle|\, \ell, m_\ell = \ell \right\rangle = -\mu_{\rm B}\ell. \tag{2.14}$$

[4] See Eqs. (B.28) and (B.29) in the appendix on Gaussian and SI units.

[5] The classical energy of an ideal electric dipole $\mathbf{p}$ in an electric field $\mathbf{E}$ is $\mathcal{H}_{\rm e} = -\mathbf{pE}$. Here, $\mathbf{p} = q\mathbf{d}$, $\mathbf{d}$ is the vector connecting the negative charge with the positive charge, and q is the magnitude of the positive charge (equal to the magnitude of the negative charge). Similarly, the classical energy of a magnetic dipole in a magnetic field is given by $\mathcal{H}_{\rm m} = -\boldsymbol{\mu}\mathbf{B}$. For an infinitesimal closed loop current the magnetic dipole moment can be written as $\boldsymbol{\mu} = I\mathbf{a}$, where I is the current and $\mathbf{a}$ the *vector area*, with the direction assigned by the usual right hand rule (fingers in the direction of the current) and the magnitude given by the area of the closed loop.

[6] It is worth to mention that this number is measured to an extraordinary accuracy, which is matched by a similar accuracy in the theoretical calculation using quantum electrodynamics (QED). From reference [14] we may quote the experimental value

$$\frac{\mu_e}{\mu_{\rm B}} - 1 = \frac{g_e - 2}{2} = 0.00115965218991 \pm 0.00000000000026,$$

which is a beautiful testimony of the very high level achieved both in experimental techniques and theoretical understanding.

Similarly, the spin magnetic moment is defined as

$$\mu_{\mathrm{S}} = \left\langle s, m_s = s \,\middle|\, \mu_z^S \,\middle|\, s, m_s = s \right\rangle$$
$$= -\frac{\mu_{\mathrm{B}}}{\hbar} g_e \left\langle s, m_s = s \,\middle|\, S_z \,\middle|\, s, m_s = s \right\rangle = -\mu_{\mathrm{B}} g_e s = -\frac{1}{2}\mu_{\mathrm{B}} g_e \approx -\mu_{\mathrm{B}}. \quad (2.15)$$

The spin eigenvalue $1/2$ approximately cancels the gyromagnetic ratio g_e. Thus the absolute value of the spin magnetic moment of the electron is one Bohr magneton. Using these orbital and spin magnetic moments defined in this way we can write, instead of $\boldsymbol{\mu}_{\mathrm{L}} = -\mu_{\mathrm{B}}\mathbf{L}/\hbar$ and $\boldsymbol{\mu}_{\mathrm{S}} = -2\mu_{\mathrm{B}}\mathbf{S}/\hbar$, the following expressions

$$\boldsymbol{\mu}_{\mathrm{L}} = \frac{\mu_{\mathrm{L}}}{\ell\hbar}\mathbf{L}, \qquad \boldsymbol{\mu}_{\mathrm{S}} = \frac{\mu_{\mathrm{S}}}{s\hbar}\mathbf{S} = \frac{2\mu_{\mathrm{S}}}{\hbar}\mathbf{S} = -\frac{g_e\mu_{\mathrm{B}}}{\hbar}\mathbf{S} \approx -\frac{2\mu_{\mathrm{B}}}{\hbar}\mathbf{S}. \quad (2.16)$$

To simplify future calculations it is convenient to assume the magnetic field to be aligned with the z-axis, $\mathbf{B} = (0, 0, B)$, therefore our Hamiltonian (2.7) reduces to

$$\mathcal{H}_{\mathrm{m}} = \frac{1}{\hbar}\mu_{\mathrm{B}} B(L_z + 2S_z). \quad (2.17)$$

If the energy change due to the application of the magnetic field is small compared with the spin-orbit coupling energy, the field is said to be *weak*. This is the regime of the *Zeeman effect*, while the case of a strong field is called *Paschen-Back* effect, although in the most recent scientific literature there is a tendency to use just the term Zeeman effect in both situations. Historically there was also a distinction between the so-called *normal* and *anomalous* Zeeman effects. The normal effect appears in transitions between atomic terms with zero spin (singlets),[7] producing an energy splitting that could be understood using a classical model developed by Lorentz. The anomalous Zeeman effect arises in transitions between states with non-zero spin. It could only be explained after the discovery of the electron spin and the development of quantum mechanics. Cadmium lamps are often used in academic labs to show the normal Zeeman effect in $^1\mathrm{D}_2 \to {}^1\mathrm{P}_1$ transitions, and the anomalous effect in $^3\mathrm{S}_1 \to {}^3\mathrm{P}_2$.

Let's start with the weak field regime, in which case we shall assume the effects of the spin-orbit coupling to have been already taken into account and we work with the eigenstates in the coupled[8] basis, $|\ell s j m\rangle$. It is necessary to evaluate only matrix elements diagonal in ℓ, s and j, since individual terms like $^2\mathrm{S}_{1/2}$, $^2\mathrm{P}_{1/2}$, $^2\mathrm{P}_{3/2}$, etc. are isolated (different energies) in the weak-field regime. We need the Landé formula (see appendix B.5.5), which is

$$\left\langle \alpha j m' \,\middle|\, \mathbf{A} \,\middle|\, \alpha j m \right\rangle = \frac{\langle \alpha j m \,|\, \mathbf{A}\mathbf{J} \,|\, \alpha j m\rangle}{j(j+1)\hbar^2} \left\langle \alpha j m' \,\middle|\, \mathbf{J} \,\middle|\, \alpha j m \right\rangle, \quad (2.18)$$

where $\mathbf{A}$ is an arbitrary vector operator and $|\alpha j m\rangle$ are eigenstates of the total angular momentum (α stands for any other quantum numbers needed to completely specify the state). In our case we are only interested in the matrix element of the third component of the vector operator $(\mathbf{L} + 2\mathbf{S})\mathbf{J}$, so this formula transforms into

$$\left\langle \ell s j m' \,\middle|\, L_z + 2S_z \,\middle|\, \ell s j m \right\rangle = \frac{\langle \ell s j m \,|\, (\mathbf{L} + 2\mathbf{S})\mathbf{J} \,|\, \ell s j m\rangle}{j(j+1)\hbar^2} \left\langle \ell s j m' \,\middle|\, J_z \,\middle|\, \ell s j m \right\rangle. \quad (2.19)$$

First note that $\langle \ell s j m' \,|\, J_z \,|\, \ell s j m\rangle = \hbar m \delta_{m'm}$. Furthermore,

$$\mathbf{L} = \mathbf{J} - \mathbf{S}, \qquad L^2 = J^2 + S^2 - 2\mathbf{S}\mathbf{J}, \qquad \mathbf{S}\mathbf{J} = \frac{1}{2}(J^2 + S^2 - L^2), \quad (2.20)$$

[7] States with zero spin can only be present in multi-electron atoms, as we will see.

[8] We are dealing with two perturbations simultaneously: Zeeman splitting and fine structure. The appropriate basis to work with is the one associated to the good quantum numbers of the dominant perturbation. If the external magnetic field is weak the dominant perturbation is the spin-orbit interaction, therefore the good quantum numbers are n, ℓ, j and m_j, but not m_ℓ and m_s because in the presence of spin-orbit coupling $\mathbf{L}$ and $\mathbf{S}$ are not separately conserved.

$$(\mathbf{L}+2\mathbf{S})\mathbf{J} = (\mathbf{J}-\mathbf{S}+2\mathbf{S})\mathbf{J} = (\mathbf{J}+\mathbf{S})\mathbf{J} = J^2 + \mathbf{S}\mathbf{J} = \frac{3}{2}J^2 + \frac{1}{2}(S^2 - L^2). \tag{2.21}$$

Therefore

$$\begin{aligned}\langle \ell s j m \,|\, (\mathbf{L}+2\mathbf{S})\mathbf{J} \,|\, \ell s j m\rangle &= \left\langle \ell s j m \,\middle|\, \frac{3}{2}J^2 + \frac{1}{2}(S^2 - L^2) \,\middle|\, \ell s j m \right\rangle \\ &= \frac{3\hbar^2}{2} j(j+1) + \frac{\hbar^2}{2} s(s+1) - \frac{\hbar^2}{2}\ell(\ell+1). \end{aligned} \tag{2.22}$$

And

$$\frac{\langle \ell s j m \,|\, (\mathbf{L}+2\mathbf{S})\mathbf{J} \,|\, \ell s j m\rangle}{j(j+1)\hbar^2} = 1 + \frac{j(j+1) + s(s+1) - \ell(\ell+1)}{2j(j+1)} \equiv g_{\mathrm{J}}, \tag{2.23}$$

where g_{J} is the so–called Landé g factor. So, finally

$$\left\langle \ell s j m' \,\middle|\, L_z + 2S_z \,\middle|\, \ell s j m \right\rangle = \hbar g_{\mathrm{J}} m \delta_{m'm}. \tag{2.24}$$

This result shows that only the diagonal elements are nonzero. The energies in a weak magnetic field, known as *Zeeman levels*, are then given by

$$E_{\mathrm{m}} = \mu_{\mathrm{B}} g_{\mathrm{J}} B m. \tag{2.25}$$

They are proportional to the magnetic quantum number m, something which was to be expected. To use the actual value of the gyromagnetic ratio, g_e, rather than the round value of 2, we need to make the substitution

$$(\mathbf{L}+2\mathbf{S})\mathbf{J} \;\longrightarrow\; (\mathbf{L}+g_e\mathbf{S})\mathbf{J} = (\mathbf{J}-\mathbf{S})\mathbf{J} + g_e\mathbf{S}\mathbf{J} = J^2 + (g_e-1)\mathbf{S}\mathbf{J},$$

which leads to write the Landé factor as

$$g_{\mathrm{J}} = 1 + (g_e - 1)\frac{j(j+1) + s(s+1) - \ell(\ell+1)}{2j(j+1)}. \tag{2.26}$$

The total magnetic moment operator is defined as $\boldsymbol{\mu}_{\mathrm{J}} = -\dfrac{g_{\mathrm{J}}\mu_{\mathrm{B}}}{\hbar}\mathbf{J}$. Taking into account that the total magnetic moment is

$$\mu_{\mathrm{J}} \;=\; \left\langle j, m=j \,\middle|\, \mu_z^J \,\middle|\, j, m=j \right\rangle \;=\; -\frac{\mu_{\mathrm{B}} g_{\mathrm{J}}}{\hbar} \langle j, m=j \,|\, J_z \,|\, j, m=j\rangle \;=\; -g_{\mathrm{J}}\mu_{\mathrm{B}} j, \tag{2.27}$$

we can write the Landé factor as $g_{\mathrm{J}} = -\mu_{\mathrm{J}}/(\mu_{\mathrm{B}} j)$, so that the magnetic moment operator becomes

$$\boldsymbol{\mu}_{\mathrm{J}} = -\frac{g_{\mathrm{J}}\mu_{\mathrm{B}}}{\hbar}\mathbf{J}, \qquad g_{\mathrm{J}} = -\frac{\mu_{\mathrm{J}}}{\mu_{\mathrm{B}} j}, \qquad \boldsymbol{\mu}_{\mathrm{J}} = \frac{\mu_{\mathrm{J}}}{j\hbar}\mathbf{J}. \tag{2.28}$$

The Hamiltonian which describes the interaction of the electron with the magnetic field is then

$$\mathcal{H}_{\mathrm{m}} = \frac{1}{\hbar}\mu_{\mathrm{B}} B(L_z + 2S_z) = -\boldsymbol{\mu}_{\mathrm{J}}\mathbf{B}. \tag{2.29}$$

Comparing this with our previous relation $\mathcal{H}_{\mathrm{m}} = -\boldsymbol{\mu}_{\mathrm{L}}\mathbf{B} - \boldsymbol{\mu}_{\mathrm{S}}\mathbf{B}$ we see that $\boldsymbol{\mu}_{\mathrm{J}} = \boldsymbol{\mu}_{\mathrm{L}} + \boldsymbol{\mu}_{\mathrm{S}}$,

$$\boldsymbol{\mu}_{\mathrm{J}} = \frac{1}{\hbar}\mu_{\mathrm{B}}(\mathbf{L} + 2\mathbf{S}), \tag{2.30}$$

so all our definitions make sense.

For an electron in a s state, $^2\mathrm{S}_{1/2}$, we have $\ell = 0$, $s = 1/2$, $j = 1/2$ and $g_{\mathrm{J}} = 2$, so the energies of the Zeeman levels given by $E_{\mathrm{m}} = g_{\mathrm{J}}\mu_{\mathrm{B}} B m$ are $E_{\mathrm{m}} = \pm\mu_{\mathrm{B}} B$, for $m = \pm 1/2$. For

	$\left\vert 1\frac{1}{2}\right\rangle$	$\left\vert 1-\frac{1}{2}\right\rangle$	$\left\vert 0\frac{1}{2}\right\rangle$	$\left\vert 0-\frac{1}{2}\right\rangle$	$\left\vert -1\frac{1}{2}\right\rangle$	$\left\vert -1-\frac{1}{2}\right\rangle$
$\left\langle 1\frac{1}{2}\right\vert$	$\frac{\hbar^2}{2}\xi_{n\ell}+2\mu_{\rm B}B$					
$\left\langle 1-\frac{1}{2}\right\vert$		$-\frac{\hbar^2}{2}\xi_{n\ell}$	$\frac{\hbar^2}{\sqrt{2}}\xi_{n\ell}$			
$\left\langle 0\frac{1}{2}\right\vert$		$\frac{\hbar^2}{\sqrt{2}}\xi_{n\ell}$	$\mu_{\rm B}B$			
$\left\langle 0-\frac{1}{2}\right\vert$				$-\mu_{\rm B}B$	$\frac{\hbar^2}{\sqrt{2}}\xi_{n\ell}$	
$\left\langle -1\frac{1}{2}\right\vert$				$\frac{\hbar^2}{\sqrt{2}}\xi_{n\ell}$	$-\frac{\hbar^2}{2}\xi_{n\ell}$	
$\left\langle -1-\frac{1}{2}\right\vert$						$\frac{\hbar^2}{2}\xi_{n\ell}-2\mu_{\rm B}B$

Table 2.1: $\mathcal{H}'_{\rm m}$ matrix elements for p ($\ell=1$) states. The notation is shortened to $|m_\ell, m_s\rangle$.

a $^2\mathrm{P}_{1/2}$ term $g_{\rm J}=2/3$ and there will be also a splitting in two levels, while for a $^2\mathrm{P}_{3/2}$ term $g_{\rm J}=4/3$ and the splitting is in four different levels.

$$E_{\rm m}(^2\mathrm{S}_{1/2}) = \pm\mu_{\rm B}B \tag{2.31}$$

$$E_{\rm m}(^2\mathrm{P}_{1/2}) = \pm\frac{1}{3}\mu_{\rm B}B \tag{2.32}$$

$$E_{\rm m}(^2\mathrm{P}_{3/2}) = \pm 2\mu_{\rm B}B,\ \pm\frac{2}{3}\mu_{\rm B}B \tag{2.33}$$

If the strength of the magnetic field is increased to the point where the splitting becomes comparable to the spin-orbit coupling, it is no longer possible to isolate a single term with a specific value of j, because terms with different values of j may have the same (or nearly the same) energy. In such a case we should include both the spin-orbit and the magnetic field interaction in the perturbation Hamiltonian

$$\mathcal{H}'_{\rm m} = \xi(r)\mathbf{LS} + \frac{\mu_{\rm B}}{\hbar}B(L_z+2S_z). \tag{2.34}$$

The matrix elements are now more easily calculated in the $|\ell s m_\ell m_s\rangle$ basis, since then

$$\left\langle \ell' s m'_\ell m'_s \right| (L_z+2S_z) \left| \ell s m_\ell m_s \right\rangle = \hbar(m_\ell+2m_s)\delta_{\ell'\ell}\delta_{m'_\ell m_\ell}\delta_{m'_s m_s}, \tag{2.35}$$

and

$$\left\langle \ell' s m'_\ell m'_s \right| \mathbf{LS} \left| \ell s m_\ell m_s \right\rangle = \left\langle \ell' s m'_\ell m'_s \right| (-L_{+1}S_{-1}+L_0S_0-L_{-1}S_{+1}) \left| \ell s m_\ell m_s \right\rangle. \tag{2.36}$$

For the multiplet of p states we have already collected the **LS** matrix elements in table (1.3). It is now straightforward to update the table including the magnetic field interactions given by equation (2.35). This is shown in table (2.1). According to the structure of this table it is clear that the secular equation factors out into the following four equations, where $\xi_{n\ell}=\langle n\ell|\xi(r)|n\ell\rangle$,

$$\left(2\mu_{\rm B}B+\frac{\hbar^2}{2}\xi_{n\ell}-E\right)=0, \tag{2.37}$$

$$\begin{vmatrix} -\frac{\hbar^2}{2}\xi_{n\ell}-E & \frac{\hbar^2}{\sqrt{2}}\xi_{n\ell} \\ \frac{\hbar^2}{\sqrt{2}}\xi_{n\ell} & \mu_{\rm B}B-E \end{vmatrix}=0, \tag{2.38}$$

$$\begin{vmatrix} -\mu_{\rm B}B-E & \frac{\hbar^2}{\sqrt{2}}\xi_{n\ell} \\ \frac{\hbar^2}{\sqrt{2}}\xi_{n\ell} & -\frac{\hbar^2}{2}\xi_{n\ell}-E \end{vmatrix}=0, \tag{2.39}$$

$$\left(\frac{\hbar^2}{2}\xi_{n\ell}-2\mu_{\rm B}B-E\right)=0. \tag{2.40}$$

The energy eigenvalues are then,

$$\frac{E_1}{\hbar^2\xi_{n\ell}} = 2\frac{\mu_{\rm B}B}{\hbar^2\xi_{n\ell}} + \frac{1}{2} \tag{2.41}$$

$$\frac{E_{2,3}}{\hbar^2\xi_{n\ell}} = \frac{1}{2}\left(\frac{\mu_{\rm B}B}{\hbar^2\xi_{n\ell}} - \frac{1}{2}\right) \pm \frac{1}{2}\sqrt{\left(\frac{\mu_{\rm B}B}{\hbar^2\xi_{n\ell}}\right)^2 + \left(\frac{\mu_{\rm B}B}{\hbar^2\xi_{n\ell}}\right) + \frac{9}{4}} \tag{2.42}$$

$$\frac{E_{4,5}}{\hbar^2\xi_{n\ell}} = -\frac{1}{2}\left(\frac{\mu_{\rm B}B}{\hbar^2\xi_{n\ell}} + \frac{1}{2}\right) \pm \frac{1}{2}\sqrt{\left(\frac{\mu_{\rm B}B}{\hbar^2\xi_{n\ell}}\right)^2 - \left(\frac{\mu_{\rm B}B}{\hbar^2\xi_{n\ell}}\right) + \frac{9}{4}} \tag{2.43}$$

$$\frac{E_6}{\hbar^2\xi_{n\ell}} = -2\frac{\mu_{\rm B}B}{\hbar^2\xi_{n\ell}} + \frac{1}{2}. \tag{2.44}$$

Now we can study these energies as a function of the magnetic field, B. This is shown in figure (2.1). The following interesting regimes are visible.

- Strong field, $\mu_{\rm B}B \gg \hbar^2\xi_{n\ell}$. This is the *Paschen–Back* region, where the spin-orbit coupling is negligible compared to the magnetic interaction. Remember that in this situation the *good* quantum numbers are n, ℓ, m_ℓ, m_s but not j and m_j, because in the presence of an external torque caused by the magnetic field the total angular momentum is not conserved, whereas L_z and S_z are. Ignoring the fine structure we can assume the energies to satisfy the relation $E_{\rm m} = \mu_{\rm B}B(m_\ell + 2m_s)$. The hydrogenic energy levels will then be

$$E_{n\ell m_\ell m_s} = \frac{E_1^{(0)}}{n^2} + \mu_{\rm B}B(m_\ell + 2m_s). \tag{2.45}$$

  In the strong field regime the fine-structure Hamiltonian $\mathcal{H}_{\rm fs} = \mathcal{H}_{\rm so} + \mathcal{H}_{\rm r} + \mathcal{H}_{\rm D}$ can be considered to produce a small perturbation in the *unperturbed* energies $E_{n\ell m_\ell m_s}$ given by the previous expression. Although the states $|\ell s m_\ell m_s\rangle$ are degenerate (their energy does not depend on ℓ and states with the same value of $m_\ell + 2m_s$ have the same energy), it can be shown that the fine-structure correction to the unperturbed energies (2.45) can be computed with ordinary first-order non-degenerated perturbation theory using the $|\ell s m_\ell m_s\rangle$ states.[9] Thus

$$E_{\rm fs}^{(1)} = \langle \ell s m_\ell m_s \,|\, \mathcal{H}_{\rm fs} \,|\, \ell s m_\ell m_s\rangle = \langle \ell s m_\ell m_s \,|\, \mathcal{H}_{\rm so} + \mathcal{H}_{\rm r} + \mathcal{H}_{\rm D} \,|\, \ell s m_\ell m_s\rangle\,. \tag{2.46}$$

  The relativistic and the Darwin contributions are still given by (see equations (1.198) and (1.210)).

$$E_{\rm r} = E_n^{(0)}\frac{(Z\alpha)^2}{n}\left[-\frac{3}{4n} + \frac{1}{(\ell+\frac{1}{2})}\right] = -E_1^{(0)}\frac{(Z\alpha)^2}{n^3}\left[\frac{3}{4n} - \frac{1}{(\ell+\frac{1}{2})}\right], \tag{2.47}$$

$$E_{\rm D} = -E_n^{(0)}\frac{(Z\alpha)^2}{n} = -E_1^{(0)}\frac{(Z\alpha)^2}{n^3}, \quad (\ell = 0 \ \text{ only}). \tag{2.48}$$

  For the spin-orbit term we need to evaluate

$$\langle \mathbf{LS}\rangle = \langle L_x\rangle\langle S_x\rangle + \langle L_y\rangle\langle S_y\rangle + \langle L_z\rangle\langle S_z\rangle = \langle L_z\rangle\langle S_z\rangle = \hbar^2 m_\ell m_s. \tag{2.49}$$

[9] In appendix B.2 (see equations (B.52) and (B.53)) it is shown that if we can find an operator $\mathcal{A}$ which commutes with both the unperturbed Hamiltonian $\mathcal{H}_0$ and the perturbation $\mathcal{H}'$, then we can apply non-degenerate perturbation theory using the simultaneous eigenstates of $\mathcal{A}$ and $\mathcal{H}_0$. In the particular case of the strong-field Zeeman effect, we can choose the operators L^2 and J_z, which commute with $\mathcal{H}_0$ and $\mathcal{H}' = \mathcal{H}_{\rm fs}$, to play the role of $\mathcal{A}$. Note that L^2 removes the degeneracy in ℓ, while J_z removes the degeneracy in $m_\ell + 2m_s = m_\ell + m_j$, and the $|\ell s m_\ell m_s\rangle$ states are simultaneous eigenstates of $\mathcal{H}_0$, L^2 and J_z.

Note that $\langle L_x\rangle = \langle S_x\rangle = \langle L_y\rangle = \langle S_y\rangle = 0$ for eigenstates of S_z and L_z. Since the expectation value of $\xi(r)$ is given by equation (1.167), we have

$$E_{\rm so} = \frac{Z^4 e^2 \hbar^2}{2m^2 c^2 a_0^3 n^3} \frac{m_\ell m_s}{\ell(\ell+\frac{1}{2})(\ell+1)} = -E_1^{(0)} \frac{(Z\alpha)^2}{n^3} \frac{m_\ell m_s}{\ell(\ell+\frac{1}{2})(\ell+1)}, \qquad (\ell \neq 0 \text{ only}). \quad (2.50)$$

Adding everything we get

$$\begin{aligned} E_{\rm fs}^{(1)} &= E_{\rm r} + E_{\rm so} + E_{\rm D} \\ &= -E_1^{(0)} \frac{(Z\alpha)^2}{n^3} \left[\frac{3}{4n} - \frac{1}{(\ell+\frac{1}{2})} + \frac{m_\ell m_s}{\ell(\ell+\frac{1}{2})(\ell+1)} (1-\delta_{\ell 0}) + \delta_{\ell 0} \right]. \end{aligned} \quad (2.51)$$

The first $(1-\delta_{\ell 0})$ factor makes sure the spin–orbit contribution is computed only for states with $\ell \neq 0$. The second $\delta_{\ell 0}$ represents the Darwin contribution, which is non–zero only for $\ell = 0$ states. The total $E_{n\ell m_\ell m_s}$ energy is given by

$$\begin{aligned} E_{n\ell m_\ell m_s} = \frac{E_1^{(0)}}{n^2} \left[1 - \frac{Z^2\alpha^2}{n} \left(\frac{3}{4n} - \frac{1}{(\ell+\frac{1}{2})} + \frac{m_\ell m_s}{\ell(\ell+\frac{1}{2})(\ell+1)} (1-\delta_{\ell 0}) + \delta_{\ell 0} \right) \right] \\ + \mu_{\rm B} B (m_\ell + 2m_s). \end{aligned} \quad (2.52)$$

- Weak field, $\mu_{\rm B} B \ll \hbar^2 \xi_{n\ell}$. If we consider only linear terms in $\mu_{\rm B} B$, the energies can now be approximated by

$$E_1 = \frac{1}{2}\hbar^2\xi_{n\ell} + 2\mu_{\rm B}B, \qquad E_2 = \frac{1}{2}\hbar^2\xi_{n\ell} + \frac{2}{3}\mu_{\rm B}B, \qquad E_3 = -\hbar^2\xi_{n\ell} + \frac{1}{3}\mu_{\rm B}B. \quad (2.53)$$

$$E_4 = \frac{1}{2}\hbar^2\xi_{n\ell} - \frac{2}{3}\mu_{\rm B}B, \qquad E_5 = -\hbar^2\xi_{n\ell} - \frac{1}{3}\mu_{\rm B}B, \qquad E_6 = \frac{1}{2}\hbar^2\xi_{n\ell} - 2\mu_{\rm B}B. \quad (2.54)$$

Here we recognize again the energies for the splitting in a weak magnetic field given by equations (2.32) and (2.33), together with the energies for the spin-orbit interaction with $\ell = 1$.

Figure 2.2 shows the correlation between the spin-orbit weak-field and strong-field splitting for a p state. In the weak-field regime $m_j = m_\ell + m_s$ is a *good* quantum number and the system is best described in the coupled representation. In a classical analogy, one would say that the spin and the orbital angular momentum are coupled to produce a total angular momentum, j, which precesses as a whole about the applied magnetic field. On the other hand, when the field is strong the good quantum numbers are m_ℓ and m_s separately, as appropriate for the uncoupled basis. In classical terms, in this case the spin angular momentum and the orbital angular momentum are each individually precessing about the magnetic field. See figure (2.3) for an illustration.

Note that in the presence of an external torque the total angular momentum is not conserved. When the atom is placed in a magnetic field, the Hamiltonian is no longer invariant under all three-dimensional rotations but only under rotations about an axis parallel to the magnetic field. The consequence of this reduced symmetry is that the Hamiltonian no longer commutes with J^2, although it does commute with J_z. In other words, in a magnetic field j is not a good quantum number, but m is. When these features are fully realized the field is regarded as strong. The field is considered weak if, to a good approximation, j is still a good quantum number.

Atoms can be exposed to electromagnetic radiation to produce electronic transitions between magnetic substates. These are known as electron spin resonance (ESR) experiments (also electron paramagnetic resonance, EPR). For an s state the energies are given by $E_{\rm m} = \pm\mu_{\rm B} B$, therefore we need a photon energy of

$$E_\gamma = \hbar\omega = 2\mu_{\rm B}B = 2\frac{e\hbar}{2m_e c}B, \qquad \nu = \frac{\omega}{2\pi} = \frac{e}{2\pi m_e c}B \quad (2.55)$$

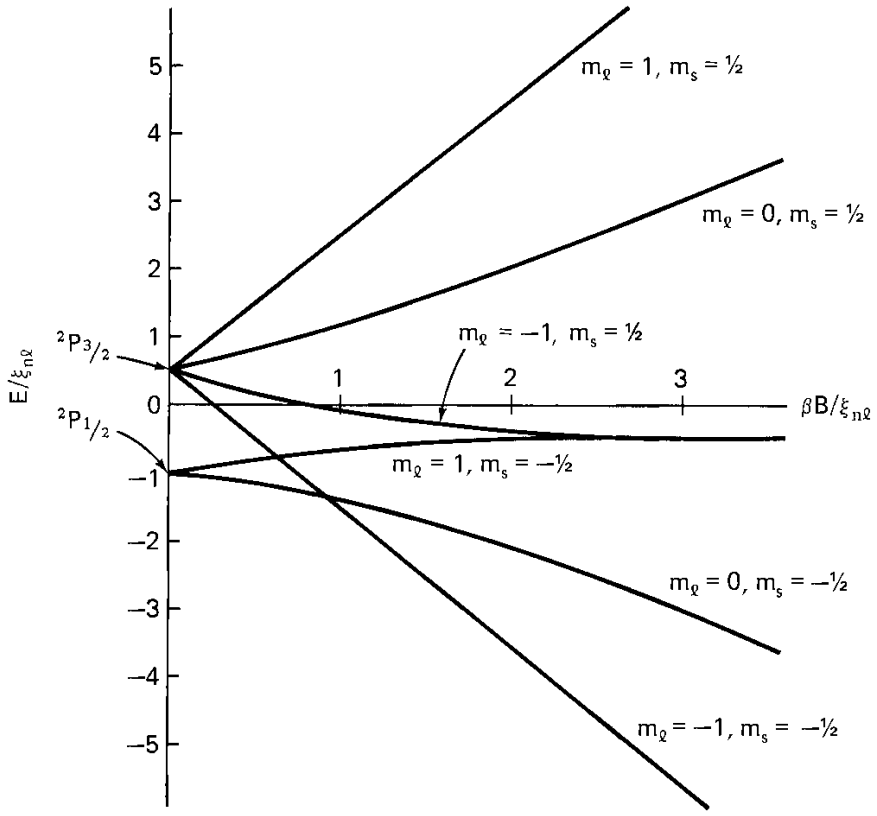


Figure 2.1: Transition from a weak to a strong magnetic field for a ^{2}P term (figure taken from Weissbluth, 1978; the parameter β is the Bohr magneton, μ_B).

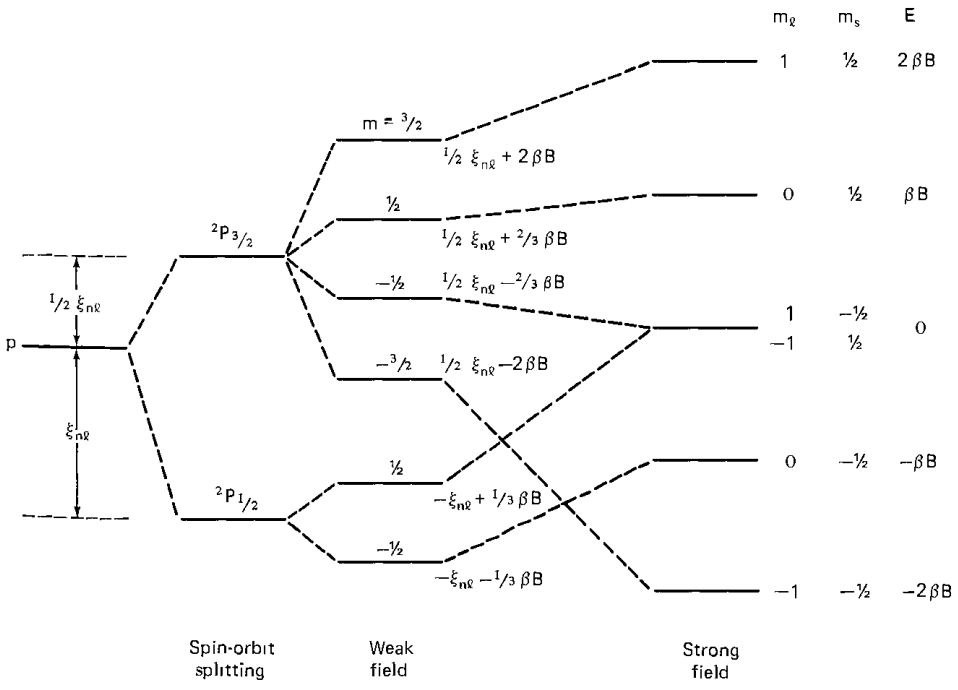


Figure 2.2: Splitting of a p state under the influence of spin-orbit coupling and a magnetic field (figure taken from Weissbluth, 1978; the parameter β is the Bohr magneton, μ_B).

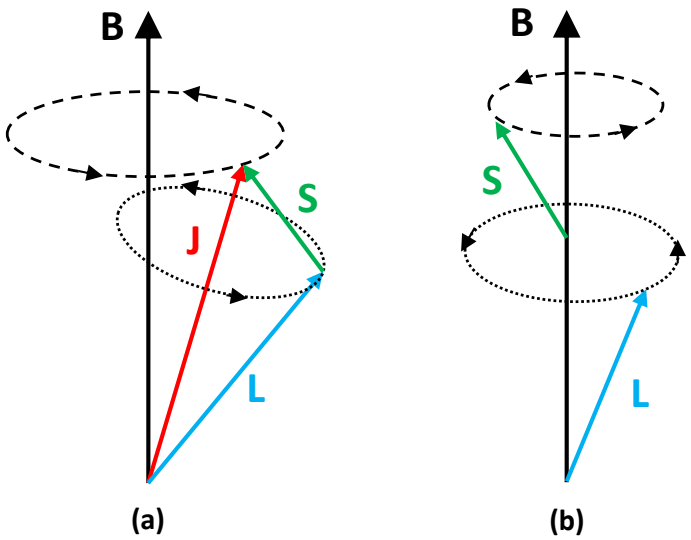


Figure 2.3: Precession of **L**, **S**, **J** in a weak (a) and strong (b) magnetic field.

To end this section, let us see how the Hamiltonian of the spin-orbit interaction can be used to make an estimation of the internal magnetic field in the Hydrogen atom. Based on this, we can make an estimation of how large should an external magnetic field be to cause a *strong* Zeeman effect. Furthermore, We will use this example to show how to convert expressions from the Gaussian system of units to the SI system.

The Hamiltonian of the spin-orbit interaction in Hydrogen is

$$\mathcal{H}_{\rm so} = \frac{e^2}{2m^2c^2r^3}\,\mathbf{LS} = \left(\frac{e}{m}\mathbf{S}\right)\frac{e}{2mc^2r^3}\,\mathbf{L} = -\boldsymbol{\mu}\mathbf{B}, \tag{2.56}$$

where $\boldsymbol{\mu} = -2\mu_{\rm B}\mathbf{S}/\hbar = -e\mathbf{S}/mc = -e\hbar/2mc$ is the electron magnetic moment and $\mathbf{B}$ is the magnetic field created by the proton at the position of the electron. To make an estimation of the value of the magnetic field we could take $L = \hbar$ and $r = a_0$, so that

$$\mathcal{H}_{\rm so} = \left(\frac{e\hbar}{2mc}\right)\frac{e\hbar}{2mca_0^3} = -\boldsymbol{\mu}\mathbf{B} \quad\Longrightarrow\quad B = \frac{e\hbar}{2mca_0^3} = \frac{\alpha\hbar^2}{2mea_0^3}\ \text{Gauss}. \tag{2.57}$$

In the last step we used the expression for the (dimensionless) structure fine constant in Gaussian units, $\alpha = e^2/\hbar c$. In this system of units, the absolute value of the electric charge of the electron is

$$e = \sqrt{\alpha\hbar c} = \sqrt{\frac{1}{137}\times(1.055\times10^{-27}\,\text{erg}\cdot\text{s})\times(2.9979\times10^{10}\,\text{cm}\cdot\text{s}^{-1})}$$
$$= 4.80\times10^{-10}\ \text{statcoulomb}, \tag{2.58}$$

so the value of the magnetic field is

$$B = \frac{\frac{1}{137}\times(1.055\times10^{-27}\,\text{erg}\cdot\text{s})^2}{2\times(9.109\times10^{-28}\,\text{gr})(4.80\times10^{-10}\,\text{statcoul})(5.291\times10^{-9}\,\text{cm})^3} = 6.2\times10^4\,\text{G}. \tag{2.59}$$

It is interesting to work out in detail the conversion from Gaussian to SI units. The electric charge and the magnetic field in both system of units are related through the expressions (see equations (B.12) and (B.16) in the appendix on Gaussian units)

$$e_{\rm g} = \frac{e_{\rm si}}{\sqrt{4\pi\epsilon_0}}, \qquad B_{\rm g} = \sqrt{\frac{4\pi}{\mu_0}}B_{\rm si}. \tag{2.60}$$

Therefore,

$$B_{\rm si} = \sqrt{\frac{\mu_0}{4\pi}} B_{\rm g} = \sqrt{\frac{\mu_0}{4\pi}} \frac{e_{\rm g}\hbar}{2mca_0^3} = \sqrt{\frac{\mu_0}{4\pi}} \frac{e_{\rm si}}{\sqrt{4\pi\epsilon_0}} \frac{\hbar}{2mca_0^3} = \frac{\sqrt{\mu_0\epsilon_0}}{4\pi\epsilon_0} \frac{e_{\rm si}\hbar}{2mca_0^3} = \frac{1}{4\pi\epsilon_0} \frac{e_{\rm si}\hbar}{2mc^2a_0^3}. \quad (2.61)$$

In the last step we have used the relation $\sqrt{\mu_0\epsilon_0} = 1/c$. Therefore, in the SI system of units the expression to use is

$$B_{\rm si} = \frac{1}{4\pi\epsilon_0} \frac{e_{\rm si}\hbar}{2mc^2a_0^3} = \frac{\alpha\hbar^2}{2me_{\rm si}ca_0^3} \text{ Tesla}, \quad (2.62)$$

where $\alpha = e_{\rm si}^2/(4\pi\epsilon_0\hbar c)$ is the fine structure constant written in the SI system. If we compare the last expression for $B_{\rm si}$ with the one for $B_{\rm g}$ given by (2.57), without paying proper attention to the electric charge, we might conclude that $B_{\rm g}/B_{\rm si} = c$, but this is wrong. The correct comparison to make is

$$\frac{B_{\rm g}}{B_{\rm si}} = \frac{e_{\rm si}}{e_{\rm g}}\, c = \sqrt{4\pi\epsilon_0}\, c = \sqrt{4\pi\epsilon_0}\, \frac{1}{\sqrt{\mu_0\epsilon_0}} = \sqrt{\frac{4\pi}{\mu_0}}. \quad (2.63)$$

Putting numbers in the SI system,

$$B = \frac{\frac{1}{137}(1.05\times10^{-34}\,\mathrm{J\cdot s})^2}{2(9.1\times10^{-31}\,\mathrm{kg})(1.6\times10^{-19}\,\mathrm{C})(3\times10^{8}\,\mathrm{m/s})(0.53\times10^{-10}\,\mathrm{m})^3} = 6.2\,\mathrm{T}. \quad (2.64)$$

A strong Zeeman field would be one such that, for example, $B_{\rm ext} \gg 10\,\mathrm{T}$, and a weak field one such that $B_{\rm ext} \ll 10\,\mathrm{T}$. Note that the Earth's magnetic field is about 10^{-4} T, which is certainly weak in the Zeeman effect sense.

2.1.2 The quadratic term. Diamagnetism.

Let us study now the quadratic term in (2.4), $\frac{e^2}{2m_ec^2}A^2$. For a vector potential given by $\mathbf{A} = \frac{1}{2}\mathbf{B}\times\mathbf{r}$ this term is

$$\mathcal{H}_{\rm d} = \frac{e^2}{2m_ec^2}\frac{1}{4}(\mathbf{B}\times\mathbf{r})(\mathbf{B}\times\mathbf{r}) = \frac{e^2}{8m_ec^2}B^2r^2\sin^2\theta, \quad (2.65)$$

where θ is the angle between $\mathbf{B}$ and $\mathbf{r}$. This is the so-called *diamagnetic term*, while that of equation (2.7) is the *paramagnetic* term.[10] Since $x = r\sin\theta\cos\varphi$ and $y = r\sin\theta\sin\varphi$, for any spherically symmetric state we have

$$\left\langle r^2\sin^2\theta\right\rangle = \left\langle x^2+y^2\right\rangle = \left\langle r^2-z^2\right\rangle = \left\langle r^2\right\rangle - \left\langle z^2\right\rangle = \left\langle r^2\right\rangle - \frac{1}{3}\left\langle r^2\right\rangle = \frac{2}{3}\left\langle r^2\right\rangle,$$

and the expectation value of $\mathcal{H}_{\rm d}$ would be

$$\langle\mathcal{H}_{\rm d}\rangle = E_{\rm d} = \frac{e^2}{8m_ec^2}B^2\left\langle r^2\sin^2\theta\right\rangle = \frac{e^2}{12m_ec^2}B^2\left\langle r^2\right\rangle, \quad (2.66)$$

where r is the distance of the electron to the nucleus.

To put numbers it is convenient to use the relations $a_0 = \hbar/(\alpha m_ec)$ and $\mu_{\rm B} = e\hbar/(2m_ec)$ to write $ea_0 = 2\mu_{\rm B}/\alpha$. Then, for $B = 1$ Tesla (10^4 Gauss) and the hydrogen ground-state, where $\langle r^2\rangle = 3a_0^2/2$, the diamagnetic energy would be

$$(E_{\rm d})_{\rm H_{gs}} = \frac{1}{8}\frac{(ea_0B)^2}{m_ec^2} = \frac{1}{2}\frac{(\mu_{\rm B}B)^2}{\alpha^2m_ec^2} = \frac{1}{2}\frac{(5.788\times10^{-5}\,\mathrm{eV\cdot T^{-1}}\times1\,\mathrm{T})^2}{\left(\frac{1}{137}\right)^2\times511\times10^3\,\mathrm{eV}} = 6.15\times10^{-11}\,\mathrm{eV}.$$

[10] The paramagnetic term gives rise to the alignment of magnetic moments in an external magnetic field, which leads to paramagnetism.

We see that, for ordinary magnetic field values, the diamagnetic energy is really small. The ratio of $E_{\rm d}$ to the energy of the Zeeman levels given by equation (2.25), $E_{\rm m} = \mu_{\rm B} g_{\rm J} B m$, is

$$\frac{E_{\rm d}}{E_{\rm m}} = \frac{e^2}{12 m_e c^2} \frac{B^2 \left\langle r^2 \right\rangle}{\mu_{\rm B} g_{\rm J} B m} = \frac{1}{12 g_{\rm J} m} \frac{e^2 \left\langle r^2 \right\rangle}{\mu_{\rm B} m_e c^2} B. \tag{2.67}$$

To estimate the size of this ratio let's put in numbers again for a magnetic field of 1 Tesla and a $\ell = 0$ ($g_{\rm J} m = 1$) hydrogenic state. From equation (1.132) we have

$$\langle r^2 \rangle_{n\ell} = \frac{a_0^2}{Z^2} \frac{n^2}{2} \left[5n^2 + 1 - 3\ell(\ell+1) \right] \approx \frac{5 n^4 a_0^2}{2 Z^2}. \tag{2.68}$$

Therefore

$$\frac{E_{\rm d}}{E_{\rm m}} = \frac{5n^4}{24 Z^2} \frac{e^2 a_0^2}{\mu_{\rm B} m_e c^2} B = \frac{5n^4}{6 Z^2} \frac{\mu_{\rm B} B}{\alpha^2 m_e c^2} = \frac{5n^4}{6Z^2} \times (2.13 \times 10^{-6}). \tag{2.69}$$

Since laboratory magnetic fields do not usually exceed a few tens of tesla, the diamagnetic term can be safely neglected unless n is large. Nevertheless, it is worth noting that in some astrophysical environments, such as at the surface of neutron stars, magnetic fields as large as 10^8 Tesla are believed to exist. In such situations the quadratic diamagnetic term cannot be ignored.

The magnetic moment may be taken as[11]

$$\mu = -\frac{\partial E_{\rm d}}{\partial B} = -\frac{e^2}{6 m_e c^2} B \left\langle r^2 \right\rangle, \tag{2.70}$$

thus the diamagnetic susceptibility is

$$\chi_{\rm d} = \frac{\mu}{B} = -\frac{e^2}{6 m_e c^2} \left\langle r^2 \right\rangle = -\frac{1}{6} \alpha^2 a_0 \left\langle r^2 \right\rangle. \tag{2.71}$$

The diamagnetic susceptibility is a measure of how much a material will become (dia)magnetized in an externally applied magnetic field. Diamagnetism is due to the orbital motion of electrons, therefore it is a general property of all atoms and molecules. However, paramagnetism arises from the alignment of the magnetic moments associated to the orbital and spin angular momenta in an applied external magnetic field. It is not uncommon for a material to have both a diamagnetic and paramagnetic susceptibility.

2.2 Electric fields.

2.2.1 Linear Stark effect.

External electric fields applied to atoms give rise to the so-called *Stark effect*. In general, this is a second-order effect, except for hydrogen, as we will see in this section. The interaction of an electron with an external electric potential $\phi(\mathbf{r})$ gives it an extra energy of $-e\phi(\mathbf{r})$. Since atoms are very small compared with the scales over which the external macroscopic potential varies, we can replace $\phi(\mathbf{r})$ with the first two terms in its Taylor series expansion, $\phi(\mathbf{r}) = \phi(0) + [\boldsymbol{\nabla}\phi(0)]\mathbf{r}$. Setting the (arbitrary) value of $\phi(\mathbf{r})$ at the position $\mathbf{r} = 0$ of the atomic nucleus equal to zero, this gives $\phi(\mathbf{r}) = [\boldsymbol{\nabla}\phi(0)]\mathbf{r} = -\boldsymbol{\mathcal{E}}\mathbf{r}$, where $\boldsymbol{\mathcal{E}} \equiv -\boldsymbol{\nabla}\phi(0)$ is the electric field at the nucleus.[12] Choosing the z-axis as the direction of the electric field, the Hamiltonian of the interaction may then be taken as

$$\mathcal{H}_s = -e\phi(\mathbf{r}) = e\mathcal{E}z = e\mathcal{E}r\cos\theta = \sqrt{\frac{4\pi}{3}}\, Y_1^0\, e\mathcal{E}r, \tag{2.72}$$

[11] Remember that the energy of a magnetic dipole immersed in an external magnetic field is $E = -\boldsymbol{\mu}\mathbf{B} = -\mu_x B_x - \mu_y B_y - \mu_z B_z$. If the direction of $\mathbf{B}$ is along the z-axis, $E = -\mu_z B_z$ and we may just take the magnetic moment to be $\mu = -\partial E/\partial B$.

[12] Note that the electric field is written in calligraphic font, $\boldsymbol{\mathcal{E}}$ and $\mathcal{E}$, to avoid confusions with the energy, E.

where we have written $\cos\theta$ in terms of the Y_1^0 spherical harmonic. For the strong-field regime, where the Stark effect is large compared to the spin-orbit splittings, the uncoupled basis is the appropriate one to perform the calculations. Let's compute then the $\mathcal{H}_s$ matrix elements in the $|n\ell s m_\ell m_s\rangle$ basis. We need the formula for the Y_L^M matrix elements in terms of the $3j$ symbol coefficients,

$$\begin{aligned}\left\langle \ell' m_{\ell'} \left| Y_L^M \right| \ell m \right\rangle \\ = (-1)^{m_{\ell'}} \sqrt{\frac{1}{4\pi}(2\ell'+1)(2L+1)(2\ell+1)} \begin{pmatrix} \ell' & L & \ell \\ -m_{\ell'} & M & m_\ell \end{pmatrix} \begin{pmatrix} \ell' & L & \ell \\ 0 & 0 & 0 \end{pmatrix}.\end{aligned} \tag{2.73}$$

According to the $3j$-symbol properties, the second $3j$-symbol in the previous equation is zero when $\ell' + L + \ell$ is odd. Since $L = 1$ (we are dealing with $Y_L^M = Y_1^0$), the above matrix element is zero if $\ell' = \ell$. This implies that, in general, there is no first-order splitting due to an electric field. This was to be expected, because the wave functions $\psi_{n\ell m}$ have a *definite* parity (even when ℓ is even and odd otherwise), therefore the matrix element $\langle \psi_{n\ell m} | e\mathcal{E}z | \psi_{n\ell m} \rangle$ vanishes because the Hamiltonian $\mathcal{H}_s = e\mathcal{E}z$ has odd parity.[13]

However, for Hydrogen things are a bit different because, in first approximation, the energy eigenvalues depend only on the principal quantum number n and not on ℓ. For $n = 2$, for example, we have four degenerate states: $(\ell = 0, m_\ell = 0)$, $(\ell = 1, m_\ell = 0, \pm 1)$. So within this collection of states there is a first-order Stark effect because there are nonzero matrix elements of Y_1^0 between the state with $\ell = 0$ and the states with $\ell = 1$. To compute the first-order Stark effect in the $n = 2$ states of Hydrogen note that

$$\begin{pmatrix} \ell' & L & \ell \\ -m_{\ell'} & M & m_\ell \end{pmatrix} = 0 \tag{2.74}$$

unless $-m_{\ell'} + M + m_\ell = 0$. This implies that $m_{\ell'} = m_\ell$, because $M = 0$. Also, since one of the ℓ values is zero, we conclude that $m_{\ell'} = m_\ell = 0$. We have then only two nonzero matrix elements left. Using the notation $|n\ell m_\ell\rangle$,

$$\left\langle 200 \left| Y_1^0 \right| 210 \right\rangle = \left\langle 210 \left| Y_1^0 \right| 200 \right\rangle = \frac{1}{\sqrt{4\pi}}. \tag{2.75}$$

To proceed further we need the matrix element of r between the radial part of the wavefunctions $|200\rangle \equiv 2s$ and $|210\rangle \equiv 2p_0$, which is given by

$$\langle R(2s) \,|\, r \,|\, R(2p_0) \rangle = \langle R(2p_0) \,|\, r \,|\, R(2s) \rangle = -\frac{9}{\sqrt{3}} \frac{a_0}{Z}. \tag{2.76}$$

Now, since $\mathcal{H}_s = e\mathcal{E} r Y_1^0 \sqrt{4\pi/3}$, we combine everything together into

$$\langle 200 \,|\, \mathcal{H}_s \,|\, 210 \rangle = -e\mathcal{E} \frac{9}{\sqrt{3}} \frac{a_0}{Z} \sqrt{\frac{4\pi}{3}} \frac{1}{\sqrt{4\pi}} = -3e\mathcal{E}a_0/Z, \tag{2.77}$$

and build table (2.2) showing the matrix elements for the Stark Effect in hydrogen with $n = 2$, $s = 1/2$, $m_s = m_s'$. From the table, we see that the secular equation leads directly to the following four eigenvalues,

$$E = \pm 3e\mathcal{E}a_0/Z,\ 0,\ 0. \tag{2.78}$$

The two states with $m_\ell = 0$ are shifted up and down symmetrically, while the states with $m_\ell = \pm 1$ are not affected by the electric field. This means that the m degeneracy is only partially lifted, as shown in figure (2.4). You may check that the eigenvalue $-3e\mathcal{E}a_0/Z$ is associated to the new normalized eigenstate $\frac{1}{\sqrt{2}}(\psi_{200} + \psi_{210})$, while the eigenstate $\frac{1}{\sqrt{2}}(\psi_{200} - \psi_{210})$ is associated to $+3e\mathcal{E}a_0/Z$.

[13] The integration over all space of an odd-parity expression vanishes. Remember that the parity of Y_ℓ^m is $(-1)^\ell$.

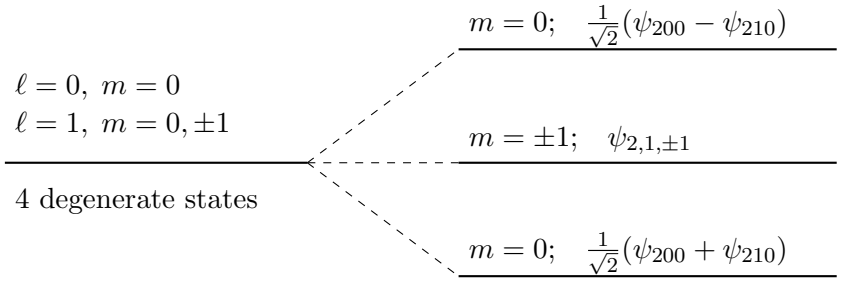


Figure 2.4: Splitting of the degenerate $n = 2$ Hydrogen level due to the linear Stark effect.

	$\|0\ 0\rangle$	$\|1\ 0\rangle$	$\|1\ 1\rangle$	$\|1\ -1\rangle$
$\langle 0\ 0\|$		$-3e\mathcal{E}a_0/Z$		
$\langle 1\ 0\|$	$-3e\mathcal{E}a_0/Z$			
$\langle 1\ 1\|$				
$\langle 1\ -1\|$				

Table 2.2: Matrix elements for the Stark effect in hydrogen with $n = 2$, $s = 1/2$, $m_s = m'_s$.

In hydrogen, assuming an electric field of $\mathcal{E} = 10^4$ V cm^{-1}, we have (remember that $a_0 = 5.29 \times 10^{-9}$ cm)

$$3e\mathcal{E}a_0 = 3 \times 10^4 \times 5.29 \times 10^{-9}\,\text{eV} = 1.59 \times 10^{-4}\,\text{eV} \;=\; 1.28\,\text{cm}^{-1}, \tag{2.79}$$

which is much larger than the fine structure splitting. The Stark effect in hydrogen is linear due to the degeneracy of states with different ℓ and the same n. At very high electric fields there appears a quadratic effect superimposed on the linear one, resulting in an asymmetric shifting of the energy levels.

2.2.2 Quadratic Stark effect.

We have seen that for the ground state of hydrogenic atoms there is no linear Stark effect. We must then go to second order in perturbation theory. The second-order correction to the (zero order) energy eigenvalue $E_n^{(0)}$ is (see equation (B.64)).

$$E_n^{(2)} = \sum_{m \neq n} \frac{\left| \left\langle \psi_n^{(0)} \right| \mathcal{H}_s \left| \psi_m^{(0)} \right\rangle \right|^2}{E_n^{(0)} - E_m^{(0)}}, \tag{2.80}$$

where $\mathcal{H}_s = e\mathcal{E}z$ is the small perturbation. The second-order correction to the hydrogenic ground state is then (we neglect now the fine structure),

$$E_{100}^{(2)} = e^2\mathcal{E}^2 \sum_{n \neq 1, \ell, m} \frac{\left| \left\langle \psi_{n\ell m}^{(0)} \right| z \left| \psi_{100}^{(0)} \right\rangle \right|^2}{E_1^{(0)} - E_n^{(0)}}. \tag{2.81}$$

The complete set of hydrogenic states is the union of a discrete set and a continuous (non-discrete) set of wavefunctions. The summation above implies a summation over the discrete set of eigenfunctions and an integration over the continuous set of eigenfunctions. Since the energy differences $E_1^{(0)} - E_n^{(0)}$ are always negative for $n \geq 2$, it is clear from (2.81) that the correction $E_{100}^{(2)}$ is always also negative. The effect of the quadratic Stark effect is to lower (make it more negative) the energy of the ground state.[14] We can easily obtain a lower limit for $E_{100}^{(2)}$

[14] This is a general feature of second-order corrections in time-independent perturbation theory, they are always negative.

by replacing the energy differences $E_1^{(0)} - E_n^{(0)}$ by $E_1^{(0)} - E_2^{(0)}$. Remember that, in Rydberg units we have

$$E_2^{(0)} - E_1^{(0)} \approx 1 - \frac{1}{4} = \frac{3}{4}, \qquad E_n^{(0)} - E_1^{(0)} \approx 1 - \frac{1}{n^2} > \frac{3}{4} \quad \text{for} \quad n > 2.$$

Therefore

$$E_{100}^{(2)} = -e^2\mathcal{E}^2 \sum_{n\neq 1,\ell,m} \frac{\left|\left\langle \psi_{n\ell m}^{(0)} \middle| z \middle| \psi_{100}^{(0)} \right\rangle\right|^2}{E_n^{(0)} - E_1^{(0)}} > \frac{-e^2\mathcal{E}^2}{E_2^{(0)} - E_1^{(0)}} \sum_{n\neq 1,\ell,m} \left|\left\langle \psi_{n\ell m}^{(0)} \middle| z \middle| \psi_{100}^{(0)} \right\rangle\right|^2. \quad (2.82)$$

To perform the summation we first note that $\left\langle \psi_{100}^{(0)} \middle| z \middle| \psi_{100}^{(0)} \right\rangle = 0$, therefore we may write

$$\begin{aligned} \sum_{n\neq 1,\ell,m} \left|\left\langle \psi_{n\ell m}^{(0)} \middle| z \middle| \psi_{100}^{(0)} \right\rangle\right|^2 &= \sum_{n,\ell,m} \left|\left\langle \psi_{n\ell m}^{(0)} \middle| z \middle| \psi_{100}^{(0)} \right\rangle\right|^2 \\ &= \sum_{n,\ell,m} \left\langle \psi_{100}^{(0)} \middle| z \middle| \psi_{n\ell m}^{(0)} \right\rangle \left\langle \psi_{n\ell m}^{(0)} \middle| z \middle| \psi_{100}^{(0)} \right\rangle = \left\langle \psi_{100}^{(0)} \middle| z^2 \middle| \psi_{100}^{(0)} \right\rangle = \left\langle z^2 \right\rangle_{100}, \end{aligned} \quad (2.83)$$

where we have used the completeness relation of the hydrogenic sates, $\sum_{n,\ell,m} \left|\psi_{n\ell m}^{(0)}\right\rangle \left\langle \psi_{n\ell m}^{(0)}\right| = I$ (summation and integration). Now we use the results

$$\left\langle r^2 \right\rangle_{n\ell m} = \frac{a_0^2}{Z^2}\frac{n^2}{2}\left[5n^2 + 1 - 3\ell(\ell+1)\right], \qquad \left\langle r^2 \right\rangle_{100} = 3\frac{a_0^2}{Z^2}, \quad (2.84)$$

and take into account the symmetry of the hydrogenic ground state to write

$$\left\langle z^2 \right\rangle_{100} = \left\langle x^2 \right\rangle_{100} = \left\langle y^2 \right\rangle_{100} = \frac{1}{3}\left\langle r^2 \right\rangle_{100}, \qquad \left\langle z^2 \right\rangle_{100} = \frac{a_0^2}{Z^2}. \quad (2.85)$$

Plugging everything in (2.82) we get

$$E_{100}^{(2)} > \frac{-e^2\mathcal{E}^2}{E_2^{(0)} - E_1^{(0)}}\frac{a_0^2}{Z^2} = -\frac{1}{E_2^{(0)} - E_1^{(0)}}\left(\frac{e\mathcal{E}a_0}{Z}\right)^2 \quad (2.86)$$

Expressing energies in Rydberg units we have

$$E_n^{(0)} = -\frac{Z}{n^2}, \qquad \frac{1}{E_2^{(0)} - E_1^{(0)}} = \frac{4}{3}\frac{1}{Z}, \qquad E_{100}^{(2)} > -1.33\,\frac{1}{Z}\left(\frac{e\mathcal{E}a_0}{Z}\right)^2. \quad (2.87)$$

Another approximation can be obtained in an easy way (see problem 14), which is

$$E_{100}^{(2)} \approx \frac{1}{E_1^{(0)}}\left(\frac{e\mathcal{E}a_0}{Z}\right)^2 = -\frac{1}{Z}\left(\frac{e\mathcal{E}a_0}{Z}\right)^2. \quad (2.88)$$

In any case, it is possible to perform a full evaluation of the sum in equation (2.81). Here we just quote the result,

$$E_{100}^{(2)} = \frac{9}{8}\frac{1}{E_1^{(0)}}\left(\frac{e\mathcal{E}a_0}{Z}\right)^2 = -\frac{9}{8}\frac{1}{Z}\left(\frac{e\mathcal{E}a_0}{Z}\right)^2 = -1.125\left(\frac{e\mathcal{E}a_0}{Z}\right)^2. \quad (2.89)$$

It is interesting to note that about one third of this value arises from the contribution of the continuum in the summation.

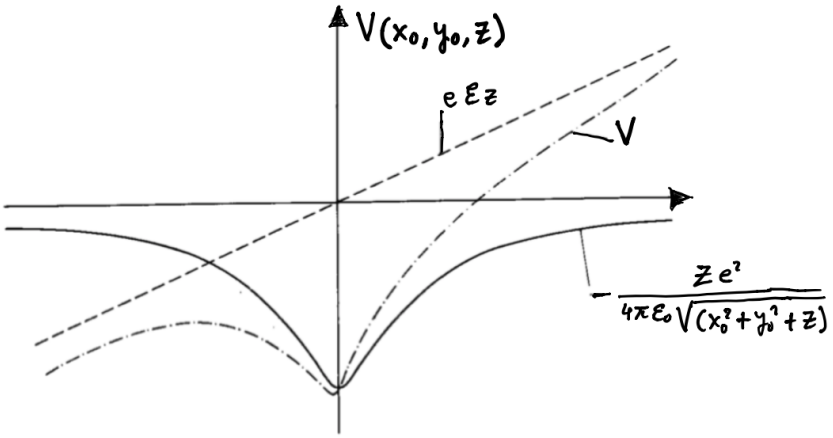


Figure 2.5: Potential experienced by an electron interacting with a nucleus of charge Ze in a uniform electric field of strength $\mathcal{E}$, as a function of z, for $x = x_0$ and $y = y_0$ fixed (adapted from Bransden, [2]).

2.2.3 Ionization by a static electric field.

An external electric field applied to the atom may have the effect of removing the electron from the atom. Consider a hydrogenic atom placed in an external electric field $\boldsymbol{\mathcal{E}}$ in the z-axis direction. The potential energy of the electron is given by

$$V = -\frac{Ze^2}{r} + e\mathcal{E}z. \tag{2.90}$$

A plot of V as a function of z, for x and y fixed (figure 2.5), shows that the nucleus represents just a local minima, and that V can become even more negative if z is negative enough (large distances from the nucleus). One could then see this potential as having two minima separated by a barrier, one at the nucleus and the other at the anode. The electron has a non–zero probability to escape from the atom by means of the so-called *tunnel effect* and be accelerated towards the anode.

This ionization can be observed when the external electric field is very strong and/or for levels with high principal quantum number n. Consequently, in the presence of an external electric field the lifetime of the discrete levels is decreased due to this tunnel effect, so the width of the spectral lines is increased. This is known as *Stark broadening*. In particular, the ground state itself is no longer stationary. It can be shown that the lifetime of the hydrogen ground state in an external electric field is (in atomic units)

$$\tau = \frac{1}{4}\mathcal{E}\exp\left(\frac{2}{3\mathcal{E}}\right). \tag{2.91}$$

For small values of the electric field strength, $\mathcal{E}$, the lifetime of the ground state is very long and can be considered essentially stable. Even if we take a high value such as, for example, $\mathcal{E} = 0.01\,\text{a.u.} = 5.14 \times 10^9\,\text{V}\cdot\text{m}^{-1}$, we get $\tau = 2 \times 10^{26}\,\text{a.u.} = 4.8 \times 10^9\,\text{s}$.

2.3 Hyperfine interactions.

An ideal point-like nucleus without intrinsic angular momentum (spin) gives rise only to a Coulomb potential of the form $-Ze/r$. But real nuclei are notoriously complicated systems, certainly not point-like, with charge distributions not necessarily spherical and different values of spin. All interactions between electrons and nuclei other than the Coulomb interaction are

referred to as *hyperfine* interactions. Among them, the most important are those due to the nuclear spin and the nuclear quadrupole moment. The nuclear spin gives rise to a nuclear magnetic moment that interacts with the electronic magnetic dipole moments associated with the orbital and/or spin angular momenta. Furthermore, a non-spherical nuclear electric charge distribution creates an electric quadrupole moment that interacts with the electric field gradient created by the electronic charge distribution at the position of the nucleus.

2.3.1 Magnetic hyperfine interaction.

The Hamiltonian for the magnetic hyperfine interaction can be derived from the Dirac equation (1.79), which we reproduce here expanded,

$$E'\psi_u = \left[q\phi + \frac{1}{2m}\left[\boldsymbol{\sigma}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)K\boldsymbol{\sigma}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)\right]\right]\psi_u \tag{2.92}$$

where

$$K = \frac{2mc^2}{E' + 2mc^2 - q\phi}, \qquad E' = E - mc^2. \tag{2.93}$$

It is convenient to write in a separate $\mathcal{H}_1$ term the dependence of the Hamiltonian on the vector potential,

$$\mathcal{H} = q\phi + \frac{1}{2m}\left[\boldsymbol{\sigma}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)K\boldsymbol{\sigma}\left(\mathbf{p} - \frac{q}{c}\mathbf{A}\right)\right] = \mathcal{H}_0 + \mathcal{H}_1 \tag{2.94}$$

$$\mathcal{H}_0 = q\phi + \frac{1}{2m}(\boldsymbol{\sigma}\mathbf{p})K(\boldsymbol{\sigma}\mathbf{p}), \tag{2.95}$$

$$\mathcal{H}_1 = -\frac{q}{2mc^2}\Big[(\boldsymbol{\sigma}\mathbf{p})K(\boldsymbol{\sigma}\mathbf{A}) + (\boldsymbol{\sigma}\mathbf{A})K(\boldsymbol{\sigma}\mathbf{p})\Big] + \frac{q^2}{2mc^2}(\boldsymbol{\sigma}\mathbf{A})K(\boldsymbol{\sigma}\mathbf{A}). \tag{2.96}$$

The last term of the previous expression is quadratic in $\mathbf{A}$ and can be neglected in a first approximation. If the potential ϕ depends only on r, $\phi = \phi(r)$, then K is also a function only of r, $K(r)$, and we may write

$$\boldsymbol{\nabla}K = \frac{dK}{dr}\frac{\mathbf{r}}{r}. \tag{2.97}$$

Furthermore remember that the operator $\mathbf{p} = -i\hbar\boldsymbol{\nabla}$ acts on everything that stands to its right, so for an arbitrary function $f(\mathbf{r})$

$$\mathbf{p}Kf(\mathbf{r}) = -i\hbar\boldsymbol{\nabla}(Kf(\mathbf{r})) = -i\hbar K(\boldsymbol{\nabla}f) - i\hbar f\boldsymbol{\nabla}K = K\mathbf{p}f - i\hbar\frac{dK}{dr}\frac{\mathbf{r}}{r}f. \tag{2.98}$$

Therefore

$$\mathbf{p}K = K\mathbf{p} - i\hbar\frac{dK}{dr}\frac{\mathbf{r}}{r}. \tag{2.99}$$

Using this relation we can write

$$(\boldsymbol{\sigma}\mathbf{p})K(\boldsymbol{\sigma}\mathbf{A}) = K(\boldsymbol{\sigma}\mathbf{p})(\boldsymbol{\sigma}\mathbf{A}) - i\hbar\frac{dK}{dr}\frac{1}{r}(\boldsymbol{\sigma}\mathbf{r})(\boldsymbol{\sigma}\mathbf{A}). \tag{2.100}$$

Inserting this result in the expression for $\mathcal{H}_1$ and neglecting the quadratic part in $\mathbf{A}$ we have

$$\mathcal{H}_1 = -\frac{q}{2mc}\Big[K(\boldsymbol{\sigma}\mathbf{p})(\boldsymbol{\sigma}\mathbf{A}) - i\hbar\frac{dK}{dr}\frac{1}{r}(\boldsymbol{\sigma}\mathbf{r})(\boldsymbol{\sigma}\mathbf{A}) + K(\boldsymbol{\sigma}\mathbf{A})(\boldsymbol{\sigma}\mathbf{p})\Big]. \tag{2.101}$$

This expression can be transformed using the rule

$$(\boldsymbol{\sigma}\mathbf{A})(\boldsymbol{\sigma}\mathbf{B}) = \mathbf{A}\mathbf{B} + i\boldsymbol{\sigma}(\mathbf{A}\times\mathbf{B}), \tag{2.102}$$

into

$$\mathcal{H}_1 = -\frac{q}{2mc}K(\mathbf{p}\mathbf{A} + \mathbf{A}\mathbf{p}) + iK\boldsymbol{\sigma}(\mathbf{p}\times\mathbf{A} + \mathbf{A}\times\mathbf{p}) - i\hbar\frac{dK}{dr}\frac{1}{r}\big[\mathbf{r}\mathbf{A} + i\boldsymbol{\sigma}(\mathbf{r}\times\mathbf{A})\big]. \tag{2.103}$$

To proceed further we need to specify the vector potential. If we assume the nucleus to be a point dipole with a magnetic dipole moment $\boldsymbol{\mu}$, the vector potential at $\mathbf{r}$ is[15]

$$\mathbf{A} = \frac{\boldsymbol{\mu} \times \mathbf{r}}{r^3} \quad \text{(Gaussian units)}. \tag{2.104}$$

Now we use the vector identities

$$\boldsymbol{\nabla}(\mathbf{a} \times \mathbf{b}) = \mathbf{b}(\boldsymbol{\nabla} \times \mathbf{a}) - \mathbf{a}(\boldsymbol{\nabla} \times \mathbf{b}), \qquad \boldsymbol{\nabla} \times f\mathbf{a} = (\boldsymbol{\nabla} f) \times \mathbf{a} + f(\boldsymbol{\nabla} \times \mathbf{a}), \tag{2.105}$$

$$\boldsymbol{\nabla}\left(\frac{\mathbf{a}}{f}\right) = \frac{f(\boldsymbol{\nabla}\mathbf{a}) - \mathbf{a}(\boldsymbol{\nabla} f)}{f^2}, \tag{2.106}$$

together with the fact that $\boldsymbol{\nabla} \times \boldsymbol{\mu} = 0$, because $\boldsymbol{\mu}$ is constant, to find that $\boldsymbol{\nabla}\mathbf{A} = 0$. Indeed, since $\boldsymbol{\nabla} \times \mathbf{r} = 0$ and $\boldsymbol{\nabla}(1/r^3) = [\partial(1/r^3)/\partial r]\,\hat{\mathbf{r}} = -3\hat{\mathbf{r}}/r^4 = -3\mathbf{r}/r^5$, we have

$$\begin{aligned}\boldsymbol{\nabla}\mathbf{A} = \boldsymbol{\nabla}\left(\boldsymbol{\mu} \times \frac{\mathbf{r}}{r^3}\right) &= \frac{\mathbf{r}}{r^3}\cancel{(\boldsymbol{\nabla} \times \boldsymbol{\mu})} - \boldsymbol{\mu}\left(\boldsymbol{\nabla} \times \frac{\mathbf{r}}{r^3}\right) \\ &= -\boldsymbol{\mu}\left[\left(\boldsymbol{\nabla}\frac{1}{r^3}\right) \times \mathbf{r} + \frac{1}{r^3}\cancel{\boldsymbol{\nabla} \times \mathbf{r}}\right] = \boldsymbol{\mu}\left(3\frac{\mathbf{r}}{r^5} \times \mathbf{r}\right) = 0.\end{aligned} \tag{2.107}$$

Furthermore, note that $\mathbf{rA} = 0$, because $\mathbf{r}(\boldsymbol{\mu} \times \mathbf{r}) = 0$, and remember that $\mathbf{p} \times \mathbf{A} + \mathbf{A} \times \mathbf{p} = -i\hbar\,(\boldsymbol{\nabla} \times \mathbf{A})$ and $\mathbf{pA} - \mathbf{Ap} = -i\hbar\boldsymbol{\nabla}\mathbf{A}$, which means in this case $\mathbf{pA} = \mathbf{Ap}$ because $\boldsymbol{\nabla}\mathbf{A} = 0$. Combining now everything together and putting $q = -e$, we get

$$\mathcal{H}_1 = \frac{e}{2mc}\left[2K\mathbf{Ap} + K\hbar\boldsymbol{\sigma}(\boldsymbol{\nabla} \times \mathbf{A}) + \hbar\frac{dK}{dr}\frac{1}{r}\boldsymbol{\sigma}(\mathbf{r} \times \mathbf{A})\right]. \tag{2.108}$$

According to the special form of the vector potential, we may write

$$\mathbf{Ap} = \frac{1}{r^3}(\boldsymbol{\mu} \times \mathbf{r})\mathbf{p} = \frac{1}{r^3}(\mathbf{r} \times \mathbf{p})\boldsymbol{\mu} = \frac{1}{r^3}\mathbf{L}\boldsymbol{\mu} = \frac{1}{r^3}\boldsymbol{\mu}\mathbf{L}. \tag{2.109}$$

Also, since $\boldsymbol{\mu}$ is constant[16],

$$\boldsymbol{\nabla} \times \left(\frac{\boldsymbol{\mu}}{r}\right) = \boldsymbol{\nabla}\left(\frac{1}{r}\right) \times \boldsymbol{\mu} + \frac{1}{r}\boldsymbol{\nabla} \times \boldsymbol{\mu} = \boldsymbol{\nabla}\left(\frac{1}{r}\right) \times \boldsymbol{\mu} = -\boldsymbol{\mu} \times \boldsymbol{\nabla}\left(\frac{1}{r}\right) = \frac{\boldsymbol{\mu} \times \mathbf{r}}{r^3} = \mathbf{A}, \tag{2.110}$$

$$\nabla^2\left(\frac{\boldsymbol{\mu}}{r}\right) = \boldsymbol{\mu}\nabla^2\left(\frac{1}{r}\right) = -4\pi\boldsymbol{\mu}\delta(\mathbf{r}), \tag{2.111}$$

$$\boldsymbol{\nabla}\left(\frac{\boldsymbol{\mu}}{r}\right) = \boldsymbol{\mu}\boldsymbol{\nabla}\left(\frac{1}{r}\right) + \frac{1}{r}\boldsymbol{\nabla}\boldsymbol{\mu} = \boldsymbol{\mu}\boldsymbol{\nabla}\left(\frac{1}{r}\right) = -\frac{\boldsymbol{\mu}\mathbf{r}}{r^3}, \tag{2.112}$$

$$\boldsymbol{\nabla}\left[\boldsymbol{\nabla}\left(\frac{\boldsymbol{\mu}}{r}\right)\right] = -\boldsymbol{\nabla}\left(\frac{\boldsymbol{\mu}\mathbf{r}}{r^3}\right) = -\left[(\boldsymbol{\mu}\mathbf{r})\boldsymbol{\nabla}\left(\frac{1}{r^3}\right) + \frac{1}{r^3}\boldsymbol{\nabla}(\boldsymbol{\mu}\mathbf{r})\right] = (\boldsymbol{\mu}\mathbf{r})\frac{3\mathbf{r}}{r^5} - \frac{1}{r^3}\boldsymbol{\mu}, \tag{2.113}$$

where in the last step we have used $\boldsymbol{\nabla}(\boldsymbol{\mu}\mathbf{r}) = \boldsymbol{\nabla}(\mu_x x + \mu_y y + \mu_z z) = \boldsymbol{\mu}$. We need also the vector identity $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{a}) = \boldsymbol{\nabla}(\boldsymbol{\nabla}\mathbf{a}) - \nabla^2\mathbf{a}$ to write

$$\boldsymbol{\nabla} \times \mathbf{A} = \boldsymbol{\nabla} \times \left[\boldsymbol{\nabla} \times \left(\frac{\boldsymbol{\mu}}{r}\right)\right] = \boldsymbol{\nabla}\left[\boldsymbol{\nabla}\left(\frac{\boldsymbol{\mu}}{r}\right)\right] - \nabla^2\left(\frac{\boldsymbol{\mu}}{r}\right) = (\boldsymbol{\mu}\mathbf{r})\frac{3\mathbf{r}}{r^5} - \frac{1}{r^3}\boldsymbol{\mu} + 4\pi\boldsymbol{\mu}\delta(\mathbf{r}). \tag{2.114}$$

The last bit of vector handling we need for the moment is $\mathbf{a} \times \mathbf{b} \times \mathbf{c} = \mathbf{b}(\mathbf{ac}) - \mathbf{c}(\mathbf{ab})$,

$$\mathbf{r} \times \mathbf{A} = \frac{1}{r^3}(\mathbf{r} \times \boldsymbol{\mu} \times \mathbf{r}) = \frac{1}{r^3}\left[\boldsymbol{\mu}(\mathbf{rr}) - \mathbf{r}(\mathbf{r}\boldsymbol{\mu})\right] = \frac{\boldsymbol{\mu}}{r} - \frac{\mathbf{r}(\mathbf{r}\boldsymbol{\mu})}{r^3}. \tag{2.115}$$

[15] See Eq. (B.313). In the SI system we would have $\mathbf{A} = \frac{\mu_0}{4\pi}\frac{\boldsymbol{\mu} \times \mathbf{r}}{r^3}$ (SI units).

[16] In the vector operations discussed here be careful to note that the gradient of a divergence, $\boldsymbol{\nabla}(\boldsymbol{\nabla}\mathbf{a})$, is not the same as the Laplacian of a vector, $\nabla^2\mathbf{a} = (\boldsymbol{\nabla}\boldsymbol{\nabla})\mathbf{a} \neq \boldsymbol{\nabla}(\boldsymbol{\nabla}\mathbf{a})$.

Now we insert everything back in the expression for $\mathcal{H}_1$ to get

$$\mathcal{H}_1 = \frac{2\mu_{\rm B}}{\hbar} K \left[\frac{\boldsymbol{\mu}(\mathbf{L}-\mathbf{S})}{r^3} + \frac{3(\boldsymbol{\mu}\mathbf{r})(\mathbf{S}\mathbf{r})}{r^5} + 4\pi\delta(\mathbf{r})\boldsymbol{\mu}\mathbf{S} \right] + \frac{2\mu_{\rm B}}{\hbar}\frac{dK}{dr}\left[\frac{\boldsymbol{\mu}\mathbf{S}}{r^2} - \frac{(\boldsymbol{\mu}\mathbf{r})(\mathbf{S}\mathbf{r})}{r^4} \right] \tag{2.116}$$

where

$$\mu_{\rm B} = \frac{e\hbar}{2mc}, \qquad \mathbf{S} = \frac{\hbar}{2}\boldsymbol{\sigma}. \tag{2.117}$$

We need to play close attention to K because when replacing ϕ by the Coulombic potential we have a divergence at $r = 0$. Indeed, setting $\phi = Ze/r$ and $q = -e$ we have

$$K = \frac{2mc^2}{E' + 2mc^2 + \dfrac{Ze^2}{r}}. \tag{2.118}$$

When computing the fine structure of Hydrogen we made a non–relativistic approximation to the Dirac equation by assuming that (see (1.85))

$$E' + \frac{Ze^2}{r} \ll 2mc^2. \tag{2.119}$$

However, this approximation is not valid for an atom in an s state, since the wavefunction for an $\ell = 0$ state has a nonzero value at $r = 0$. In such a case we will follow an argument adapted from the original derivation by Fermi and assume that the kinetic energy of the electron is negligible compared to the electron rest mass energy, $E' = E - mc^2 \ll 2mc^2$, but that the terms Ze^2/r and $2mc^2$ may have a similar size. Therefore

$$K(r) \approx \frac{1}{1 + \dfrac{Ze^2}{2mc^2 r}} \equiv \frac{1}{1 + \dfrac{r_0}{r}}, \qquad r_0 = \frac{Ze^2}{2mc^2}. \tag{2.120}$$

For $Z = 1$, for example, the ratio r_0/a_0 is

$$\frac{r_0}{a_0} = \frac{e^2}{2mc^2}\frac{me^2}{\hbar^2} = \frac{1}{2}\left(\frac{e^2}{\hbar c}\right)^2 = \frac{1}{2}\alpha^2 \approx \frac{1}{2}\left(\frac{1}{137}\right)^2. \tag{2.121}$$

Since $a_0 = 0.52 \times 10^{-10}\,\text{m}$, we get $r_0 = 1.4 \times 10^{-15}\,\text{m}$. This value is in the range of nuclear sizes, much smaller than the Bohr radius. In the plot (2.6) we see that $K(r)$ rises from zero at $r = 0$ to almost unity in a distance of several units of r_0, so we may approximate it by a step function,

$$K(r) \approx \begin{cases} 0, & \text{when } r = 0 \\ 1 & \text{when } r \neq 0. \end{cases} \tag{2.122}$$

We can show that, under this approximation, the δ-function term in $\mathcal{H}_1$ can safely be ignored. Indeed, since $K(0) = 0$, for any state (with any value of ℓ) we have

$$\langle \psi \,|\, K\delta(\mathbf{r}) \,|\, \psi \rangle = K(0)|\psi(0)|^2 = 0. \tag{2.123}$$

We write then $\mathcal{H}_1$ without the δ-function,

$$\mathcal{H}_1 = \mathcal{H}_a + \mathcal{H}_b, \tag{2.124}$$

$$\mathcal{H}_a = \frac{2\mu_{\rm B}}{\hbar}\left[\frac{\boldsymbol{\mu}(\mathbf{L}-\mathbf{S})}{r^3} + \frac{3(\boldsymbol{\mu}\mathbf{r})(\mathbf{S}\mathbf{r})}{r^5} \right], \qquad \mathcal{H}_b = \frac{2\mu_{\rm B}}{\hbar}\frac{dK}{dr}\left[\frac{\boldsymbol{\mu}\mathbf{S}}{r^2} - \frac{(\boldsymbol{\mu}\mathbf{r})(\mathbf{S}\mathbf{r})}{r^4} \right]. \tag{2.125}$$

We have set $K = 1$ in $\mathcal{H}_a$. This is a valid approximation because, as it will be shown later, there is no contribution from $\mathcal{H}_a$ in s states, which are the only ones with non-vanishing amplitudes at $r = 0$ (where $K = 0$).

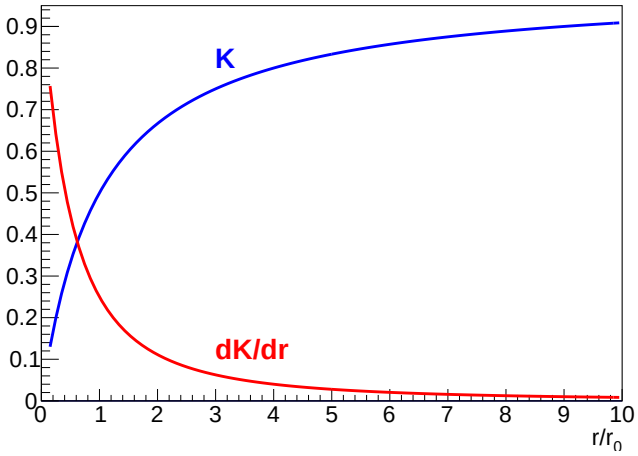


Figure 2.6: Variation of K and K' with r, in units of r_0.

We turn our attention now to the term $K' = dK/dr$. We expect that, since K was approximated by a step function, K might be approximated by a δ-function. Figure (2.6) shows that the function

$$K' = \frac{r_0/r^2}{\left(1 + \frac{r_0}{r}\right)^2} \tag{2.126}$$

goes close to zero within just a few nuclear radii, so we could take $K' = 0$ for $r \neq 0$. On the other hand, assuming that $K(r)$ is the step function given by (2.122), we have

$$\int_0^\infty K' dr = K(\infty) - K(0) = 1 - 0 = 1. \tag{2.127}$$

The behavior of K and K' at negative r has no physical meaning, so we may as well set $K = K' = 0$ for all $r < 0$ and write

$$\int_{-\infty}^{+\infty} K' dr = 1. \tag{2.128}$$

This last equation, together with the condition $K' = 0$ for $r \neq 0$, are the properties of a δ-function. Let us write $K' = 4\pi r^2 \delta(\mathbf{r})$, because then

$$\int_{-\infty}^{+\infty} K'\, dr = \int_{-\infty}^{+\infty} 4\pi\delta(\mathbf{r})\, r^2 dr = \int_{-\infty}^{+\infty} \delta(\mathbf{r})\, d\mathbf{r} = 1. \tag{2.129}$$

Concerning the $\mathcal{H}_b$ term, let's write it as

$$\mathcal{H}_b = -\boldsymbol{\mu}\mathbf{B}, \qquad \text{with} \qquad \mathbf{B} = -\frac{2\mu_B}{\hbar}\frac{dK}{dr}\left[\frac{\mathbf{S}}{r^2} - \frac{\mathbf{r}(\mathbf{Sr})}{r^4}\right], \tag{2.130}$$

and apply the Landé formula (let $\mathbf{A} \to \mathbf{B}$, $\mathbf{J} \to \mathbf{S}$ and $j \to s$ in equation (2.18)),

$$\left\langle sm \,\middle|\, \mathbf{B} \,\middle|\, sm' \right\rangle = \frac{\langle sm \,|\, \mathbf{BS} \,|\, sm\rangle}{s(s+1)\hbar^2} \left\langle sm \,\middle|\, \mathbf{S} \,\middle|\, sm' \right\rangle, \tag{2.131}$$

with

$$\langle sm \,|\, \mathbf{BS} \,|\, sm\rangle = -\frac{2\mu_B}{\hbar}\left\langle sm \,\middle|\, \frac{dK}{dr}\left[\frac{S^2}{r^2} - \frac{(\mathbf{Sr})^2}{r^4}\right] \,\middle|\, sm \right\rangle, \tag{2.132}$$

and

$$(\mathbf{Sr})(\mathbf{Sr}) = \frac{\hbar^2}{4}(\boldsymbol{\sigma}\mathbf{r})(\boldsymbol{\sigma}\mathbf{r}) = \frac{\hbar^2}{4}r^2, \tag{2.133}$$

to get (for $s = 1/2$)

$$\langle sm \,|\, \mathbf{BS} \,|\, sm\rangle \equiv \langle \mathbf{BS}\rangle = -\frac{2\mu_{\mathrm{B}}}{\hbar}\left\langle sm \left| \frac{dK}{dr}\left[\frac{S^2}{r^2} - \frac{\hbar^2}{4r^2}\right] \right| sm \right\rangle$$
$$= -\frac{2\mu_{\mathrm{B}}}{\hbar}\left\langle \frac{dK}{dr}\frac{1}{r^2}\right\rangle \left[s(s+1)\hbar^2 - \frac{\hbar^2}{4}\right] = -\mu_{\mathrm{B}}\hbar\left\langle \frac{dK}{dr}\frac{1}{r^2}\right\rangle, \quad (2.134)$$

where $\langle (dK/dr)(1/r^2)\rangle$ represents an integral taken over the radial part of the wave function. But since $K' = 4\pi r^2\delta(\mathbf{r})$, we have

$$\left\langle \frac{dK}{dr}\frac{1}{r^2}\right\rangle = 4\pi\left\langle \delta(\mathbf{r})\right\rangle, \qquad \text{thus} \qquad \langle \mathbf{BS}\rangle = -4\pi\mu_{\mathrm{B}}\hbar\left\langle \delta(\mathbf{r})\right\rangle. \quad (2.135)$$

Inserting this back into the Landé formula (2.131) gives

$$\left\langle sm \,\middle|\, \mathbf{B} \,\middle|\, sm'\right\rangle = -\frac{16\pi}{3}\frac{\mu_{\mathrm{B}}}{\hbar}\left\langle \delta(\mathbf{r})\right\rangle \left\langle sm \,\middle|\, \mathbf{S} \,\middle|\, sm'\right\rangle. \quad (2.136)$$

We may express this as an operator relationship

$$\mathbf{B} = -\frac{16\pi}{3}\frac{\mu_{\mathrm{B}}}{\hbar}\delta(\mathbf{r})\,\mathbf{S}, \quad (2.137)$$

so that $\mathcal{H}_b$ can be written as

$$\mathcal{H}_b = -\boldsymbol{\mu}\mathbf{B} = \frac{16\pi}{3}\frac{\mu_{\mathrm{B}}}{\hbar}\delta(\mathbf{r})\,\boldsymbol{\mu}\mathbf{S} = a\,\boldsymbol{\mu}\mathbf{S}, \qquad \text{with} \qquad a = \frac{16\pi}{3}\frac{\mu_{\mathrm{B}}}{\hbar}\delta(\mathbf{r}). \quad (2.138)$$

The term $\mathbf{B} = -a\mathbf{S}$ is sometimes regarded as an effective magnetic field. Inserting this expression for $\mathcal{H}_b$ into equation (2.125), the magnetic hyperfine Hamiltonian can be written as

$$\mathcal{H}_1 \equiv \mathcal{H}_h = \frac{2\mu_{\mathrm{B}}}{\hbar}\left[\frac{\boldsymbol{\mu}(\mathbf{L}-\mathbf{S})}{r^3} + \frac{3(\boldsymbol{\mu}\mathbf{r})(\mathbf{S}\mathbf{r})}{r^5} + \frac{8\pi}{3}\delta(\mathbf{r})\,\boldsymbol{\mu}\mathbf{S}\right] \quad (2.139)$$

We want to write now the nuclear magnetic moment operator $\boldsymbol{\mu}$ in terms of the **nuclear spin** operator $\mathbf{I}$. It is possible to write an expression similar to the one used for the magnetic moment of the electron,

$$\boldsymbol{\mu} = \frac{1}{\hbar}g_{\mathrm{N}}\mu_{\mathrm{N}}\mathbf{I}, \quad (2.140)$$

where

$$\mu_{\mathrm{N}} = \frac{e\hbar}{2Mc} = \frac{m}{M}\mu_{\mathrm{B}} = 5.0505\times10^{-24}\,\mathrm{erg/G} \;=\; 5.0505\times10^{-27}\,\mathrm{J/T} \quad (2.141)$$

is the nuclear Bohr magneton. Here M is the mass of the proton and g_{N} is a factor whose magnitude and sign depend on the nucleus under consideration. It is also common to write the nuclear magnetic moment as

$$\boldsymbol{\mu} = \frac{g_{\mathrm{N}}\mu_{\mathrm{N}}}{\hbar}\mathbf{I} = \gamma\mathbf{I}, \quad (2.142)$$

where the relation $\gamma = g_{\mathrm{N}}\mu_{\mathrm{N}}/\hbar$ is known as the nuclear gyromagnetic ratio. We define the magnetic moment of the nucleus, μ, as the expected value of the $\boldsymbol{\mu}$ operator in a nuclear state which has the maximum value of its spin third component,

$$\mu = \langle I(m_I = I)|\mu_z|I(m_I = I)\rangle = \frac{g_{\mathrm{N}}\mu_{\mathrm{N}}}{\hbar}\langle I(m_I = I)|I_z|I(m_I = I)\rangle = g_{\mathrm{N}}\mu_{\mathrm{N}}I = \gamma\hbar I. \quad (2.143)$$

This allows us to write the magnetic moment operator as $\boldsymbol{\mu} = \dfrac{\mu}{I\hbar}\mathbf{I}$.

Using the nuclear spin operator, $\mathbf{I}$, the magnetic hyperfine Hamiltonian can be written as

$$\mathcal{H}_h = \frac{2\mu_{\mathrm{B}}g_{\mathrm{N}}\mu_{\mathrm{N}}}{\hbar^2}\left[\frac{\mathbf{I}(\mathbf{L}-\mathbf{S})}{r^3} + \frac{3(\mathbf{I}\mathbf{r})(\mathbf{S}\mathbf{r})}{r^5} + \frac{8\pi}{3}\delta(\mathbf{r})\,\mathbf{I}\mathbf{S}\right] \quad (2.144)$$

This is the most common expression used for the magnetic hyperfine interaction. The first two terms in the brackets taken together are known as the dipole-dipole interaction, in analogy with the corresponding classical expression. The last term is the Fermi contact interaction and has no classic analog. In the absence of orbital angular momentum, the magnetic hyperfine Hamiltonian can be written as

$$\mathcal{H}_h = \mathbf{IAS}, \tag{2.145}$$

where

$$\mathbf{A} = \frac{2\mu_\mathrm{B} g_\mathrm{N} \mu_\mathrm{N}}{\hbar^2} \left[-\frac{1}{r^3} + \frac{3(\mathbf{rr})}{r^5} + \frac{8\pi}{3}\delta(\mathbf{r}) \right], \tag{2.146}$$

is the magnetic hyperfine coupling tensor.

2.3.2 Magnetic hyperfine interaction in one-electron systems.

To simplify the dipole-dipole part of the magnetic hyperfine Hamiltonian we start by writing the Hamiltonian (2.144) as

$$\mathcal{H}_h = -\frac{g_\mathrm{N}\mu_\mathrm{N}}{\hbar} \left[-\mathbf{I}\frac{2\mu_\mathrm{B}}{\hbar} \left(\frac{\mathbf{L}-\mathbf{S}}{r^3} + \frac{3\mathbf{r}(\mathbf{Sr})}{r^5} \right) - \frac{2\mu_\mathrm{B}}{\hbar}\frac{8\pi}{3}\delta(\mathbf{r})\,\mathbf{IS} \right]. \tag{2.147}$$

Denoting

$$\mathbf{B}' = -\frac{2\mu_\mathrm{B}}{\hbar} \left[\frac{\mathbf{L}-\mathbf{S}}{r^3} + \frac{3\mathbf{r}(\mathbf{Sr})}{r^5} \right], \tag{2.148}$$

we have

$$\mathcal{H}_h = -\frac{g_\mathrm{N}\mu_\mathrm{N}}{\hbar} \left[\mathbf{IB}' - \frac{2\mu_\mathrm{B}}{\hbar}\frac{8\pi}{3}\delta(\mathbf{r})\,\mathbf{IS} \right]. \tag{2.149}$$

We will show now that the operator $\mathbf{B}'$ is in fact proportional to the total angular momentum of the electron, $\mathbf{J}$. We need to use the Landé formula,

$$\left\langle jm \,\middle|\, \mathbf{B}' \,\middle|\, jm' \right\rangle = \frac{\left\langle jm \,\middle|\, \mathbf{B}'\mathbf{J} \,\middle|\, jm \right\rangle}{j(j+1)\hbar^2} \left\langle jm \,\middle|\, \mathbf{J} \,\middle|\, jm' \right\rangle, \tag{2.150}$$

with

$$\left\langle jm \,\middle|\, \mathbf{B}'\mathbf{J} \,\middle|\, jm \right\rangle = -\frac{2\mu_\mathrm{B}}{\hbar} \left\langle jm \,\middle|\, \frac{(\mathbf{L}-\mathbf{S})\mathbf{J}}{r^3} + \frac{3(\mathbf{Sr})(\mathbf{rJ})}{r^5} \,\middle|\, jm \right\rangle. \tag{2.151}$$

Now take into account that

$$(\mathbf{L}-\mathbf{S})\mathbf{J} = (\mathbf{L}-\mathbf{S})(\mathbf{L}+\mathbf{S}) = L^2 - S^2, \tag{2.152}$$

and that $|jm\rangle \equiv |\ell s j m\rangle$ is an eigenfunction of L^2, S^2, J^2 and J_z, to write

$$\langle jm|(\mathbf{L}-\mathbf{S})\mathbf{J}|jm\rangle = \left\langle jm \,\middle|\, L^2 - S^2 \,\middle|\, jm \right\rangle = \ell(\ell+1)\hbar^2 - s(s+1)\hbar^2. \tag{2.153}$$

Also,

$$\mathbf{rJ} = \mathbf{r}(\mathbf{L}+\mathbf{S}) = \mathbf{r}(\mathbf{r}\times\mathbf{p}) + \mathbf{rS} = \mathbf{rS} = \mathbf{Sr}, \tag{2.154}$$

therefore

$$(\mathbf{Sr})(\mathbf{rJ}) = (\mathbf{rS})(\mathbf{rS}) = \frac{\hbar^2}{4}(\boldsymbol{\sigma}\mathbf{r})(\boldsymbol{\sigma}\mathbf{r}) = \frac{\hbar^2}{4}r^2. \tag{2.155}$$

The integration over the radial part of the wave function gives then

$$\left\langle jm \,\middle|\, \mathbf{B}'\mathbf{J} \,\middle|\, jm \right\rangle = -2\mu_\mathrm{B}\hbar \left[\ell(\ell+1) - s(s+1) + \frac{3}{4} \right] \left\langle \frac{1}{r^3} \right\rangle, \tag{2.156}$$

and for $s = 1/2$,

$$\left\langle jm \,\middle|\, \mathbf{B}'\mathbf{J} \,\middle|\, jm \right\rangle = -2\mu_\mathrm{B}\hbar\ell(\ell+1) \left\langle \frac{1}{r^3} \right\rangle, \qquad (s = 1/2). \tag{2.157}$$

Inserting this back into equation (2.150) we have

$$\langle jm \,|\, \mathbf{B}' \,|\, jm' \rangle = -\frac{2\mu_{\mathrm{B}}}{\hbar} \frac{\ell(\ell+1)}{j(j+1)} \left\langle \frac{1}{r^3} \right\rangle \langle jm \,|\, \mathbf{J} \,|\, jm' \rangle \,, \tag{2.158}$$

which means that, as an operational relation, we may write

$$\mathbf{B}' = -\frac{2\mu_{\mathrm{B}}}{\hbar} \frac{\ell(\ell+1)}{j(j+1)} \left\langle \frac{1}{r^3} \right\rangle \mathbf{J}, \qquad (s = 1/2). \tag{2.159}$$

Now we replace (2.148) with (2.159) and put it back into (2.149) to get the following Hamiltonian for the case $s = 1/2$.

$$\mathcal{H}_h = \frac{2\mu_{\mathrm{B}} g_{\mathrm{N}} \mu_{\mathrm{N}}}{\hbar^2} \left[\frac{\ell(\ell+1)}{j(j+1)} \left\langle \frac{1}{r^3} \right\rangle \mathbf{IJ} + \frac{8\pi}{3} |\psi(0)|^2 \mathbf{IS} \right] = \frac{2\mu_{\mathrm{B}} g_{\mathrm{N}} \mu_{\mathrm{N}}}{\hbar^2} \frac{\ell(\ell+1)}{j(j+1)} \left\langle \frac{1}{r^3} \right\rangle \mathbf{IJ} + A_F \mathbf{IS}, \tag{2.160}$$

where

$$A_F = \frac{16\pi}{3} \frac{\mu_{\mathrm{B}} g_{\mathrm{N}} \mu_{\mathrm{N}}}{\hbar^2} |\psi(0)|^2, \qquad |\psi(0)|^2 = \langle \delta(\mathbf{r}) \rangle \,. \tag{2.161}$$

The first (dipole) term in (2.160) is zero when $\ell = 0$, while the second (contact) term is zero when $\ell \neq 0$, which means the two parts of the magnetic hyperfine interaction do not overlap. For s states we only need to consider the contact interaction. For non-s states, only the dipole term contributes. Similarly to what happens in the spin-orbit coupling, the form of the $\mathcal{H}_h$ Hamiltonian suggests that it will be convenient to couple the nuclear spin, $\mathbf{I}$, and the electron total angular momentum, $\mathbf{J}$, to form a new angular momentum operator (the total angular momentum of the atom),

$$\mathbf{F} = \mathbf{I} + \mathbf{J}. \tag{2.162}$$

We can use the notation $|Im_I\rangle$, $|jm\rangle$, $|Fm_F\rangle$, to designate the eigenfunctions corresponding to $\mathbf{I}$, $\mathbf{J}$ and $\mathbf{F}$, respectively.

The Fermi contact term in Hydrogen.

Let's consider now the hydrogen in an s state, where only the contact term contributes. We set $\ell = 0$ and replace $\mathbf{J}$ by $\mathbf{S}$, so that now $\mathbf{F} = \mathbf{I} + \mathbf{S}$ and $F^2 = I^2 + S^2 + 2\mathbf{IS}$. Thus

$$\mathbf{IS} = \frac{1}{2} \left(F^2 - S^2 - I^2 \right). \tag{2.163}$$

The spin of the proton is 1/2, therefore

$$I = \frac{1}{2}, \qquad s = \frac{1}{2}, \qquad F = 0, 1, \tag{2.164}$$

and

$$\begin{aligned} \langle Fm_F \,|\, \mathbf{IS} \,|\, Fm_F \rangle \equiv \langle \mathbf{IS} \rangle &= \left\langle Fm_F \left| \frac{1}{2} \left(F^2 - S^2 - I^2 \right) \right| Fm_F \right\rangle \\ &= \frac{\hbar^2}{2} \Big[F(F+1) - s(s+1) - I(I+1) \Big] = \frac{\hbar^2}{2} \Big[F(F+1) - \frac{3}{2} \Big], \end{aligned} \tag{2.165}$$

which shows that only diagonal elements are nonzero. For the two possible values of F we have

$$\langle \mathbf{IS} \rangle = \begin{cases} \dfrac{\hbar^2}{4}, & F = 1, \\ -\dfrac{3\hbar^2}{4}, & F = 0. \end{cases} \tag{2.166}$$

Since for an s-state $\mathcal{H}_h^{(\ell=0)} = A_F \mathbf{IS}$, the energies are

$$E = \begin{cases} \dfrac{\hbar^2}{4} A_F = \dfrac{4}{3}\pi\mu_{\rm B} g_{\rm N}\mu_{\rm N}|\psi(0)|^2, & F = 1, \\ -\dfrac{3\hbar^2}{4} A_F = -4\pi\mu_{\rm B} g_{\rm N}\mu_{\rm N}|\psi(0)|^2, & F = 0. \end{cases} \tag{2.167}$$

For Hydrogen we have $\mathbf{I}_p = \mathbf{S}_p$, so[17]

$$\boldsymbol{\mu}_p = \frac{1}{\hbar} g_{\rm N}\mu_{\rm N}\mathbf{I} = \frac{1}{\hbar} g_p \mu_{\rm N}\mathbf{I}_p = \frac{g_p e}{2m_p c}\mathbf{S}_p, \qquad g_p = 2 \times 2.7928 = 5.5857 \tag{2.168}$$

$$|\psi(0)|^2 = \frac{1}{\pi a_0^3}, \qquad \gamma\hbar = \mu_{\rm N} g_{\rm N} = 2.82 \times 10^{-23}\ \text{erg/G}, \qquad a_0 = 0.529 \times 10^{-8}\ \text{cm}. \tag{2.169}$$

Using the previous numerical values and $\mu_{\rm B} = 0.927 \times 10^{-20}$ erg/G, the energy difference between the two levels of the s state hyperfine splitting is

$$\begin{aligned} \Delta E = \hbar^2 A_F &= \frac{16\pi}{3}\mu_{\rm B} g_{\rm N}\mu_{\rm N}|\psi(0)|^2 = \frac{16}{3}\frac{\mu_{\rm B} g_{\rm N}\mu_{\rm N}}{a_0^3} \\ &= 9.42 \times 10^{-18}\,\text{erg} = 9.42 \times 10^{-25}\,\text{J} = 5.88 \times 10^{-6}\,\text{eV} \longrightarrow 0.04738\ \text{cm}^{-1}. \end{aligned} \tag{2.170}$$

To evaluate ΔE directly in the SI system we need to express the Bohr and nuclear magnetons in SI units. Equation (B.30) in the appendix on Gaussian units tells us that each magneton gets a factor of $\sqrt{\mu_0/4\pi}$, so we should mutiply the previous expression for ΔE by a factor of $\mu_0/4\pi$.[18] Thus

$$\begin{aligned} (\Delta E)_{\rm SI} = \left(\frac{\mu_0}{4\pi}\right)\frac{16}{3}\frac{\mu_{\rm B} g_{\rm N}\mu_{\rm N}}{a_0^3} &= \frac{4\pi \times 10^{-7}\,\text{H m}^{-1}}{4\pi} \\ \times \frac{16}{3}\frac{9.27 \times 10^{-24}\,\text{J T}^{-1} \times 5.586 \times 5.05 \times 10^{-27}\,\text{J T}^{-1}}{(5.29 \times 10^{-11})^3\,\text{m}^3} &= 9.42 \times 10^{-25}\,\text{J}. \end{aligned} \tag{2.171}$$

Using the fact that in the SI system $\varepsilon_0\mu_0 = 1/c^2$ and $a_0 = 4\pi\varepsilon_0\hbar^2/(m_e e^2) = \hbar/(mc\alpha)$, the energy difference can also be written as

$$\Delta E = \frac{4}{3}\frac{g_p\hbar^4}{m_p m_e^2 c^2 a_0^4} = \frac{4}{3} g_p \frac{m_e}{m_p}\alpha^4 m_e c^2. \tag{2.172}$$

This hyperfine splitting for Hydrogen is shown in figure (2.7). Since $\Delta E = h\nu$ for electromagnetic radiation, this corresponds to a frequency of $\nu = 1420.4058$ MHz, or to a wavelength of 21cm. Radiation of this frequency can be detected by radio-frequency techniques, and its observation can be used to map the concentration of cold atomic hydrogen in the interstellar regions of our galaxy.[19] It is found that the spiral arms of the galaxy contain the greatest concentrations of hydrogen atoms and the different arms can be identified by making use of the Doppler shift,

[17] The proton magnetic moment is measured to an extraordinary accuracy, $\mu_p = 2.7928473446 \pm 0.0000000008\,\mu_{\rm N}$ [14]. If we write $g_p = 2\mu_p c$, then $\boldsymbol{\mu}_p = \dfrac{\mu_p e}{m_p c}\mathbf{S}_p$, and the expectation value of the third component of $\boldsymbol{\mu}_p$ in a state with $m_s = 1/2$ is precisely $\langle(\boldsymbol{\mu}_p)_z\rangle = \mu_p\dfrac{e\hbar}{2m_p} = \mu_p\mu_{\rm N}$, so μ_p is the proton magnetic moment in units of $\mu_{\rm N}$.

[18] The permeability of free space (vacuum) in the SI system is $\mu_0 = 4\pi \times 10^{-7}\,\text{H m}^{-1}$. The energy and magnetic flux density equivalences are $1\,\text{erg} = 10^{-7}\,\text{J}$ and $1\,\text{gauss} = 10^{-4}\,\text{T}$.

[19] This radiation was theoretically predicted by the Dutch astronomer H.C. van de Hulstin in 1944 and was experimentally detected by American physicists Harold Ewen and Edward Purcell at Harvard University in 1951. The hyperfine transition responsible for this radiation is very rare, but there is so much atomic Hydrogen in our galaxy that can be detected on Earth. Furthermore, radio-frequency radiation can easily go through clouds of interstellar dust that block optical observations.

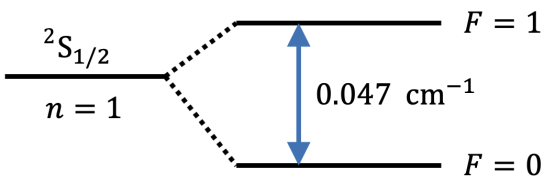


Figure 2.7: Magnetic hyperfine splitting of the ground state of hydrogen. This splitting is due only to the Fermi contact term.

since each arm possesses a different radial velocity with respect to the Earth. In contrast to atomic hydrogen, molecular hydrogen does not emit 21cm radiation and this technique cannot be applied to find regions with significant concentration of molecular hydrogen. Still, high-density galaxy regions contain other types of molecules with a suitable wavelength emission to be detected using radio-frequency techniques. In deuterium, with a nucleus consisting of a proton and a neutron, this wavelength has been measured to be about 92 cm. This energy splitting is much smaller than the separation energy between atomic states ($\sim$ eV) and nuclear states ($\sim$ MeV), therefore it is safe to assume that only the ground states of the atom and the nucleus play a role in the magnetic hyperfine interaction.

It is very instructive to evaluate the magnitude of the effective magnetic fields at the position of the electron and the proton due to their spins. For the electron we have seen (equations (2.25) and (2.31)) that the $^2S_{1/2}$ level is split into two states when the atom is placed in an external magnetic field. The two states are separated by an energy $\Delta E = 2\mu_B B$. The magnetic field due to the proton spin produces also a splitting which we could write as $\Delta E = g_e \mu_B B^e_{\text{eff}}$. It is found that B^e_{eff} comes out to be about 508 G.

$$B^e_{\text{eff}} = \frac{\Delta E}{g_e \mu_B} = \frac{9.42 \times 10^{-25}\,\text{J}}{2 \times 9.27 \times 10^{-24}\,\text{J/T}} = 5.08 \times 10^{-2}\,\text{T} = 508\,\text{G} \tag{2.173}$$

Similarly, we can find the effective magnetic field at the proton location by writing $\Delta E = g_N \mu_N B^p_{\text{eff}} = \gamma \hbar B^p_{\text{eff}}$, where $g_N \mu_N = \gamma \hbar$ is the magnetic moment of the proton, in analogy with $g_e \mu_B$, which is the magnetic moment of the electron.

$$B^p_{\text{eff}} = \frac{\Delta E}{g_p \mu_N} = \frac{9.42 \times 10^{-25}\,\text{J}}{5.586 \times 5.05 \times 10^{-27}\,\text{J/T}} = 33.4\,\text{T} = 3.34 \times 10^5\,\text{G}. \tag{2.174}$$

This shows that very high magnetic fields exist at the position of the nucleus due to the large magnetic moment of the electron. As a comparison, note that the magnitude of the earth's magnetic field at its surface ranges from 25 to 65 microteslas. Another interesting comparison can be made with the superconducting magnets of the Large Hadron Collider at CERN, which are able to deliver 8.3 Tesla, a bit less that the 33 Tesla that the electron creates at the center of the atom. At CERN, the electromagnets use a current of 1.108×10^4 amperes to produce the field, and a superconducting coil working at 2 K (-271.3 C) allows the high currents to flow without losing energy due to electrical resistance.[20]

The dipole term.

Consider now the dipole term in (2.160). For hydrogenic atoms we have

$$\left\langle \frac{1}{r^3} \right\rangle = \frac{Z^3}{a_0^3} \frac{1}{n^3(\ell+1)(\ell+\frac{1}{2})\ell}. \tag{2.175}$$

[20] Much more powerful electromagnets working at room temperature have been built. However, it is very difficult to design *reusable* magnets delivering more than 100 Tesla, since the enormous mechanical stress generated when they are switched on and off can easily tear them apart.

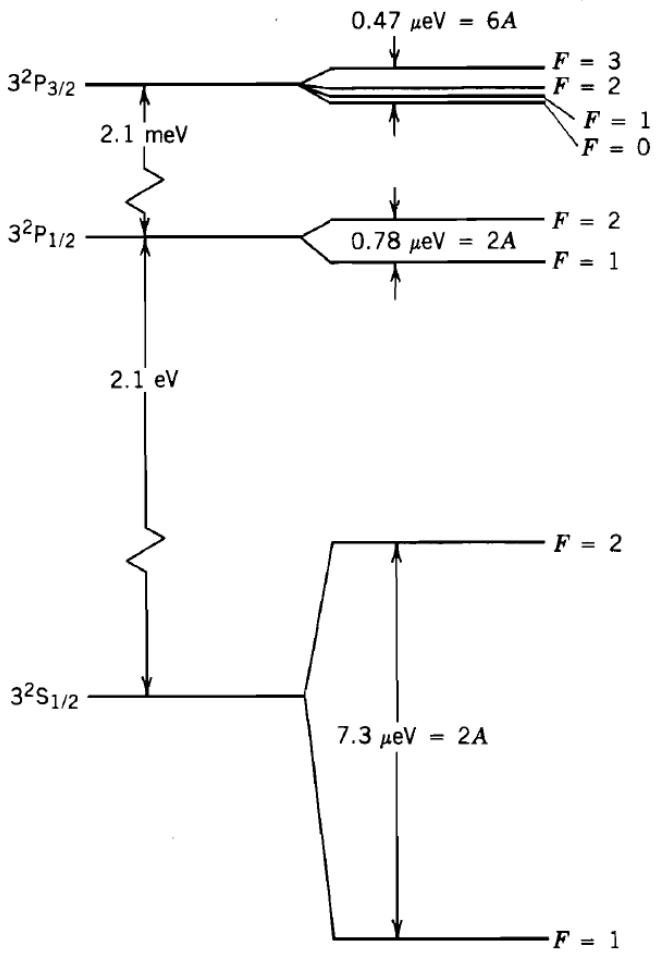


Figure 2.8: Hyperfine splitting in sodium ($I = 3/2$). Each multiplet of F states has a different value of the splitting parameter A. Figure taken from *Krane, Introductory Nuclear Physics, John Wiley & Sons, 1988.*

We need also the relations (remember that $\mathbf{F} = \mathbf{I} + \mathbf{J}$),

$$\mathbf{IJ} = \frac{1}{2}\left(F^2 - I^2 - J^2\right), \qquad \langle F m_F \,|\, \mathbf{IJ} \,|\, F m_F \rangle = \frac{\hbar^2}{2}\Big[F(F+1) - I(I+1) - j(j+1)\Big]. \quad (2.176)$$

Here again only diagonal matrix elements are nonzero, therefore the energies are given by

$$E = \frac{Z^3 \mu_{\rm B} g_{\rm N} \mu_{\rm N}}{a_0^3 n^3 j(j+1)(\ell+\frac{1}{2})}\Big[F(F+1) - I(I+1) - j(j+1)\Big]. \quad (2.177)$$

For $2\,^2{\rm P}_{1/2}$, $Z = 1$, $j = 1/2$, $I = 1/2$, $n = 2$ and $\ell = 1$, we have

$$(E)_{F=1} = \frac{\mu_{\rm B} g_{\rm N} \mu_{\rm N}}{18 a_0^3}, \qquad (E)_{F=0} = -\frac{\mu_{\rm B} g_{\rm N} \mu_{\rm N}}{6 a_0^3}, \qquad \Delta E = -\frac{2\mu_{\rm B} g_{\rm N} \mu_{\rm N}}{9 a_0^3}, \quad (2.178)$$

which is a factor 24 smaller than the splitting due to the contact interaction. Similarly, for $2\,^2{\rm P}_{3/2}$ the energies are

$$(E)_{F=2} = \frac{\mu_{\rm B} g_{\rm N} \mu_{\rm N}}{30 a_0^3}, \qquad (E)_{F=1} = -\frac{\mu_{\rm B} g_{\rm N} \mu_{\rm N}}{18 a_0^3}, \qquad \Delta E = \frac{4\mu_{\rm B} g_{\rm N} \mu_{\rm N}}{45 a_0^3}. \quad (2.179)$$

In figure (2.8) we can see the hyperfine splitting in several levels of ^{23}Na. The electronic configuration of Na is that of Ne plus a single electron in a $3s$ orbital, therefore the Na ground

state is $^2S_{1/2}$. The two first excited Na states correspond to the spin–orbit doublet $^2P_{1/2}$ and $^2P_{3/2}$. These states are separated by about 2.1×10^{-3} eV, an amount considerably larger than the fine structure splitting in Hydrogen, but this is expected because the spin-orbit interaction goes like Z^4 and we are dealing here with $Z = 11$. The ^{23}Na nucleus has a spin of $I = 3/2$, therefore the coupling with $J = 1/2$ and $J = 3/2$ generates two doublets ($F = 1, 2$) and a quartet ($F = 0, 1, 2, 3$), as indicated in the figure. The hyperfine structure can be measured with high precision using laser beams.

Magnetic hyperfine interaction in the presence of an external magnetic field.

An important practical case is that of magnetic resonance experiments, where we need to consider the magnetic hyperfine interaction in the presence of an external magnetic field. Let's study this situation for Hydrogen in an s state, where only the Fermi contact term contributes. We need to use equation (2.7), namely $\mathcal{H}_{\rm m} = \frac{1}{\hbar}\mu_{\rm B}\,\mathbf{B}(\mathbf{L} + 2\mathbf{S})$, with $\mathbf{L} = 0$, add the contact term and take into account the interaction of the nuclear spin, $\mathbf{I}$, with the external magnetic field, $\mathbf{B}$. Therefore the Hamiltonian may be written as

$$\mathcal{H} \;=\; \frac{2\mu_{\rm B}}{\hbar}\mathbf{BS} + A_F\mathbf{IS} - \gamma\mathbf{BI} \;=\; \frac{g_e\mu_{\rm B}}{\hbar}\mathbf{BS} + A_F\mathbf{IS} - \frac{g_{\rm N}\mu_{\rm N}}{\hbar}\mathbf{BI}. \tag{2.180}$$

In any case, the nuclear magnetic moment is very small compared to the electronic magnetic moment, so the contribution of the term $\gamma\mathbf{BI}$ can be neglected for our purposes here. The magnetic field felt by the s electron due to the nuclear magnetic moment is about 500 G. If the external magnetic field is large compared to this value, the electronic and nuclear angular momenta will be decoupled. The argument here is similar to the one used in the case of strong versus weak fields to study the spin-orbit coupling. To find the eigenvalues of $\mathcal{H} = (2\mu_{\rm B}/\hbar)\mathbf{BS} + A_F\mathbf{IS}$ over a wide range of external magnetic field amplitudes, it will be more convenient to use the uncoupled representation, $|sIm_sm_I\rangle$. Since both s and I are fixed to 1/2, the notation can be shortened to $|m_sm_I\rangle$. Choosing the external magnetic field $\mathbf{B}$ along the z-axis, we can write $\mathcal{H}$ in the following form,

$$\mathcal{H} = \frac{2\mu_{\rm B}}{\hbar}BS_z + A_F\big[-I_{+1}S_{-1} + I_0S_0 - I_{-1}S_{+1}\big], \tag{2.181}$$

where we have used the spherical components of $\mathbf{I}$ and $\mathbf{S}$ to write $\mathbf{IS} = -I_{+1}S_{-1} + I_0S_0 - I_{-1}S_{+1}$, as we did before in equation (1.186) for $\mathbf{LS}$ when computing the spin–orbit coupling. We need to fill up a Hamiltonian 4×4 matrix for m_s and m_I, each taking on the values $\pm1/2$. The matrix elements for the $2\mu_{\rm B}BS_z/\hbar$ term are easy to get, they are all zero except those in the diagonal. For the terms containing the spherical operators we need to use the well–known relations that can be seen in appendix B.5. The results are collected in table (2.3), and the corresponding secular equations deliver the following *four* eigenvalues,

$$E = \frac{\hbar^2}{4}A_F \pm \mu_{\rm B}B, \qquad E = -\frac{\hbar^2}{4}A_F \pm \frac{\hbar^2}{2}A_F\sqrt{1 + \left(\frac{2\mu_{\rm B}B}{\hbar^2A_F}\right)^2},$$

which can also be written as

$$E = \frac{\hbar^2}{4}A_F\left[1 \pm 2\left(\frac{2\mu_{\rm B}B}{\hbar^2A_F}\right)\right], \qquad E = \frac{\hbar^2}{4}A_F\left(-1 \pm 2\sqrt{1 + \left(\frac{2\mu_{\rm B}B}{\hbar^2A_F}\right)^2}\right). \tag{2.182}$$

These energy eigenvalues are shown in figure (2.9). We should examine several interesting limiting cases:

- For $B = 0$, the first of equations (2.182) gives $\frac{\hbar^2}{4}A_F$ (twice), while the second gives $\frac{\hbar^2}{4}A_F$ and $-\frac{3\hbar^2}{4}A_F$. The eigenvalue $\frac{\hbar^2}{4}A_F$ is three-fold degenerate due to the degeneracy associated with $F = 1$, and we recover exactly the results in equation (2.167).

	$\left\vert\frac{1}{2}\frac{1}{2}\right\rangle$	$\left\vert\frac{1}{2}-\frac{1}{2}\right\rangle$	$\left\vert-\frac{1}{2}\frac{1}{2}\right\rangle$	$\left\vert-\frac{1}{2}-\frac{1}{2}\right\rangle$
$\left\langle\frac{1}{2}\frac{1}{2}\right\vert$	$\mu_{\mathrm{B}}B+\frac{\hbar^2}{4}A_F$			
$\left\langle\frac{1}{2}-\frac{1}{2}\right\vert$		$\mu_{\mathrm{B}}B-\frac{\hbar^2}{4}A_F$	$\frac{\hbar^2}{2}A_F$	
$\left\langle-\frac{1}{2}\frac{1}{2}\right\vert$		$\frac{\hbar^2}{2}A_F$	$-\mu_{\mathrm{B}}B-\frac{\hbar^2}{4}A_F$	
$\left\langle-\frac{1}{2}-\frac{1}{2}\right\vert$				$-\mu_{\mathrm{B}}B+\frac{\hbar^2}{4}A_F$

Table 2.3: Matrix elements of the magnetic hyperfine interaction in the hydrogen s state in the presence of an external magnetic field.

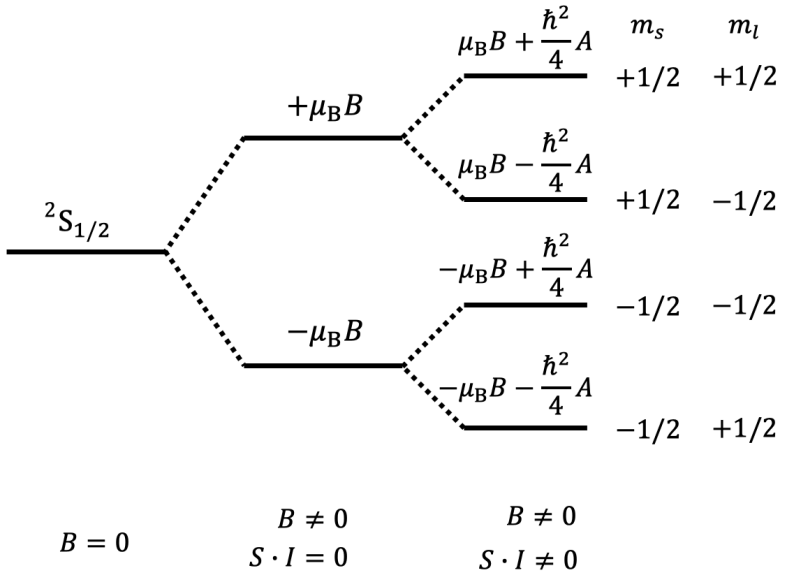


Figure 2.9: First–order splitting of the hydrogen ground state in a magnetic field with hyperfine interaction (adapted from Weissbluth, 1978).

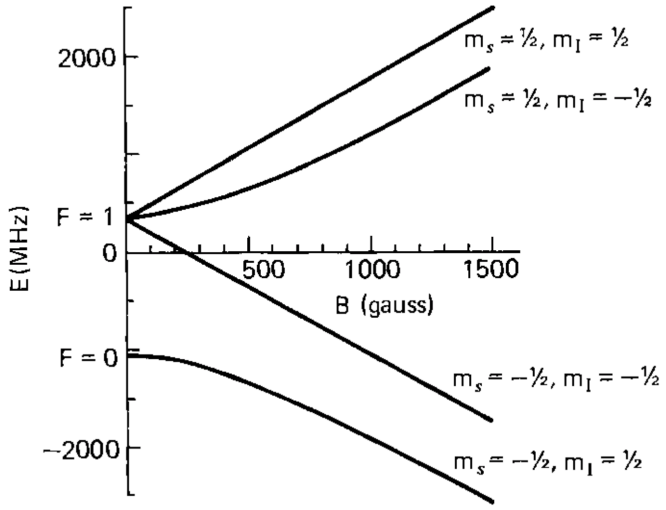


Figure 2.10: Variation with an external magnetic field of the energy levels of the hydrogen ground state in a magnetic field with hyperfine splitting (adapted from Weissbluth, 1978).

- If $\mu_{\rm B}B \ll \hbar^2 A_F$ we take the approximation $\sqrt{1+x^2} \approx 1 + x^2/2$ for small x in the square root of (2.182). Two of the energies get thereby a quadratic dependence on the field B.

$$E = \frac{\hbar^2}{4}A_F \pm \mu_{\rm B}B, \qquad E = \frac{\hbar^2}{4}A_F + \frac{{\mu_{\rm B}}^2 B^2}{\hbar^2 A_F}, \qquad E = -\frac{3\hbar^2}{4}A_F - \frac{{\mu_{\rm B}}^2 B^2}{\hbar^2 A_F}.$$

- if $\mu_{\rm B}B \gg \hbar^2 A_F$ we can do $\sqrt{1+x^2} \approx x$ for large x in (2.182) and get the following list of values, which appear also as the diagonal elements in the matrix shown in table (2.3). Using that table we can associate each energy eigenvalue with its corresponding values of m_s and m_I, as shown below

E	m_s	m_I
$\mu_{\rm B}B + (\hbar^2 A_F/4)$	$1/2$	$1/2$
$\mu_{\rm B}B - (\hbar^2 A_F/4)$	$1/2$	$-1/2$
$-\mu_{\rm B}B - (\hbar^2 A_F/4)$	$-1/2$	$1/2$
$-\mu_{\rm B}B + (\hbar^2 A_F/4)$	$-1/2$	$-1/2$

  Note that, when $\mu_{\rm B}B \gg \hbar^2 A_F$, we can replace the Hamiltonian (2.181) by

$$\mathcal{H} = \frac{2\mu_{\rm B}}{\hbar}BS_z + A_F I_z S_z, \tag{2.183}$$

  because this gives also the same energy eigenvalues shown in the previous table. The case $\mu_{\rm B}B \gg \hbar^2 A_F$ may be regarded as the Paschen–Back region for the magnetic hyperfine interaction.

All this is somehow similar to what we have observed when treating simultaneously the spin-orbit and external magnetic field interactions. When the magnetic field is weak ($\ll$ 500 G in hydrogen), the electronic angular momentum and the nuclear spin are coupled to form states with $F = 0, 1$, and the levels spread linearly with the field. At larger field values the dependence is quadratic. At still larger field values ($\gg$ 500 G) the electronic angular momentum and the nuclear spin become uncoupled and the dependence on the field becomes linear again. The *good* quantum numbers at very low fields are s, I, F, m_F. At very high fields the *good* quantum numbers are s, I, m_s, m_I. At intermediate fields the levels are mixed and neither collection of quantum numbers provides an adequate set of labels. Figure (2.11) shows the level scheme for an s electron interacting with a nucleus with spin $I = 1$ through the magnetic hyperfine interaction. This occurs, for example, in ^{14}N, which has a nucleus with spin $I = 1$.

2.3.3 Electric quadrupole interaction.

In classical electrodynamics the electric potential created by any shape of charge distribution can be described by a multipole expansion, and this is also the case in the quantum domain. The electric potential generated by a spherical distribution of charge is just of the simple Coulombian form, $-Ze/r$, which corresponds to the monopole term in the multipole expansion. But nuclei are not necessarily spherical, specially the heavy ones. They can be classified as having either a prolate ellipsoidal (football) shape exhibiting a positive electric quadrupole moment (to be defined later), or an oblate shape creating a negative quadrupole moment. The dipole term vanishes, as we will soon show.

In a coordinate system centered inside the nucleus the electrostatic interaction energy between a single electron and a nucleus containing Z protons is

$$W = -\sum_{p=1}^{Z} \frac{e^2}{|\mathbf{r}_e - \mathbf{r}_p|}, \tag{2.184}$$

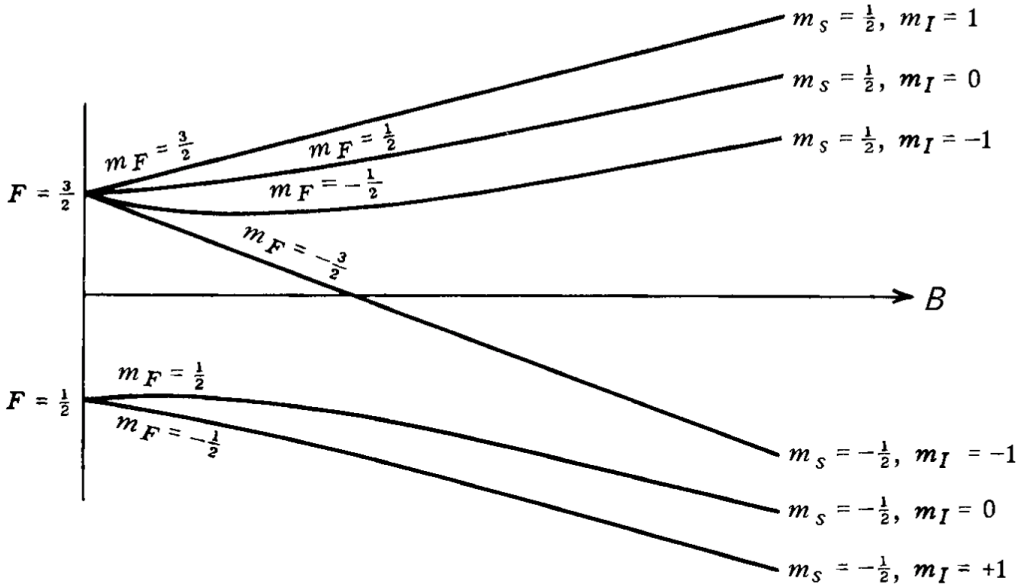


Figure 2.11: Magnetic hyperfine levels for a nucleus with spin $I = 1$ and positive coupling constant as a function of the magnetic field (taken from Weissbluth, 1978, which is adapted from *Introduction to Electron Paramagnetic Resonance*, by M. Bersohn and J. C. Baird, Addison-Wesley / W. A. Benjamin Inc.).

where $\mathbf{r}_e$ is the electron position vector and $\mathbf{r}_p$ is the position vector of the pth proton. The sum is extended to all Z protons. It is found convenient to use the spherical harmonic expansion for $1/|\mathbf{r}_e - \mathbf{r}_p|$,

$$\frac{1}{|\mathbf{r}_e - \mathbf{r}_p|} = 4\pi \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \frac{1}{2\ell+1} \frac{r_<^\ell}{r_>^{\ell+1}} Y_\ell^{m*}(\theta_p, \varphi_p) Y_\ell^m(\theta_e, \varphi_e), \tag{2.185}$$

where the notation $r_>$ means the greater of the two distances (r_e, r_p), and $r_<$ means the smaller. Later we will show that there is no quadrupole interaction for s states, therefore we can assume that the electron does never venture inside the nucleus and take always $r_e > r_p$, so the electrostatic energy becomes

$$W = -4\pi e^2 \sum_{p=1}^{Z} \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \frac{1}{2\ell+1} \frac{r_p^\ell}{r_e^{\ell+1}} Y_\ell^{m*}(\theta_p, \varphi_p) Y_\ell^m(\theta_e, \varphi_e). \tag{2.186}$$

It will be easier to compute the matrix elements if we write the Hamiltonian in terms of the scalar product of irreducible tensor operators.[21] The set of the $2\ell + 1$ spherical harmonic operators $\mathbf{Y}^{(\ell)} = \{Y_\ell^{-\ell}, Y_\ell^{-\ell+1}, \ldots, Y_\ell^{\ell-1}, Y_\ell^\ell\}$ is an example of irreducible tensor operator of rank ℓ. Under a coordinate rotation, $P_R \mathbf{r} = \mathbf{r}'$, the components of this tensor transform as a linear combination of themselves,

$$P_R Y_\ell^m P_R^{-1} = \sum_{k=-\ell}^{\ell} D_{km} Y_\ell^k. \tag{2.187}$$

The scalar product of two irreducible tensor operators of rank k is defined as

$$\mathbf{T}^{(k)} \cdot \mathbf{U}^{(k)} = \sum_q (-1)^q \, T_q^{(k)} U_{-q}^{(k)}. \tag{2.188}$$

[21] For a short review on this topic see appendix (B.6).

Since $Y_\ell^{-m} = (-1)^m Y_\ell^{m*}$, we can write

$$\sum_{m=-\ell}^{\ell} Y_\ell^{m*} Y_\ell^m = \sum_{m=-\ell}^{\ell} (-1)^m Y_\ell^{-m} Y_\ell^m = \mathbf{Y}^{(\ell)}(\mathbf{\Omega}_p)\mathbf{Y}^{(\ell)}(\mathbf{\Omega}_e), \tag{2.189}$$

$$\mathbf{\Omega}_p \equiv (\theta_p, \varphi_p), \qquad \mathbf{\Omega}_e \equiv (\theta_e, \varphi_e), \tag{2.190}$$

Defining

$$\mathbf{Q}^{(\ell)} \equiv e \sum_{p=1}^{Z} \sqrt{\frac{4\pi}{2\ell+1}}\, r_p^\ell \mathbf{Y}^{(\ell)}(\mathbf{\Omega}_p), \qquad \mathbf{U}^{(\ell)} \equiv -e\sqrt{\frac{4\pi}{2\ell+1}}\, \frac{1}{r_e^{\ell+1}} \mathbf{Y}^{(\ell)}(\mathbf{\Omega}_e). \tag{2.191}$$

We have,

$$W = \sum_{\ell=0}^{\infty} \mathbf{Q}^{(\ell)}\, \mathbf{U}^{(\ell)}. \tag{2.192}$$

Note that when $\ell = 0$ we have (since $Y_0^0 = 1/\sqrt{4\pi}$),

$$\mathbf{Q}^{(0)} = e \sum_{p=1}^{Z} \sqrt{4\pi}\, \mathbf{Y}^{(0)}(\mathbf{\Omega}_p) = Ze, \qquad \mathbf{U}^{(0)} = -e\sqrt{4\pi}\, \frac{1}{r_e} \mathbf{Y}^{(0)}(\mathbf{\Omega}_e) = -\frac{e}{r_e}, \tag{2.193}$$

$$\mathbf{Q}^{(0)}\, \mathbf{U}^{(0)} = -\frac{Ze^2}{r_e}. \tag{2.194}$$

This is just the Coulomb interaction energy of the electron with a point-like nucleus of charge Ze (Z being the sum over the protons), that is, the monopole contribution.

The contribution of the dipole term, $\mathbf{Q}^{(1)}\, \mathbf{U}^{(1)}$, is zero because it corresponds to the interaction of a nuclear electric dipole moment with the electric field created by the electrons, but nuclei states do not have electric dipole moments. This is true to the extent that they have a well-defined parity (we neglect here any possible tiny contribution from weak interactions). In other words, the electric dipole moment $\mathbf{p}$ is an odd operator (it changes sign under a spatial inversion $\mathbf{p} \to -\mathbf{p}$), therefore its expectation value in a nuclear state with a well-defined parity vanishes,

$$\langle \mathbf{p} \rangle = \int \psi^*\, \mathbf{p}\, \psi\, \mathrm{d}\tau = 0. \tag{2.195}$$

The next term we should then consider is the quadrupole interaction, with $\ell = 2$,

$$\mathcal{H}_Q = \mathbf{Q}^{(2)}\, \mathbf{U}^{(2)}. \tag{2.196}$$

We take the electronic state to be characterized by the angular momentum quantum numbers j, m_j, and the nuclear state by I, m_I. To compute the interaction energy associated with $\mathcal{H}_Q$ we will use the coupled representation, with the total angular momentum given by $\mathbf{F} = \mathbf{I} + \mathbf{J}$ and the states specified by the quantum numbers I, j, F, m_F. The matrix element for the scalar product of two tensor operators of rank k, taken between two states with angular momentum quantum numbers $|j_1 j_2 j m\rangle$ and $|j_1' j_2' j' m'\rangle$ is given by

$$\begin{aligned}\left\langle j_1 j_2 j m \left| \mathbf{T}^{(k)} \mathbf{U}^{(k)} \right| j_1' j_2' j' m' \right\rangle \\ = (-1)^{j+j_2+j_1'}\, \delta_{jj'} \delta_{mm'} \left\{ \begin{array}{ccc} j_1' & j_2' & j \\ j_2 & j_1 & k \end{array} \right\} \left\langle j_1 \left\| \mathbf{T}^{(k)} \right\| j_1' \right\rangle \left\langle j_2 \left\| \mathbf{U}^{(k)} \right\| j_2' \right\rangle,\end{aligned} \tag{2.197}$$

where the quantity in curly brackets is a $6j$ symbol defined in (B.193). For our particular case this equation transforms into

$$\begin{aligned}\left\langle I j F m_F \left| \mathbf{Q}^{(2)} \mathbf{U}^{(2)} \right| I j F' m_F' \right\rangle \\ = (-1)^{F+j+I}\, \delta_{FF'} \delta_{m_F m_F'} \left\{ \begin{array}{ccc} I & j & F \\ j & I & 2 \end{array} \right\} \left\langle I \left\| \mathbf{Q}^{(2)} \right\| I \right\rangle \left\langle j \left\| \mathbf{U}^{(2)} \right\| j \right\rangle.\end{aligned} \tag{2.198}$$

Here the $6j$ symbol has the value

$$\left\{ \begin{array}{ccc} I & j & F \\ j & I & 2 \end{array} \right\} = \frac{(-1)^{F+I+j}\ \ 2\big[3X(X-1) - 4I(I+1)j(j+1)\big]}{\sqrt{(2I-1)2I(2I+1)(2I+2)(2I+3)(2j-1)2j(2j+1)(2j+2)(2j+3)}}, \tag{2.199}$$

where

$$X = I(I+1) + j(j+1) - F(F+1). \tag{2.200}$$

To proceed, the reduced matrix elements must be computed. To evaluate $\left\langle I \left\| \mathbf{Q}^{(2)} \right\| I \right\rangle$ we are going to relate it to the nuclear electric quadrupole moment, which is a second-order symmetric tensor whose components $\mathcal{Q}_{ij}$ are defined in the following way (see equation (B.301) for the classical definition of the quadrupole moment tensor). Let $\mathbf{r}_p$ be the position vector of a proton with respect to the center of mass of the nucleus, and let $x_p = (\mathrm{x}_p)_1$, $y_p = (\mathrm{x}_p)_2$, $z_p = (\mathrm{x}_p)_3$ be its Cartesian components. Then

$$\mathcal{Q}_{ij} = \sum_p 3(\mathrm{x}_p)_i(\mathrm{x}_p)_j - \delta_{ij} r_p^2, \tag{2.201}$$

where the sum is taken over all the protons in the nucleus. Note that $\mathcal{Q}_{ij}$ is a traceless tensor, $\mathcal{Q}_{11}+\mathcal{Q}_{22}+\mathcal{Q}_{33} = 0$. It is customary to define the *magnitude* $\mathcal{Q}$ of the nuclear electric quadrupole moment as the average value of the component $\mathcal{Q}_{zz} = \mathcal{Q}_{33}$ in the state with the maximum value of the projection quantum number, $|I, m_I = I\rangle$,

$$\mathcal{Q} = \left\langle I, m_I = I \left| \sum_p (3z_p^2 - r_p^2) \right| I, m_I = I \right\rangle. \tag{2.202}$$

This quantity $\mathcal{Q}$ has dimensions of area and is often measured in barns.[22] To quote a number, the deuteron has an electric quadrupole moment of about $\mathcal{Q} = 0.0028$ barns. For a charge distribution with spherical symmetry one would have

$$\left\langle r_p^2 \right\rangle = \left\langle x_p^2 + y_p^2 + z_p^2 \right\rangle = 3\left\langle z_p^2 \right\rangle, \tag{2.203}$$

which leads to $\mathcal{Q} = 0$ for a sphere. If the charge distribution is elongated along the direction of $\mathbf{I}$, $\mathcal{Q}$ is positive. If the charge distribution is flattened (oblate form) $\mathcal{Q}$ is negative. The $\mathcal{Q}_{zz}$ component can be written using the $Y_2^0 = \frac{1}{4}\sqrt{\frac{5}{\pi}}(3\cos^2\theta - 1)$ spherical harmonic as

$$\sum_p (3z_p^2 - r_p^2) = 2\sum_p \sqrt{\frac{4\pi}{5}} r_p^2\, Y_2^0(\mathbf{\Omega}_p). \tag{2.204}$$

On the other hand, from (2.191) we have, for $\ell = 2$ and $m = 0$,

$$Q_0^{(2)} = e\sum_p \sqrt{\frac{4\pi}{5}}\, r_p^2\, Y_2^0(\mathbf{\Omega}_p). \tag{2.205}$$

Comparing the last two expressions we see that

$$Q_0^{(2)} = \frac{1}{2} e \sum_p (3z_p^2 - r_p^2). \tag{2.206}$$

Therefore the matrix element of the 0–component of the tensor $\mathbf{Q}^{(2)}$ is related to the electric quadrupole moment in the following way.

$$\left\langle I, m_I = I \left| Q_0^{(2)} \right| I, m_I = I \right\rangle = \frac{1}{2} e \left\langle I, m_I = I \left| \sum_p (3z_p^2 - r_p^2) \right| I, m_I = I \right\rangle = \frac{1}{2} e\mathcal{Q}. \tag{2.207}$$

[22] $1\,\mathrm{barn} = 10^{-24}\,\mathrm{cm}^2 = 10^{-28}\,\mathrm{m}^2$.

This can be used to compute $\left\langle I \left\| \mathbf{Q}^{(2)} \right\| I \right\rangle$ by means of the Wigner-Eckart theorem. Indeed, we have (see equation (B.229) in the appendix)

$$\left\langle Im_I \left| Q_0^{(2)} \right| Im_I \right\rangle = (-1)^{I-m_I} \begin{pmatrix} I & 2 & I \\ -m_I & 0 & m_I \end{pmatrix} \left\langle I \left\| \mathbf{Q}^{(2)} \right\| I \right\rangle, \tag{2.208}$$

with the $3j$ symbol

$$\begin{pmatrix} j & j & 2 \\ m & -m & 0 \end{pmatrix} = \begin{pmatrix} j & 2 & j \\ -m & 0 & m \end{pmatrix} = (-1)^{j-m} \frac{3m^2 - j(j+1)}{\sqrt{(2j+3)(j+1)(2j+1)j(2j-1)}}. \tag{2.209}$$

Setting $m_I = I$,

$$\begin{aligned} \begin{pmatrix} I & 2 & I \\ -I & 0 & I \end{pmatrix} &= (-1)^{I-I} \frac{3I^2 - I(I+1)}{\sqrt{(2I+3)(I+1)(2I+1)I(2I-1)}} \\ &= \frac{I(2I-1)}{\sqrt{(2I+3)(I+1)(2I+1)I(2I-1)}} = \sqrt{\frac{I(2I-1)}{(2I+3)(I+1)(2I+1)}}. \end{aligned} \tag{2.210}$$

Note that this coefficient is zero for $I = 0, \frac{1}{2}$, which means that nuclei with spins $I = 0, \frac{1}{2}$ have no quadrupole moment. Inserting this back into the Wigner-Eckart theorem,

$$\left\langle I, m_I = I \left| Q_0^{(2)} \right| I, m_I = I \right\rangle = \sqrt{\frac{I(2I-1)}{(2I+3)(I+1)(2I+1)}} \left\langle I \left\| \mathbf{Q}^{(2)} \right\| I \right\rangle = \frac{1}{2} e\mathcal{Q}, \tag{2.211}$$

so

$$\left\langle I \left\| \mathbf{Q}^{(2)} \right\| I \right\rangle = \frac{1}{2} \sqrt{\frac{(2I+3)(I+1)(2I+1)}{I(2I-1)}}\, e\mathcal{Q} \tag{2.212}$$

To compute $\left\langle j \left\| \mathbf{U}^{(2)} \right\| j \right\rangle$ we follow a similar approach. We define a quantity eq such that it is equal to twice the expectation value of $U_0^{(2)}$ in a state with the maximum value of the projection quantum number, $|j, m_j = j\rangle$.

$$\left\langle j, m_j = j \left| U_0^{(2)} \right| j, m_j = j \right\rangle = \frac{1}{2} eq. \tag{2.213}$$

From (2.191) we have, for $\ell = 2$ and $m = 0$,

$$U_0^{(2)} = -e\sqrt{\frac{4\pi}{5}} \frac{1}{r_e^3} Y_2^0(\mathbf{\Omega}_e) = -e\sqrt{\frac{4\pi}{5}} \frac{1}{r_e^3} \sqrt{\frac{5}{4\pi}} \left(\frac{3}{2} \cos^2 \theta_e - \frac{1}{2} \right) = -\frac{e}{2} \frac{3z_e^2 - r_e^2}{r_e^5}. \tag{2.214}$$

This can be related to the zz component of an electric field gradient tensor. Remember that the components of this (symmetric) tensor are given by

$$\frac{\partial^2}{\partial r_i \partial r_j} \left(\frac{-e}{r} \right) = -e \left(\frac{3r_i r_j - r^2 \delta_{ij}}{r^5} \right) = V_{r_i r_j} \equiv V_{ij} = V_{ji}, \qquad (r_i, r_j = x, y, z). \tag{2.215}$$

Therefore the expression

$$\frac{\partial^2}{\partial z^2} \left(-\frac{e}{r} \right) = -e \frac{3z^2 - r^2}{r^5} = V_{zz}, \tag{2.216}$$

is the zz component of the electric field gradient tensor produced by the electron at a point whose coordinates with respect to that electron are (x, y, z). We see then that, since we are working in a reference frame centered at the nucleus, the quantity $2U_0^{(2)}$ is the zz component of the electric field gradient tensor at the nucleus produced by an electron at $\mathbf{r}_e$,

$$U_0^{(2)} = \frac{1}{2} \left(\frac{\partial^2 V}{\partial z^2} \right)_0 = -\frac{e}{2} \frac{3z^2 - r^2}{r^5} = \frac{1}{2} V_{zz}, \tag{2.217}$$

where $V = -e/r$ is the potential due to the electron, and the second derivative is evaluated at the origin (nucleus). Therefore

$$eq = \langle j, m_j = j \,|\, V_{zz} \,|\, j, m_j = j\rangle = \langle jj \,|\, V_{zz} \,|\, jj\rangle = \langle V_{zz}\rangle\,, \tag{2.218}$$

which is the expectation value of V_{zz} over the electronic state $|jj\rangle$. Now we apply again the Wigner–Eckart theorem to get, similarly to equation (2.212),

$$\left\langle j \left\| \mathbf{U}^{(2)} \right\| j \right\rangle = \frac{1}{2}\sqrt{\frac{(2j+3)(j+1)(2j+1)}{j(2j-1)}}\, eq. \tag{2.219}$$

The insertion of (2.219), (2.212) and (2.199) into (2.198), leads to

$$\begin{aligned}
&\left\langle IjFm_F \left| \mathbf{Q}^{(2)}\mathbf{U}^{(2)} \right| IjF'm'_F \right\rangle \\
&\quad= \frac{(-1)^{F+j+I}\,\delta_{FF'}\delta_{m_F m'_F}\,(-1)^{F+I+j}\;\; 2\big[3X(X-1) - 4I(I+1)j(j+1)\big]}{\sqrt{(2I-1)2I(2I+1)(2I+2)(2I+3)(2j-1)2j(2j+1)(2j+2)(2j+3)}} \\
&\qquad\times \frac{1}{4}\sqrt{\frac{(2I+3)(I+1)(2I+1)}{I(2I-1)}}\;\sqrt{\frac{(2j+3)(j+1)(2j+1)}{j(2j-1)}}\, e^2q\mathcal{Q} \\
&\quad= \frac{\big[3X(X-1) - 4I(I+1)j(j+1)\big]}{2\sqrt{(2I-1)2I(2I+2)(2j-1)2j(2j+2)}}\;\sqrt{\frac{(I+1)}{I(2I-1)}}\;\sqrt{\frac{(j+1)}{j(2j-1)}}\, e^2q\mathcal{Q} \\
&\qquad= \frac{e^2q\mathcal{Q}}{2I(2I-1)j(2j-1)}\left[\frac{3}{4}X(X-1) - I(I+1)j(j+1)\right],
\end{aligned} \tag{2.220}$$

where $e^2q\mathcal{Q}$ is known as the *quadrupole coupling constant*, which may be positive or negative.

The Hamiltonian $\mathcal{H}_Q = \mathbf{Q}^{(2)}\mathbf{U}^{(2)}$ can be written in another useful form. Using the relations

$$\mathbf{F} = \mathbf{I} + \mathbf{J}, \qquad -2\mathbf{IJ} = I^2 + J^2 - F^2, \tag{2.221}$$

we have

$$\langle IjFm_F \,|\, -2\mathbf{IJ} \,|\, IjFm_F\rangle = \hbar^2\Big[I(I+1) + j(j+1) - F(F+1)\Big] = \hbar^2 X. \tag{2.222}$$

$$\langle \mathbf{IJ}\rangle = -\frac{\hbar^2}{2}X, \qquad 3\left\langle(\mathbf{IJ})^2\right\rangle = \frac{3\hbar^4}{4}X^2, \qquad \frac{3\hbar^2}{2}\langle \mathbf{IJ}\rangle = -\frac{3\hbar^4}{4}X. \tag{2.223}$$

Therefore,

$$\frac{1}{\hbar^4}\left\langle IjFm_F \left| 3(\mathbf{IJ})^2 + \frac{3\hbar^2}{2}(\mathbf{IJ}) - I^2J^2 \right| IjFm_F \right\rangle = \frac{3}{4}X(X-1) - I(I+1)j(j+1), \tag{2.224}$$

which allows us to write

$$\mathcal{H}_Q = \mathbf{Q}^{(2)}\mathbf{U}^{(2)} = \frac{e^2q\mathcal{Q}}{\hbar^4 2I(2I-1)j(2j-1)}\left[3(\mathbf{IJ})^2 + \frac{3\hbar^2}{2}(\mathbf{IJ}) - I^2J^2\right]. \tag{2.225}$$

The first–order energy shift due to this electric quadrupole interaction is then

$$\langle IjFm_F \,|\, \mathcal{H}_Q \,|\, IjFm_F\rangle = \frac{e^2q\mathcal{Q}}{2I(2I-1)j(2j-1)}\left[\frac{3}{4}X(X-1) - I(I+1)j(j+1)\right]. \tag{2.226}$$

Since $V_{zz} = -e(3z^2 - r^2)/r^5$, the expectation value $\langle V_{zz}\rangle$ vanishes when the electron charge distribution is spherically symmetric, therefore there is no quadrupole energy shift for s states. We recall that nuclei having no spin ($I = 0$) or a spin $I = 1/2$ have no electric quadrupole moment, so the energy shift vanishes also in those cases.

We have seen that the magnetic dipole correction (2.177) is

$$E = \frac{Z^3 \mu_{\mathrm{B}} g_{\mathrm{N}} \mu_{\mathrm{N}}}{a_0^3 n^3 j(j+1)(\ell+\frac{1}{2})}\Big[F(F+1) - I(I+1) - j(j+1)\Big] = -\frac{Z^3 \mu_{\mathrm{B}} g_{\mathrm{N}} \mu_{\mathrm{N}} X}{a_0^3 n^3 j(j+1)(\ell+\frac{1}{2})}, \quad (2.227)$$

therefore the total (magnetic dipole plus electric quadrupole) hyperfine structure energy correction is given by

$$\Delta E = \frac{e^2 q \mathcal{Q}}{2I(2I-1)j(2j-1)}\left[\frac{3}{4}X(X-1) - I(I+1)j(j+1)\right] - \frac{Z^3 \mu_{\mathrm{B}} g_{\mathrm{N}} \mu_{\mathrm{N}} X}{a_0^3 n^3 j(j+1)(\ell+\frac{1}{2})}. \quad (2.228)$$

The hyperfine energy levels obtained after this correction are still independent of the quantum number m_F, and therefore are $(2F+1)$–fold degenerate. This degeneracy can be removed by applying an external magnetic field.

The electric quadrupole interaction $\mathcal{H}_Q$ can also be written using only nuclear spin operators. This is particularly interesting if the matrix elements of the quadrupolar hyperfine hamiltonian are to be computed in the uncoupled basis. This can be of use in cases when, for instance, an interaction with an external field is to be simultaneously treated with the quadrupolar hyperfine interaction. We quote the result here without derivation

$$\mathcal{H}_Q = \frac{e^2 q \mathcal{Q}}{\hbar^2 4I(2I-1)}\Big[(3I_0^2 - I^2) + \eta(I_{+1}^2 + I_{-1}^2)\Big], \quad (2.229)$$

where the parameter $0 \leq \eta \leq 1$ is defined through the relation

$$\langle V_{xx} - V_{yy}\rangle = \eta \langle V_{zz}\rangle = \eta eq, \qquad 0 \leq \eta \leq 1, \quad (2.230)$$

and V_{xx}, V_{yy}, V_{zz}, are the diagonal elements of the electric field gradient tensor[23] produced by the electron at a point whose coordinates with respect to that electron are (x, y, z). For s electrons the charge distributions has spherical symmetry and $V_{xx} = V_{yy} = V_{zz}$. But since V must satisfy the Laplace equation[24], $\nabla^2 V = 0$, the principal components are all zero and in particular $eq = \langle V_{zz}\rangle = 0$. From this, it follows that there is no quadrupole interaction with s electrons (as stated before).

Let's examine a few other particular cases. Assuming axial symmetry, $V_{xx} = V_{yy}$, we have $\eta = 0$, so

$$\mathcal{H}_Q = \frac{e^2 q \mathcal{Q}}{\hbar^2 4I(2I-1)}\Big[(3I_0^2 - I^2)\Big], \quad (2.231)$$

$$E_Q = \langle Im_I \,|\, \mathcal{H}_Q \,|\, Im_I\rangle = \frac{e^2 q \mathcal{Q}}{4I(2I-1)}\Big[3m_I^2 - I(I+1)\Big]. \quad (2.232)$$

In this situation there are only diagonal elements with twofold degeneracies in m_I, which means that states with $\pm m_I$ values have the same energy. For $I = 0, \frac{1}{2}$ the quadrupole energy vanishes. For $I = 1$ the energies are,

$$E_Q = \frac{1}{4}e^2 q \mathcal{Q}(3m_I^2 - 2) = \begin{cases} e^2 q\mathcal{Q}/4, & m_I = \pm 1, \\ -e^2 q\mathcal{Q}/2, & m_I = 0. \end{cases} \quad (2.233)$$

For $I = 3/2$,

$$E_Q = \frac{1}{12}e^2 q \mathcal{Q}\left(3m_I^2 - \frac{15}{4}\right) = \begin{cases} e^2 q\mathcal{Q}/4, & m_I = \pm 3/2, \\ -e^2 q\mathcal{Q}/4, & m_I = \pm 1/2. \end{cases} \quad (2.234)$$

[23] This is a symmetric tensor, therefore diagonalizable.

[24] In the most general case the Poisson equation is the one that holds ($\nabla^2 V = -\rho/\epsilon_0$ in SI units and $\nabla^2 V = -4\pi\rho$ in Gaussian units), but if we are interested in finding the potential in a region where there is no charge (the charge in this case is elsewhere, is in the electron position) we need to use the Laplace equation, $\nabla^2 V = 0$.

2.3.4 Isotope shifts.

The so-called *isotope shifts* in atomic energy levels are caused by two effects: the *mass effect* due to the fact that the nuclear mass is finite, and the *volume effect* arising from the distribution of the electric charge of the nucleus within a finite volume. For one-electron atoms the mass effect can be taken into account by introducing the reduced mass,

$$\mu = \frac{mM}{m+M}, \tag{2.235}$$

together with a new unit of length and atomic energy,

$$a_\mu = \frac{m}{\mu} a_0, \qquad E_\mu = \frac{\mu}{m} E_\infty, \tag{2.236}$$

where E_∞ stands for the energy in the case of a nucleus of *infinite* mass, and $a_0 = \hbar/(mc\alpha)$ is the Bohr radius (sometimes denoted also as a_∞). It can be seen that the reduced-mass effect modifies all the energy levels of the atom in the same manner, leading to all the frequencies of the spectral lines being reduced by a factor $\mu/m \approx 1 - m/M$. If we call ΔE the reduced-mass correction to be applied to the energy levels E_∞, we obtain

$$\Delta E = E_\mu - E_\infty = E_\infty \left(\frac{\mu}{m} - 1 \right) \approx -\frac{m}{M} E_\infty. \tag{2.237}$$

Notice that ΔE is positive. For atoms with more than one electron the mass effect causes an additional energy shift called the *mass polarization correction*, which we skip for the moment.

To make an estimation of the *volume effect* let us assume a simple model of the nucleus where the electric charge is distributed in a uniform way within a sphere of radius $R = r_0 A^{1/3}$. Here A stands for the nuclear mass number and r_0 is a constant that can be taken to be approximately $r_0 \approx 1.2 \times 10^{-15}\,\mathrm{m}$. The electrostatic potential created by such a uniform sphere can be shown to be

$$\phi(r) = \frac{Ze}{2R}\left(3 - \frac{r^2}{R^2}\right) \quad \text{for} \quad r \le R, \qquad \phi(r) = \frac{Ze}{r} \quad \text{for} \quad r > R. \tag{2.238}$$

For simplicity, we shall assume that the unperturbed Hamiltonian is just the hydrogenic Hamiltonian,

$$\mathcal{H}_0 = \frac{p^2}{2\mu} - \frac{Ze^2}{r}, \tag{2.239}$$

and that the perturbation $\mathcal{H}'$ is the difference between the interaction induced by (2.238) and the Coulomb interaction $-Ze^2/r$ (we ignore here all other effects such as relativistic corrections etc.). The Hamiltonian of the perturbation is then

$$\mathcal{H}'(r) = -\frac{Ze^2}{2R}\left(3 - \frac{r^2}{R^2}\right) + \frac{Ze^2}{r} = \frac{Ze^2}{2R}\left(\frac{r^2}{R^2} + \frac{2R}{r} - 3\right) \quad \text{for} \quad r \le R.$$

$$\mathcal{H}'(r) = 0 \quad \text{for} \quad r > R.$$

Using now $\psi_{n\ell m} = R_{n\ell}(r) Y_\ell^m(\theta, \varphi)$ and the fact that the spherical harmonics are normalized, we can write the first-order energy shift due to this perturbation as

$$\Delta E = \langle \psi_{n\ell m} \,|\, \mathcal{H}' \,|\, \psi_{n\ell m} \rangle = \frac{Ze^2}{2R} \int_0^R \left(\frac{r^2}{R^2} + \frac{2R}{r} - 3\right) |R_{n\ell}(r)|^2 \, r^2 \, \mathrm{d}r. \tag{2.240}$$

The region $r \le R$ is very small compared to atomic dimensions, so we take $R_{n\ell}(r)$ to be the constant $R_{n\ell}(0)$ and pull it out of the integral. Furthermore, $R_{n\ell}(0)$ vanishes unless $\ell = 0$, therefore we are left with

$$\Delta E = \frac{Ze^2}{2R} |R_{n0}(0)|^2 \int_0^R \left(\frac{r^2}{R^2} + \frac{2R}{r} - 3\right) r^2 \, \mathrm{d}r = \frac{Ze^2 R^2}{10} |R_{n0}(0)|^2. \tag{2.241}$$

We have $\psi_{n00}(0) = R_{n0}(0)Y_0^0(\theta,\varphi) = R_{n0}(0)/\sqrt{4\pi}$, and also the relations

$$|\psi_{n00}(0)|^2 = \frac{1}{4\pi}|R_{n0}(0)|^2 = \frac{Z^3}{\pi a_\mu^3 n^3}, \qquad |R_{n0}(0)|^2 = \frac{4Z^3}{a_\mu^3 n^3}, \tag{2.242}$$

which allows to write finally, for $\ell = 0$ (we have $\Delta E \approx 0$ for $\ell \neq 0$)

$$\Delta E \approx \frac{2e^2 Z^4 R^2}{5a_\mu^3 n^3}. \tag{2.243}$$

In the SI system we have

$$\Delta E = \left(\frac{1}{4\pi\epsilon_0}\right)\frac{2e^2 Z^4 R^2}{5a_\mu^3 n^3} = \left(\frac{e^2}{4\pi\epsilon_0 \hbar c}\right)\frac{2\hbar c Z^4 R^2}{5a_\mu^3} = \frac{2\alpha\hbar c Z^4 R^2}{5a_\mu^3}. \tag{2.244}$$

What is measured in the laboratory is the difference, δE, of energy shifts between two isotopes whose charge distributions have radii R and $R + \delta R$, respectively. To first order in δR we have then

$$\delta E = \Delta E(R + \delta R) - \Delta E(R) = \frac{2e^2 Z^4}{5a_\mu^3 n^3}\left[(R + \delta R)^2 - R^2\right] \approx \frac{4e^2 Z^4}{5a_\mu^3 n^3} R\delta R. \tag{2.245}$$

Note that the isotope with the larger radius has the higher energy value (less tightly bound). This is confirmed by experiment. Another feature to observe is that δE increases rapidly with Z and decreases with n, so that the most important volume effects occur for low–lying s states of hydrogenic atoms with large Z, in particular for the ground state.

2.4 Problems.

1. Compute the matrix elements which appear in table (2.1). Compute the energy eigenvalues using the secular equations (2.37) to (2.40) and find the approximations for the strong-field and weak-field cases.

2. Work out the *weak* Zeeman effect assuming that it is possible to compute the expectation value of $\mathbf{L} + 2\mathbf{S} = \mathbf{J} + \mathbf{S}$ by noting that $\mathbf{L}$ and $\mathbf{S}$ precess about the *constant* $\mathbf{J}$ vector, so that the time-average value of $\mathbf{S}$ is just its projection along $\mathbf{J}$.

3. Consider the eight $n = 2$ Hydrogen states, $|2\ell j m_j\rangle$ and find the energy for each of them under the weak-field Zeeman splitting assumption. Make a diagram with the energy in the vertical axis and the intensity of the external magnetic field applied in the horizontal axis. Draw in the diagram the corresponding eight straight lines in the weak-field regime specifying the eight slopes.

4. For each of the eight $n = 2$ Hydrogen states, $|2\ell m_\ell m_s\rangle$, find its energy under the strong-field Zeeman splitting. Express each answer as the sum of three terms: the Bohr (zero order) energy, the fine–structure (proportional to α^2), and the Zeeman contribution (proportional to $\mu_{\rm B} B_{\rm ext}$). If you ignore the fine structure altogether, how many distinct levels are there and what are their degeneracies?

5.
   - Determine the total number of possible states with principal quantum number $n = 4$ for the electron in the hydrogen atom. Give the values of the orbital angular momentum ℓ and the total angular momentum j for each of the states. In other words, write them in the basis $|\ell j m_j\rangle$. Write them also in the basis $|\ell m_\ell m_s\rangle$.
   - All these states are degenerate if the fine structure is not taken into account. Compute the energy shift induced by the fine structure for all the states. Write the result in Rydberg units and give also the ratio of the energy shift to the *unperturbed* energy $E_4^{(0)}$.
   - If the hydrogen atom is surrounded by a weak magnetic field B pointing along the z axis, a further energy shift due to the Zeeman effect takes place. Compute this energy shift and draw a diagram (not to scale) showing this splitting on top of the fine structure splitting for the energy levels with $n = 4$.
   - Make a rough estimation of how large the external magnetic field B would have to be to produce an energy shift comparable to that of the fine structure. Give your result in Tesla and Gauss.

6. An electron in a hydrogen atom is under the influence of the perturbation

$$V = \alpha(x^2 + y^2) + \beta z^2,$$

with $0 < \alpha < -\beta \ll e^2/a_0^3$. Determine the energy correction to the five hydrogen eigenstates with $n < 3$ and $\ell < 2$, without taking into account the spin of the electron. Calculate the Zeeman effect for $\mathbf{B}$ parallel to the z axis and $\mathbf{B}$

parallel to the x axis, in the regime of strong magnetic field assuming that $\mu_B B \ll V$.

Hint: you may find useful the following relations:

$$\langle r^2 \rangle = \int R_{n\ell}^2(r) r^4 dr = \frac{n^2}{2}(5n^2 + 1 - 3\ell(\ell+1))a_0^2$$

$$\begin{aligned}(\cos^2\theta)\, Y_\ell^m(\theta,\phi) = &\frac{1}{2\ell+3}\left[\frac{((\ell+1)^2-m^2)((\ell+2)^2-m^2)}{(2\ell+1)(2\ell+5)}\right]^{1/2} Y_{\ell+2}^m(\theta,\phi) \\ &+ \left[\frac{(\ell+1)^2-m^2}{(2\ell+1)(2\ell+3)} + \frac{\ell^2-m^2}{4\ell^2-1}\right] Y_\ell^m(\theta,\phi) \\ &+ \frac{1}{2\ell+1}\left[\frac{((\ell+1)^2-m^2)(\ell^2-m^2)}{(2\ell-1)(2\ell-3)}\right]^{1/2} Y_{\ell-2}^m(\theta,\phi).\end{aligned}$$

7. For the linear Stark effect in Hydrogen, compute the matrix elements given in equations (2.75) and (2.76). Show that the eigenstates corresponding to the eigenvalues $\pm 3e\mathcal{E}a_0/Z$ are $\frac{1}{\sqrt{2}}(\psi_{200} \mp \psi_{210})$.

8. Consider the linear Stark effect in the $n = 3$ levels of the Hydrogen atom.

   a) Show that the energy splitting is symmetric around the unperturbed energy.

   b) Specify the degeneration of each level after the perturbation is considered.

   c) Show that the levels are equally spaced in energy.

9. A Hydrogen atom in its ground state is exposed to a uniform electric field $\mathcal{E}$ in the x-axis direction. The Hamiltonian of the system can be taken to be $\mathcal{H} = \mathcal{H}_0 + e\mathcal{E}x$, where $\mathcal{H}_0 = p^2/(2m) - e^2/r$ is the unperturbed Hydrogen Hamiltonian. To study the ground level of the perturbed system use the trial function $\psi = N(1+\beta x)\phi$, where β is a real variational parameter and $\phi = e^{-r/a_0}/\sqrt{\pi a_0^3}$ is the Hydrogen ground state wave function.

   (a) Taking into account that $\langle \phi \,|\, r^2 \,|\, \phi \rangle = 3a_0^2$ prove that the normalization constant is given by $|N|^2 = 1/(1+\beta^2 a_0^2)$. Prove also the following relations

   $$[\mathcal{H}, x] = -\frac{i\hbar}{m}p_x, \qquad \frac{\partial\phi}{\partial x} = \frac{x}{r}\frac{\mathrm{d}\phi}{\mathrm{d}r}.$$

   (b) Using the above relations and the fact that $\langle \phi \,|\, r \,|\, \phi \rangle = 3a_0/2$, you should be able to find the following expression for the energy of the perturbed system as a function of the variational parameter

   $$E(\beta) = \frac{1}{1+\beta^2 a_0^2}\left[E_0 + \beta^2\left(E_0 a_0^2 + \frac{\hbar^2}{2m}\right) + 2\beta e\mathcal{E}a_0^2\right].$$

   (c) If the electric field is not very strong we could expand $E(\beta)$ up to second-order in the small parameter β. In this approximation, prove that the correction to the ground-state energy is

   $$(\Delta E)^{(\mathrm{va})} = E - E_0 = \frac{(e\mathcal{E}a_0)^2}{E_0}.$$

You may find useful the relation $1/(1+x^2) \approx 1 - x^2$ for small x.

(d) The energy correction computed with the variational method, $(\Delta E)^{(\mathrm{va})}$, is quadratic in the electric field. Explain why this was to be expected and compare this result with the value which is obtained using second-order time-independent perturbation theory. Make the ratio of both values and give a numerical result.

10. In this exercise we will compute the energy splitting produced by a weak external electric field $V(\mathbf{r}) = e\mathcal{E}z$ (weak Stark effect) that is much smaller than the fine structure in the Hydrogen atom. In such a case we need to use the coupled basis $|n, j, m\rangle$. Remember that the first-order Stark effect arises only in degenerate states, therefore states with the same n and j quantum numbers. In this scenario, and since $[J_z, z] = 0$, only matrix elements between states with the same m are different from zero.

a) To compute the matrix elements we will need to express the states of the coupled base as a linear combination of the states of the uncoupled base, $|n, \ell, m_\ell, m_s\rangle$,

$$|n, j, m\rangle = \sum_{m_\ell, m_s} \langle n, \ell, m_\ell, m_s | n, j, m\rangle \; |n, \ell, m_\ell, m_s\rangle \, .$$

For systems such as the Hydrogen atom, with $s = 1/2$, states couple in doublets (except for $\ell = 0$) and it is possible to derive the following general expressions that relate the $|n, j, m\rangle$ basis to states with arbitrary ℓ and m in the uncoupled representation:

$$|\psi_1\rangle = \left| n, j = \ell + \frac{1}{2}, m_j \right\rangle = \frac{\sqrt{\ell + m + \frac{1}{2}}}{\sqrt{2\ell + 1}} \left| n, \ell, m_\ell = m_j - \frac{1}{2}, m_s = \frac{1}{2} \right\rangle + \frac{\sqrt{\ell - m + \frac{1}{2}}}{\sqrt{2\ell + 1}} \left| n, \ell, m_\ell = m_j + \frac{1}{2}, m_s = -\frac{1}{2} \right\rangle ,$$

$$|\psi_2\rangle = \left| n, j = \ell - \frac{1}{2}, m_j \right\rangle = \frac{\sqrt{\ell - m + \frac{1}{2}}}{\sqrt{2\ell + 1}} \left| n, \ell, m_\ell = m_j - \frac{1}{2}, m_s = \frac{1}{2} \right\rangle - \frac{\sqrt{\ell + m + \frac{1}{2}}}{\sqrt{2\ell + 1}} \left| n, \ell, m_\ell = m_j + \frac{1}{2}, m_s = -\frac{1}{2} \right\rangle .$$

Show that $J^2 |\psi_i\rangle = \hbar^2 j_i (j_i + 1) |\psi_i\rangle$, which demonstrates that $|\psi_1\rangle$ and $|\psi_2\rangle$ are indeed J^2 eigenstates.

In the space-spin representation these states can be written as (for $\ell \neq 0$)

$$|\psi_1\rangle = \frac{R_{n,\ell}(r)}{\sqrt{2\ell + 1}} \begin{pmatrix} \sqrt{\ell + m + \frac{1}{2}} \, Y_\ell^{m - \frac{1}{2}} \\ \sqrt{\ell - m + \frac{1}{2}} \, Y_\ell^{m + \frac{1}{2}} \end{pmatrix} ,$$

$$|\psi_2\rangle = \frac{R_{n,\ell}(r)}{\sqrt{2\ell + 1}} \begin{pmatrix} \sqrt{\ell - m + \frac{1}{2}} \, Y_\ell^{m - \frac{1}{2}} \\ -\sqrt{\ell + m + \frac{1}{2}} \, Y_\ell^{m + \frac{1}{2}} \end{pmatrix} .$$

Hydrogen levels with the same values of n and j are degenerate (have the same energy), therefore we need to identify, using the previous expression, states with the same values of n and j, but different values of ℓ. These are the doublets corresponding to $\ell = j \pm \frac{1}{2}$

$$|\phi_1\rangle = \frac{R_{n,j-1/2}(r)}{\sqrt{2j}} \begin{pmatrix} \sqrt{j+m}\, Y^{m-\frac{1}{2}}_{j-\frac{1}{2}} \\ \sqrt{j-m}\, Y^{m+\frac{1}{2}}_{j-\frac{1}{2}} \end{pmatrix}, \qquad \ell = j - \frac{1}{2}, \quad \ell \neq 0,$$

$$|\phi_2\rangle = \frac{R_{n,j+1/2}(r)}{\sqrt{2(j+1)}} \begin{pmatrix} \sqrt{j-m+1}\, Y^{m-\frac{1}{2}}_{j+\frac{1}{2}} \\ -\sqrt{j+m+1}\, Y^{m+\frac{1}{2}}_{j+\frac{1}{2}} \end{pmatrix}, \qquad \ell = j + \frac{1}{2}, \quad \ell \neq 0.$$

b) Use the previous expressions for $|\phi_1\rangle$ and $|\phi_2\rangle$ to show that the energy shifts produced by $V(\mathbf{r})$ are

$$\Delta E = \pm \frac{3}{4} e\mathcal{E} a_0 \frac{mn\sqrt{n^2 - (j+1/2)^2}}{j(j+1)}.$$

You might need the two following relations:

$$\int_0^\infty r^3 R_{n,j-1/2}(r) R_{n,j+1/2}(r)\, \mathrm{d}r = -\frac{3}{2} n \sqrt{n^2 - (j+1/2)^2}\, a_0,$$

$$\cos\theta\, Y^m_\ell = \sqrt{\frac{(\ell+m+1)(\ell-m+1)}{(2\ell+1)(2\ell+3)}}\, Y^m_{\ell+1} + \sqrt{\frac{(\ell+m)(\ell-m)}{(2\ell-1)(2\ell+1)}}\, Y^m_{\ell-1}.$$

c) What is the magnitude of the electric field (E) that would produce a Stark effect shift of 10% the fine-structure shift in a state with $n = 2$, $j = 1/2$ and $m = -1/2$?

11. We saw in the previous problem that the degenerate hydrogen levels with the same values of n and j (but different values of ℓ) are the doublets:

$$|\phi_1\rangle = \frac{R_{n,j-1/2}(r)}{\sqrt{2j}} \begin{pmatrix} \sqrt{j+m}\, Y^{m-1/2}_{j-1/2} \\ \sqrt{j-m}\, Y^{m+1/2}_{j-1/2} \end{pmatrix}, \qquad \ell = j - \frac{1}{2}, \quad \ell \neq 0,$$

$$|\phi_2\rangle = \frac{R_{n,j+1/2}(r)}{\sqrt{2(j+1)}} \begin{pmatrix} \sqrt{j-m+1}\, Y^{m-1/2}_{j+1/2} \\ -\sqrt{j+m+1}\, Y^{m+1/2}_{j+1/2} \end{pmatrix}, \qquad \ell = j + \frac{1}{2}, \quad \ell \neq 0.$$

Consider now a perturbation that is smaller than the fine structure of hydrogen and is given by $V = \lambda z^2$. Demonstrate that the energy shift of the hydrogen spectrum due to this perturbation is:

$$\Delta E_{n\ell jm} = \lambda \frac{a_0^2}{Z^2} \frac{n^2}{4} \left[5n^2 + 1 - 3\ell(\ell+1)\right] \left(1 - \frac{m^2}{(j+1)j}\right).$$

Hint 1: It is useful to write z^2 as

$$z^2 = r^2 \cos^2\theta = \frac{2\sqrt{\pi}}{3} r^2 \left(\frac{2}{\sqrt{5}} Y^0_2 + Y^0_0\right).$$

Hint 2: Use a conservation law to justify that $\langle \phi_2 | V | \phi_1 \rangle = 0$.

Hint 3: Consider

$$\langle Y_{\ell'}^{m'}|Y_2^0|Y_\ell^m\rangle = (-1)^m\sqrt{\frac{5}{4\pi}}\sqrt{(2\ell'+1)(2\ell+1)}\begin{pmatrix} \ell' & 2 & \ell \\ -m' & 0 & m\end{pmatrix}\begin{pmatrix} \ell' & 2 & \ell \\ 0 & 0 & 0\end{pmatrix}.$$

Hint 4: Also consider

$$\begin{pmatrix} \ell & 2 & \ell \\ -m & 0 & m\end{pmatrix} = (-1)^{\ell-m}\frac{2[3m^2-\ell(\ell+1)]}{\sqrt{(2\ell+3)(2\ell+2)(2\ell+1)(2\ell)(2\ell-1)}}.$$

12. a) Given a 2×2 hermitian operator

$$\mathcal{H} = \begin{pmatrix} \mathcal{H}_{11} & \mathcal{H}_{12} \\ \mathcal{H}_{12}^* & \mathcal{H}_{22}\end{pmatrix},$$

prove that the solutions of the secular equation are

$$\lambda_\pm = \frac{1}{2}\left[\mathcal{H}_{11}+\mathcal{H}_{22} \pm \sqrt{(\mathcal{H}_{11}-\mathcal{H}_{22})^2+4|\mathcal{H}_{12}|^2}\right].$$

In LS coupling, the degenerate levels of a spin 1/2 system can be described using the basis $|\alpha, L, S, M_L = M_J - \frac{1}{2}, M_S = \frac{1}{2}\rangle$ and $|\alpha, L, S, M_L = M_J + \frac{1}{2}, M_S = -\frac{1}{2}\rangle$, where α stands for other quantum numbers that might be needed to completely specify the states. Consider an atom under the influence of an external electric field that produces a perturbation (quadratic Stark effect) comparable in magnitude to the spin-orbit interaction. Ignoring constant terms that do not contribute to the energy splitting, the combined perturbation can be described using the Hamiltonian

$$\mathcal{H} = \frac{a}{\hbar^2}\mathbf{L}\cdot\mathbf{S} - \frac{b}{\hbar^2}L_z^2,$$

where a and b are constant parameters (in fact b is proportional to the modulus squared of the electric field).

b) Prove that, if $|M_J| < L + \frac{1}{2}$ (which is equivalent to require $|M_J| \leq L - \frac{1}{2}$), the energy of the levels is shifted by the amount

$$\Delta E = -\frac{(b+a)}{4} - bM_J^2 \pm \sqrt{\frac{a^2}{4}\left(L+\frac{1}{2}\right)^2 + b(b+a)M_J^2}.$$

Hint: compute the $\mathcal{H}$ matrix elements and use the result of part a).

c) For $M_J = L + \frac{1}{2}$ the state $|\alpha, L, S, M_J + \frac{1}{2}, -\frac{1}{2}\rangle$ does not exist, because M_L is not allowed to have the value $M_L = M_J + \frac{1}{2} = L + 1$. Similarly, for $M_J = -(L + \frac{1}{2})$ the state $|\alpha, L, S, M_J - \frac{1}{2}, \frac{1}{2}\rangle$ does not exist. In any of these two cases we have a single level, not a doublet. Prove that, for $M_J = \pm(L + \frac{1}{2})$, the energy shift is

$$\Delta E = \frac{a}{2}L - bL^2.$$

13. Deuterium is a hydrogen isotope with a nucleus of spin 1 consisting of a proton and a neutron. The interaction between the nucleus and electron spins leads to a hyperfine splitting of the deuterium ground state. Transitions between these hyperfine levels lead to emission electromagnetic radiation which can be detected on earth to map

the concentration of atomic deuterium in our galaxy. Determine the wavelength of that radiation. (Data: dipole magnetic moment of the deuterium nucleus $\mu_d = g_{\rm d}\mu_{\rm N} = 0.857\,\mu_{\rm N}$; Planck constant $h = 6.626 \times 10^{-34}\,{\rm J}\cdot{\rm s}$; permeability of free space $\mu_0 = 4\pi \times 10^{-7}\,{\rm H\,m^{-1}}$; nuclear magneton $\mu_{\rm N} = 5.05 \times 10^{-27}\,{\rm J\,T^{-1}}$; Bohr magneton $\mu_{\rm B} = 9.27 \times 10^{-24}\,{\rm J\,T^{-1}}$).

14. The second-order correction to the energy in time-independent perturbation theory is

$$E_n^{(2)} = \sum_{m\neq n} \frac{|\langle n\,|\,V\,|\,m\rangle|^2}{E_n^{(0)} - E_m^{(0)}} = \sum_{m\neq n} \frac{|V_{nm}|^2}{E_n^{(0)} - E_m^{(0)}}. \tag{2.246}$$

Show that this can also be written as

$$E_n^{(2)} = \frac{(V^2)_{nn}}{E_n^{(0)}} - \frac{(V_{nn})^2}{E_n^{(0)}} + \sum_{m\neq n} \frac{E_m^{(0)}\,|V_{nm}|^2}{E_n^{(0)}\left(E_n^{(0)} - E_m^{(0)}\right)}. \tag{2.247}$$

By neglecting the sum of the right-hand side of the previous expression obtain an approximation for the quadratic Stark effect on the Hydrogen ground state.

15. For an electric field strength of $\mathcal{E} = 10^7\ {\rm V\,m^{-1}}$, compute the linear and quadratic Stark effect in Hydrogen. Give the answers in eV and ${\rm cm^{-1}}$.

16. When an atom is placed in a constant magnetic field, $\mathbf{B}$, the Hamiltonian must be modified by adding a perturbation given by

$$\mathcal{H}_{\rm m} = \frac{1}{\hbar}\mu_{\rm B}\,\mathbf{B}(\mathbf{L} + 2\mathbf{S}) + \frac{e^2}{2mc^2}A^2,$$

where $\mathbf{A} = \frac{1}{2}\mathbf{B}\times\mathbf{r}$.

a) Assuming a very strong magnetic field (but less than de 50 Tesla), study the effect of the quadratic term (diamagnetic perturbation) on the $2p$ states of the Hydrogen atom. Compare the result (make a ratio) with the largest perturbation due to the linear term (Paschen-Back effect).

b) Compute the corresponding energy shifts for a 50 Tesla magnetic field.

c) Determine the intensity of the magnetic field needed for the diamagnetic perturbation to be 10% of the Paschen-Back effect. Is perturbation theory applicable to such an intense magnetic field?

d) Suppose now that the only perturbation to the Coulomb interaction in the Hydrogen atom was given by the Hamiltonian

$$\mathcal{H}_1 = \frac{(eA)^2}{2mc^2}(1 + 2\cos(2\varphi)),$$

where φ is the azimuthal angle in spherical coordinates. Neglecting the spin degree of freedom, study the effect of this perturbation on the states $|21\bar{1}\rangle$, $|210\rangle$, $|211\rangle$ (the notation $|n\ell m_\ell\rangle = |21\bar{1}\rangle$ means $n = 2$, $\ell = 1$, $m_\ell = -1$).
Hint: you may use the integrals over r and θ computed in part a).

17. Compute the matrix elements of the magnetic hyperfine interaction in the hydrogen s state in the presence of an external magnetic field and check that you obtain the results shown in table (2.3). Compute the corresponding eigenvalues.

18. Assuming axial symmetry for the diagonal elements of the electric field gradient tensor, $V_{xx} = V_{yy} \neq V_{zz}$, compute the quadrupole splitting induced by a nucleus of spin $I = 2$.

19. The most abundant Aluminium isotope, ^{27}Al, has a nuclear spin of $I = 5/2$, a quadrupole electric moment of $\mathcal{Q} = 0.147$ barns, and a nuclear quadrupole coupling constant of $e^2q\mathcal{Q} = 2 \times 10^{-7}$ eV. Assuming axial symmetry for the diagonal elements of the electric field gradient tensor, $V_{xx} = V_{yy} \neq V_{zz}$, compute the quadrupole splitting induced by the nuclear spin on the electronic energy levels, giving the results in eV and cm^{-1}. Compute also the value of the quantity q in the ^{27}Al nuclear quadrupole coupling constant, giving the result in units of fm^{-3}.

20. For an atom of Xenon, $^{132}_{54}$Xe, compute the energy of the $1s$ state using the hydrogenic atom wavefunction for a point–like nucleus. Next, assume that the nuclear electric charge is uniformly distributed within a sphere of radius $R = r_0 A^{1/3}$, where $r_0 = 1.2 \times 10^{-15}$ m, and compute the energy shift due to the finite size of the nucleus. What fraction of the *point–like* hydrogenic energy does this shift represent? Repeat the calculations for the case of a muon in the $1s$ state of the Xenon atom and comment on the results ($m_\mu = 105.66\,\mathrm{MeV}/c^2$). For such a heavy nucleus you can ignore the effect of the reduced mass.

21. If the proton could be treated as a uniformly charged spherical shell of radius b, the potential energy would be the constant $-e^2/b$ inside the shell (for $r \leq b$) and $-e^2/r$ outside the shell (for $r > b$).

    (a) Use this very simple model and first-order perturbation theory to prove that the correction to the ground state of the hydrogen atom due to the finite size of the proton is $\Delta E = 2e^2b^2/(3a_\mu^3)$. (Hint: take $R(r) \approx R(0)$ to simplify the integrals, where $R(r)$ is the radial part of the ground state).

    (b) Evaluate ΔE assuming that the proton has a radius of about $b = 0.88$ fm. Express the result both in eV and as a ratio with respect to the unperturbed energy. Is the size of the correction comparable to the fine structure?

    (c) Repeat the calculation for the muonic hydrogen (a muon bound to a proton) and answer the same question as before.

    Data: $m_e = 0.511$ MeV; $m_p = 938.27$ MeV; $m_\mu = 105.66$ MeV; hydrogen ground state wave function: $\psi = \left(\frac{1}{a_\mu}\right)^{3/2} \frac{1}{\sqrt{\pi}} e^{-r/a_\mu}$; Bohr radius $a_0 = 0.529 \times 10^{-10}$ m.

22. A particle of mass m and charge $q = -e$ is placed in a 3-dimensional box of sides

(a, a, b), where $b \ll a$. A weak electric field

$$\boldsymbol{\mathcal{E}} = \mathcal{E}\,(y/a, x/a, 0), \qquad \mathcal{E} > 0,$$

is applied to this particle. Find the energy of the ground state and first excited states to order $\mathcal{E}$ (first order in perturbation theory).

Chapter 3

Atoms with several electrons and molecules.

3.1 The helium atom and the hydrogen molecule ion.

Before going to more complicated systems let us have a look to the simplest multi-electron atom. Even in this case no exact solution of the Schrödinger equation has been found. Ignoring spin-orbit coupling, the Hamiltonian for helium ($Z = 2$) consists of just two hydrogenic Hamiltonians and an additional term describing the repulsion between the two electrons,

$$\mathcal{H} = \left(\frac{p_1^2}{2m} - \frac{2e^2}{r_1}\right) + \left(\frac{p_2^2}{2m} - \frac{2e^2}{r_2}\right) + \frac{e^2}{r_{12}}. \tag{3.1}$$

If we could ignore the e^2/r_{12} term, the solution to the Schrödinger equation would be the product of two hydrogen-like wave functions,

$$\psi(\mathbf{r_1}, \mathbf{r_2}) = \psi_{n\ell m}(\mathbf{r}_1)\psi_{n'\ell'm'}(\mathbf{r}_2). \tag{3.2}$$

Since $Z = 2$, the Bohr radius for helium would have to be divided by two with respect to hydrogen,

$$(a_0)_{\rm H} = \frac{\hbar^2}{me^2} \quad \longrightarrow \quad (a_0)_{\rm He} = \frac{\hbar^2}{2me^2}, \tag{3.3}$$

and the energy of each electron multiplied by 4,

$$E_n = -\frac{1}{2}\alpha^2 mc^2 \frac{Z^2}{n^2} = -\frac{1}{2}\alpha^2 mc^2 \frac{4}{n^2} = 4(E_n)_{\rm H} \tag{3.4}$$

Since we have two electrons, the total energy would be

$$E_{n,n'} = -4\left[\frac{1}{2}\alpha^2 mc^2 \left(\frac{1}{n^2} + \frac{1}{(n')^2}\right)\right]. \tag{3.5}$$

In particular, the ground state energy for this *simplified* helium atom would be

$$E_0 = -8 \times 13.6\,\text{eV} = -109\,\text{eV}, \tag{3.6}$$

and its wavefunction

$$\begin{aligned}\psi_0(\mathbf{r_1}, \mathbf{r_2}) &= \psi_{100}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2) \\ &= \left(\frac{Z}{a_0}\right)^{3/2} \frac{1}{\sqrt{\pi}} e^{-Zr_1/a_0} \left(\frac{Z}{a_0}\right)^{3/2} \frac{1}{\sqrt{\pi}} e^{-Zr_2/a_0} = \frac{8}{\pi a_0^3} e^{-2(r_1+r_2)/a_0}.\end{aligned} \tag{3.7}$$

Since this is a symmetric function under the interchange of the two electrons, the corresponding spin state has to be antisymmetric to make a total overall antisymmetric state, as required for a system of two identical fermions,

$$\text{Parahelium} \qquad |0,0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle). \tag{3.8}$$

The actual ground state of helium is indeed a spin singlet (parahelium), but its measured energy value is -78.975 eV, quite a bit apart from the estimation we got by ignoring the electronic repulsion. The excited states of helium consist of one electron in the hydrogenic ground state and the other in an excited state, $\psi_{n\ell m}\psi_{100}$. Both electrons in excited states is a very unstable configuration: one electron would decay immediately to the ground state releasing enough energy to knock out the other, leading to ionized helium.

Although the ground state is necessarily parahelium due to Pauli's exclusion principle,[1] the excited helium states may be found with the electrons in the singlet configuration or in the triplet (symmetric) configuration. This latter is called *orthohelium*.

$$\text{Orthoelium} \qquad |1,1\rangle = |\uparrow\uparrow\rangle\,, \qquad |1,0\rangle = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle), \qquad |1,-1\rangle = |\downarrow\downarrow\rangle\,. \tag{3.9}$$

Since the symmetric spatial state brings the electrons closer together, we expect a higher interaction energy in parahelium, leading to higher energy levels (atom less tightly bound) than their orthoehelium counterparts. This can be seen in figure (3.1). The values of the energy levels shown in that figure are relative to the ground state of ionized helium (with just one electron, so the binding energy is $-4 \times 13.6\,\text{eV} = -54.4\,\text{eV}$). To get the total energy we have to subtract 54.4 eV; for example, from the figure we see that the energy of the ground state would be $-24.5 - 54.4 = -78.9\,\text{eV}$.

In chapter 4 it will be shown that radiative transitions between singlet and triplet spin states (known as intercombination lines), are not allowed in the electric dipole approximation provided the spin-orbit interactions can be neglected. Therefore, for low Z ($Z < 40$) where spin–orbit effects are negligible, the energy spectrum of two–electron atoms (or ions) consists of two nearly independent systems of levels, one for the *para* (singlet) states and the other for the *ortho* (triplet) states. This is precisely what happens in the helium spectrum. The presence of two independent types of spectra for helium led to the first spectroscopists to talk about two different species of helium, parahelium and orthohelium. This terminology is still used now, even if there is only one helium and the origin of the two spectra is understood.

3.1.1 The ground state of helium through the variational principle.

As stated before, the ground state energy of helium has been measured to be $E_{\text{gs}} = -78.975$ eV. We don't quote any uncertainty here because in this section we are going to make only an estimation of this energy applying the variational principle,[2] not a detailed calculation to be compared with experiment. We have seen already that if we ignore the electronic repulsion term in the Hamiltonian we get an approximate solution for the ground state which is given as a product of two hydrogenic wave functions, $\psi_0 = (8/\pi a^3)e^{-2(r_1+r_2)/a}$, with an energy of about $8E_1 = -109$ eV (we write here a instead of a_0 for the Hydrogen Bohr radius, for brevity). Recall that $E_n = -13.6/n^2$ eV for Hydrogen and that for a nucleus with atomic number Z we should do the transformation $E_n \to Z^2 E_n$.

Before going to the variational principle method let us compute the effect of the electronic repulsion using the ψ_0 wave function of the previous section. Ignoring fine structure and smaller

[1] The spatial wave function for orthohelium states must be antisymmetric but, for the ground state, an antisymmetric spatial wave function is identically zero (see equation 3.67).

[2] There is no known exact solution to this problem.

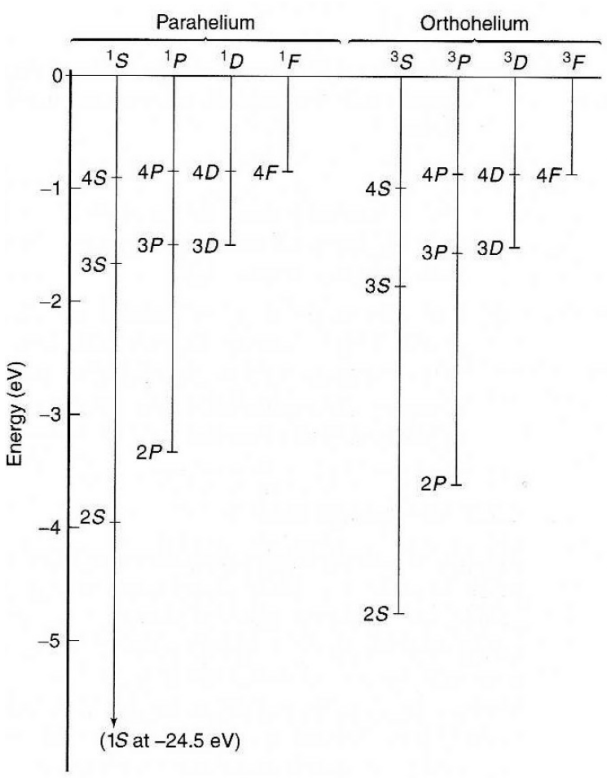


Figure 3.1: Energy level diagram for helium. The parahelium energies are uniformly higher than their orthohelium counterparts. The numerical values on the vertical scale are shifted by 54.4 eV (to get the true total energy subtract 54.4 eV from each level). Taken from Griffiths, 2018.

corrections the Helium Hamiltonian is

$$\mathcal{H} = -\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) - e^2\left(\frac{2}{r_1} + \frac{2}{r_2} - \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\right). \tag{3.10}$$

Calling $V_{\text{ee}} = e^2/|\mathbf{r_1} - \mathbf{r_2}|$, we may write

$$\mathcal{H} = \left[-\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) - e^2\left(\frac{2}{r_1} + \frac{2}{r_2}\right)\right] + V_{\text{ee}}, \tag{3.11}$$

therefore

$$\mathcal{H}\psi_0 = 8E_1\psi_0 + V_{\text{ee}}\psi_0, \qquad \langle\mathcal{H}\rangle = 8E_1 + \langle V_{\text{ee}}\rangle\,, \tag{3.12}$$

where, using atomic units (see problem (1)),

$$\langle V_{\text{ee}}\rangle = \langle\psi_0\,|\,V_{\text{ee}}\,|\,\psi_0\rangle = \left(\frac{8}{\pi a^3}\right)^2 e^2 \int \frac{e^{-4(r_1+r_2)/a}}{|\mathbf{r_1} - \mathbf{r_2}|}\, d^3\mathbf{r}_1 d^3\mathbf{r}_2 \;=\; \frac{5e^2}{4a} = -\frac{5}{2}E_1 = 34\,\text{eV}. \tag{3.13}$$

Here we have used the relation $a = -e^2/2E_1$, which is easily seen to be consistent with the more familiar expression $E_1 = -\alpha^2 mc^2/2$,

$$E_1 = -\frac{1}{2}e^2\frac{1}{a} = -\frac{1}{2}e^2\frac{\alpha mc}{\hbar} = -\frac{1}{2}\frac{e^2}{\hbar c}\alpha mc^2 = -\frac{1}{2}\alpha^2 mc^2. \tag{3.14}$$

So we get

$$\langle\mathcal{H}\rangle = -109\,\text{eV} + 34\,\text{eV} = -75\,\text{eV}, \tag{3.15}$$

a value much closer to the measured one.

We employ now the variational method to find a better approximation to the ground state by assuming that the effect of one of the helium electrons over the other is to partially screen the nuclear charge. In other words, each electron represents a cloud of negative charge which partially shields the nucleus, so that the other electron actually sees an effective nuclear charge ζ that is a bit less than $Z = 2$. Let's try then the following functional form,

$$\psi_1(\mathbf{r}_1, \mathbf{r}_2) = \frac{\zeta^3}{\pi a^3} e^{-\zeta(r_1+r_2)/a}, \tag{3.16}$$

where ζ is a variational parameter. We must find the value of ζ that minimizes $\langle \mathcal{H} \rangle$, where $\mathcal{H}$ is the full Hamiltonian (3.10) without any approximation.[3] We rewrite this Hamiltonian in the following manner,

$$\mathcal{H} = \underbrace{-\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2) - e^2\left(\frac{\zeta}{r_1} + \frac{\zeta}{r_2}\right)}_{\psi_1 \text{ is an eigenfunction of this part}} + e^2\left(\frac{(\zeta-2)}{r_1} + \frac{(\zeta-2)}{r_2} + \frac{1}{|\mathbf{r_1} - \mathbf{r_2}|}\right). \tag{3.17}$$

Note that ψ_1 is the ground state eigenfunction of the partial Hamiltonian indicated by the under-brace symbol in the previous equation. The expectation value of that partial Hamiltonian computed for ψ_1 is then $2\zeta^2 E_1$ (compare with $2Z^2 E_1$). And the expectation value of the full $\mathcal{H}$ is

$$\langle \mathcal{H} \rangle = 2\zeta^2 E_1 + 2(\zeta - 2)e^2 \left\langle \frac{1}{r} \right\rangle + \langle V_{\text{ee}} \rangle, \tag{3.18}$$

where $\langle 1/r \rangle$ is the expectation value of $1/r$ in the one-particle hydrogenic ground state ψ_{100} with nuclear charge ζ. This can be found by taking equation (1.133) and replacing Z/a_0 by $\zeta/a_0 = \zeta/a$,

$$\left\langle \frac{1}{r} \right\rangle = \frac{Z}{a} \quad \longrightarrow \quad \left\langle \frac{1}{r} \right\rangle = \frac{\zeta}{a} = -\frac{2\zeta}{e^2} E_1. \tag{3.19}$$

For $\langle V_{\text{ee}} \rangle$ we may take the result we got before for $\zeta = 2$ (equation (3.13), $5e^2/4a$) and multiply the parameter a by $2/\zeta$ to make the equation valid for any value of ζ,

$$\langle V_{\text{ee}} \rangle = \frac{5e^2}{4\frac{2}{\zeta}a} = \frac{5\zeta e^2}{8a} = -\frac{5\zeta}{4} E_1. \tag{3.20}$$

Putting everything together we find

$$\langle \mathcal{H} \rangle = \left[2\zeta^2 - 4\zeta(\zeta - 2) - \frac{5\zeta}{4}\right] E_1 = \left[-2\zeta^2 + \frac{27}{4}\zeta\right] E_1. \tag{3.21}$$

According to the variational principle, the value of this expression exceeds the actual ground–state energy of helium, E_{gs}, for any value of ζ. To find the lowest bound we minimize $\langle \mathcal{H} \rangle$ with respect to ζ,

$$\frac{d\langle \mathcal{H} \rangle}{d\zeta} = \left[-4\zeta + \frac{27}{4}\right] E_1 = 0 \quad \Longrightarrow \quad \zeta = \frac{27}{16} = \frac{1}{2}\left(\frac{3}{2}\right)^3 = 1.69. \tag{3.22}$$

The value of ζ is slightly less than 2, in agreement with the screening hypothesis. Inserting this value of ζ in (3.21) we find

$$\langle \mathcal{H} \rangle = \frac{1}{2}\left(\frac{3}{2}\right)^6 E_1 = -77.5\,\text{eV}. \tag{3.23}$$

Indeed, we are a little closer to the measured value of -78.975 eV. The Helium calculation was first done in 1927, and historically was considered to have provided one of the earliest signs that Schrödinger's wave mechanics was on the right track. Using more complicated trial functions with more adjustable parameters it is possible to calculate the helium ground state energy with good precision.

[3] Remember that in the variational method the Hamiltonian is not modified. For a brief review you may read appendix B.4.

3.1.2 The hydrogen molecule ion, H_2^+.

As another application of the variational principle we discuss now the hydrogen molecule ion, H_2^+, consisting of a single electron in the Coulomb field of two protons. Let us first assume the protons are fixed at a distance R apart from each other. The Hamiltonian is then

$$\mathcal{H} = -\frac{\hbar^2}{2m}\nabla^2 - e^2\left(\frac{1}{r_1}+\frac{1}{r_2}\right), \tag{3.24}$$

where r_1 and r_2 are the distances of the electron from the respective protons. In fact, the Schrödinger equation for this simple Hamiltonian (with R taken constant) can be solved analytically using spheroidal coordinates. However, this approach is not applicable to more complicated cases and we go directly to apply the variational method. For a trial ground state function, we take a linear combination of the two hydrogenic states which would be associated to the electron being bound to one proton or the other.[4]

$$\psi = A[\psi_0(\mathbf{r}_1)+\psi_0(\mathbf{r}_2)], \qquad \psi_0(r) = \frac{1}{\sqrt{\pi a^3}}e^{-r/a}. \tag{3.25}$$

We now normalize this trial function (the $\psi_0(\mathbf{r}_i)$ wavefunctions are already normalized),

$$\begin{aligned} 1 = \int |\psi|^2\, d^3\mathbf{r} = |A|^2\left[\int |\psi_0(\mathbf{r}_1)|^2\, d^3\mathbf{r} + \int |\psi_0(\mathbf{r}_2)|^2\, d^3\mathbf{r} + 2\int \psi_0(\mathbf{r}_1)\psi_0(\mathbf{r}_2)\, d^3\mathbf{r}\right] \\ = 2|A|^2\left[1+\int \psi_0(\mathbf{r}_1)\psi_0(\mathbf{r}_2)\, d^3\mathbf{r}\right] = 2|A|^2\Big(1+I\Big), \end{aligned} \tag{3.26}$$

where (see problem 4)

$$I \equiv \langle\psi_0(\mathbf{r}_1)|\psi_0(\mathbf{r}_2)\rangle = \frac{1}{\pi a^3}\int e^{-(r_1+r_2)/a}\, d^3\mathbf{r} = e^{-R/a}\left[1+\left(\frac{R}{a}\right)+\frac{1}{3}\left(\frac{R}{a}\right)^2\right]. \tag{3.27}$$

The parameter I, called the overlap integral, measures the amount by which $\psi_0(\mathbf{r}_1)$ overlaps $\psi_0(\mathbf{r}_2)$. Note that I goes to zero as $R\to\infty$, and goes to 1 as $R\to 0$. The normalization constant in terms of I is

$$|A|^2 = \frac{1}{2(I+1)}. \tag{3.28}$$

Now we should compute the expectation value of $\mathcal{H}$ in the trial state $\psi = A[\psi_0(\mathbf{r}_1)+\psi_0(\mathbf{r}_2)]$. Since $\psi_0(\mathbf{r}_1)$ and $\psi_0(\mathbf{r}_2)$ are hydrogen ground eigenstates corresponding to the electron being bound to proton 1 or proton 2, respectively, we have

$$\left(-\frac{\hbar^2}{2m}\nabla^2-\frac{e^2}{r_1}\right)\psi_0(\mathbf{r}_1) = E_1\psi_0(\mathbf{r}_1), \qquad \left(-\frac{\hbar^2}{2m}\nabla^2-\frac{e^2}{r_2}\right)\psi_0(\mathbf{r}_2) = E_1\psi_0(\mathbf{r}_2), \tag{3.29}$$

where $E_1 = -13.6$ eV. Therefore we can write

$$\begin{aligned} \mathcal{H}\psi &= A\left[-\frac{\hbar^2}{2m}\nabla^2 - e^2\left(\frac{1}{r_1}+\frac{1}{r_2}\right)\right]\Big[\psi_0(\mathbf{r}_1)+\psi_0(\mathbf{r}_2)\Big] \\ &= A\left(-\frac{\hbar^2}{2m}\nabla^2-\frac{e^2}{r_1}\right)\psi_0(\mathbf{r}_1) + A\left(-\frac{\hbar^2}{2m}\nabla^2-\frac{e^2}{r_2}\right)\psi_0(\mathbf{r}_2) - Ae^2\left[\frac{1}{r_2}\psi_0(\mathbf{r}_1)+\frac{1}{r_1}\psi_0(\mathbf{r}_2)\right] \\ &= E_1\psi - Ae^2\left[\frac{1}{r_2}\psi_0(\mathbf{r}_1)+\frac{1}{r_1}\psi_0(\mathbf{r}_2)\right]. \end{aligned} \tag{3.30}$$

[4] This strategy of taking a *linear combination of atomic orbitals* (LCAO) is a standard procedure that we will study later in this chapter from a more general point of view.

Before computing $\langle\psi|\mathcal{H}|\psi\rangle$ note that, due to the symmetry of the problem, we have

$$\left\langle A\psi_0(\mathbf{r}_1)+A\psi_0(\mathbf{r}_2)\left|\frac{Ae^2}{r_2}\psi_0(\mathbf{r}_1)+\frac{Ae^2}{r_1}\psi_0(\mathbf{r}_2)\right.\right\rangle =$$
$$|A|^2e^2\left\langle\psi_0(\mathbf{r}_1)\left|\frac{1}{r_2}\right|\psi_0(\mathbf{r}_1)\right\rangle+|A|^2e^2\left\langle\psi_0(\mathbf{r}_2)\left|\frac{1}{r_1}\right|\psi_0(\mathbf{r}_2)\right\rangle$$
$$+|A|^2e^2\left\langle\psi_0(\mathbf{r}_1)\left|\frac{1}{r_1}\right|\psi_0(\mathbf{r}_2)\right\rangle+|A|^2e^2\left\langle\psi_0(\mathbf{r}_2)\left|\frac{1}{r_2}\right|\psi_0(\mathbf{r}_1)\right\rangle =$$
$$2|A|^2e^2\left\langle\psi_0(\mathbf{r}_1)\left|\frac{1}{r_2}\right|\psi_0(\mathbf{r}_1)\right\rangle+2|A|^2e^2\left\langle\psi_0(\mathbf{r}_1)\left|\frac{1}{r_1}\right|\psi_0(\mathbf{r}_2)\right\rangle. \quad (3.31)$$

Therefore

$$\langle\mathcal{H}\rangle = E_1 - 2e^2|A|^2\left[\left\langle\psi_0(\mathbf{r}_1)\left|\frac{1}{r_2}\right|\psi_0(\mathbf{r}_1)\right\rangle+\left\langle\psi_0(\mathbf{r}_1)\left|\frac{1}{r_1}\right|\psi_0(\mathbf{r}_2)\right\rangle\right]. \quad (3.32)$$

The so-called direct integral is given by (problem (5))

$$D \equiv a\left\langle\psi_0(\mathbf{r}_1)\left|\frac{1}{r_2}\right|\psi_0(\mathbf{r}_1)\right\rangle = \frac{a}{R}-\left(1+\frac{a}{R}\right)e^{-2R/a}, \quad (3.33)$$

and the exchange integral,

$$X \equiv a\left\langle\psi_0(\mathbf{r}_1)\left|\frac{1}{r_1}\right|\psi_0(\mathbf{r}_2)\right\rangle = \left(1+\frac{R}{a}\right)e^{-R/a}. \quad (3.34)$$

Note that both D and X go to zero when $R\to\infty$. When performing the direct and exchange integrals it might help you to use a more explicit (but cumbersome) notation, to make clear that there is only one $\mathbf{r}$ to integrate over. If we write the electron wavefunction bound to proton 1 as $\psi_0^{(\mathrm{p}_1)}(\mathbf{r})$ instead of $\psi_0(\mathbf{r}_1)$ and, similarly, $\psi_0^{(\mathrm{p}_2)}(\mathbf{r})$ instead of $\psi_0(\mathbf{r}_2)$, the exchange integral can be written as

$$X = a\left\langle\psi_0(\mathbf{r}_1)\left|\frac{1}{r_1}\right|\psi_0(\mathbf{r}_2)\right\rangle = a\left\langle\psi_0^{(\mathrm{p}_1)}(\mathbf{r})\left|\frac{1}{r}\right|\psi_0^{(\mathrm{p}_2)}(\mathbf{r})\right\rangle.$$

To proceed, we could take a reference frame such that proton 1 is at the origin and proton 2 is on the z–axis at the point $z=R$. In this frame, the wavefunction $\psi_0^{(\mathrm{p}_2)}$ can be written as

$$\psi_0^{(\mathrm{p}_2)}(\mathbf{r}) = \frac{1}{\sqrt{\pi a^3}}\,e^{-\frac{1}{a}\sqrt{r^2+R^2-2rR\cos\theta}},$$

where $\mathbf{r}=(r,\theta,\varphi)$ is the electron position vector referred to proton 1 (the origin).

Collecting everything we have (since $1/a=-2E_1/e^2$)[5]

$$\langle\mathcal{H}\rangle = E_1 - 2e^2|A|^2\left(\frac{D}{a}+\frac{X}{a}\right) = E_1+4E_1|A|^2(D+X) = \left[1+2\left(\frac{D+X}{1+I}\right)\right]E_1. \quad (3.35)$$

According to the variational principle, the actual ground state energy has to be *less* than the value for $\langle\mathcal{H}\rangle$ that we have just calculated. Furthermore, we should not forget that there is also potential energy associated with the proton-proton repulsion, which is given by (remember that $E_1<0$)

$$V_{\mathrm{pp}} = \frac{e^2}{R} = -\frac{2a}{R}E_1. \quad (3.36)$$

[5]We are using $E_1 = -\left(\frac{e^2}{4\pi\epsilon_0}\right)\frac{1}{2a} = -13.6\,\mathrm{eV}$, or, since in our expressions we are taking $\frac{1}{4\pi\epsilon_0}=1$, we use $E_1=-\frac{e^2}{2a}$.

Therefore the total energy of the system would be

$$E_{\rm tot} = -\frac{2a}{R}E_1 + \left[1 + 2\left(\frac{D+X}{1+I}\right)\right]E_1. \tag{3.37}$$

For $R \to \infty$ we recover $E_{\rm tot} \to E_1$, which is just the ground-state hydrogen binding energy. This is expected because in that situation we have two infinitely separated protons with an electron bound to one of them. In units of $-E_1$ the total energy is

$$\frac{E_{\rm tot}}{-E_1} = \frac{2a}{R} - \left[1 + 2\left(\frac{D+X}{1+I}\right)\right]. \tag{3.38}$$

Defining a new variable, $x = R/a$, we have

$$D = \frac{1}{x} - \left(1 + \frac{1}{x}\right)e^{-2x}, \qquad X = (1+x)\,e^{-x}, \qquad I = \left(1 + x + \frac{1}{3}x^2\right)e^{-x}. \tag{3.39}$$

Therefore the total energy (in units of $-E_1$) can be written as

$$\frac{E_{\rm tot}}{-E_1} = F(x) = \frac{2}{x} - 1 - 2\left(\frac{D+X}{1+I}\right) = -1 + \left(\frac{2}{x}\right)\frac{\left(1 - \frac{2}{3}x^2\right)e^{-x} + (1+x)e^{-2x}}{1 + \left(1 + x + \frac{1}{3}x^2\right)e^{-x}}. \tag{3.40}$$

It is very instructive to plot this function and see that there is a region where the graph goes below -1, with a minimum around 2.4 Bohr radii for the proton–proton separation, as shown in figure (3.2). This means that the binding energy of the system is less than that of a neutral hydrogen atom plus a free proton ($-13.6\,\text{eV}$) and the bonding of the H_2^+ molecule ion is possible in that region. It is the prototype of a covalent bond, with the electron shared equally by the two protons. The equilibrium separation of the two protons is about 2.4 Bohr radii, or 1.3 Å (the experimental value is 1.06 Å). The variational principle only gives an upper bound for the ground state energy (overestimates the energy), which means that underestimates the strength of the bond. More accurate estimates of the energy and proton separation can be obtained using more ellaborate trial functions. The experimental value for the energy required to completely dissociate the molecule according to $H_2^+ \to H^+ + H^+ + e^-$ is $E_{\rm eq} = -16.39\,\text{eV}$.

3.2 N-electron atoms.

To study atoms with an arbitrary number of electrons we need to extend the one–particle Schrödinger equation by summing over N electrons and adding all the relevant electron–electron interactions. We will assume that it is not necessary to include operators involving more than two electrons. The most important terms in the Hamiltonian of a N-electron atom in the absence of external fields are then

$$\mathcal{H} = \underbrace{\sum_{i=1}^{N}\left(\frac{p_i^2}{2m} - \frac{Ze^2}{r_i}\right)}_{\mathcal{H}_0} + \underbrace{\sum_{i<j}\frac{e^2}{r_{ij}}}_{\mathcal{H}_1} + \underbrace{\sum_{i=1}^{N}\xi(r_i)\,\boldsymbol{\ell}_i\mathbf{s}_i}_{\mathcal{H}_2} = \mathcal{H}_0 + \mathcal{H}_1 + \mathcal{H}_2, \tag{3.41}$$

where r_i is the distance from the nucleus of charge Ze to the ith electron with mass m, charge $-e$ and momentum $\mathbf{p}_i$. The distance between the ith and the jth electron is $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. The term $\mathcal{H}_0$ contains the kinetic and potential energy of every electron in the field of the nucleus. The term $\mathcal{H}_1$ is a sum over all electron pairs and contains the Coulomb repulsion between every pair of electrons. The term $\mathcal{H}_2$ contains the spin–orbit interaction. Other effects such as those due to hyperfine interactions or relativistic corrections will be treated as perturbations.

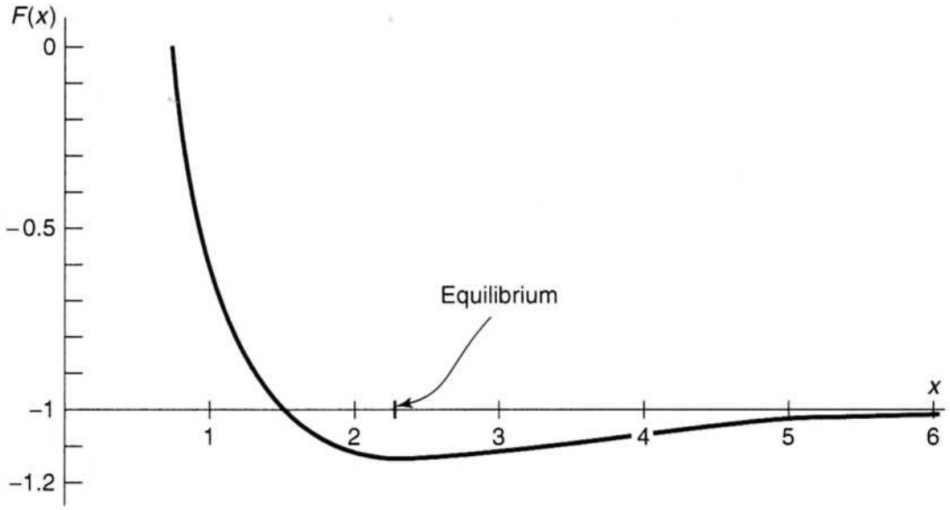


Figure 3.2: Plot of the function $F(x)$ given by equation (3.40). This shows the existence of a bound state at an equilibrium x value of about 2.4. The variable x is the distance between the protons in units of the Bohr radius. Figure taken from Griffiths [7].

The relative importance of the $\mathcal{H}_1$ and $\mathcal{H}_2$ terms depends on the atomic number Z. For low Z, $\mathcal{H}_1$ is the dominant term and $\mathcal{H}_2$ can be treated as a perturbation. For high Z it is the other way around, since the spin–orbit interaction goes with Z^4 (equation 1.177). The basic Hamiltonians for these two limiting cases are then

$$\mathcal{H}_{\rm I} \;=\; \mathcal{H}_0 + \mathcal{H}_1 = \sum_{i=1}^{N}\left(\frac{p_i^2}{2m} - \frac{Ze^2}{r_i}\right) + \sum_{i<j}\frac{e^2}{r_{ij}} \qquad \text{(for low } Z\text{)}, \tag{3.42}$$

$$\mathcal{H}_{\rm II} = \mathcal{H}_0 + \mathcal{H}_2 \;=\; \sum_{i=1}^{N}\left(\frac{p_i^2}{2m} - \frac{Ze^2}{r_i}\right) + \sum_{i=1}^{N}\xi(r_i)\,\boldsymbol{\ell}_i\mathbf{s}_i \qquad \text{(for high } Z\text{)}. \tag{3.43}$$

The properties of most atoms can be described on the basis of one of these Hamiltonians, especially $\mathcal{H}_{\rm I}$, which is often regarded as the fundamental Hamiltonian in atomic physics.

3.2.1 Symmetries of the Hamiltonian. LS coupling and jj coupling.

For a system of N identical particles, the Hamiltonian must be invariant under an interchange in the coordinates (spatial and spin) of any two particles. This forces the eigenstates of the Hamiltonian to be symmetric or antisymmetric with respect to an interchange of two particles. Fermions obey the Fermi-Dirac statistics, therefore the total wavefunction for a system of N electrons must be antisymmetric with respect to an interchange of any two of them. The Hamiltonian given by equation (3.41) is also invariant under a parity transformation (spatial reflection with respect to the origin, $\mathbf{r} \to -\mathbf{r}$). This implies that the eigenstates have a definite parity, either positive or negative. In other words, parity is a good quantum number of atomic states.

Due to the inter-electronic repulsion, the Hamiltonian $\mathcal{H}_{\rm I}$

$$\mathcal{H}_{\rm I} \;=\; \sum_{i=1}^{N}\left(\frac{p_i^2}{2m} - \frac{Ze^2}{r_i}\right) + \sum_{i<j}\frac{e^2}{r_{ij}} \;=\; \mathcal{H}_0 \;+\; \mathcal{H}_1, \tag{3.44}$$

does not commute with the orbital angular momentum of the individual electrons, $\boldsymbol{\ell}_i$, but we can prove (problem 7) that it does commute with the total orbital angular momentum of the

entire electronic system,[6]

$$\mathbf{L} = \sum_{i=1}^{N} \boldsymbol{\ell}_i, \qquad [\mathcal{H}_{\rm I}, \mathbf{L}] = 0. \tag{3.45}$$

The absence of spin operators in $\mathcal{H}_{\rm I}$ seems to suggest that $[\mathcal{H}_{\rm I}, \mathbf{S}] = 0$, where $\mathbf{S} = \sum_i \mathbf{s}_i$ and $\mathbf{s}_i$ is the spin operator for the ith electron. This is correct, but the rigorous proof (which we do not provide here) cannot rely on such a simple argument because the antisymmetry requirement on the wave function prevents the spin and orbital angular momentum from being completely independent. The Hamiltonian $\mathcal{H}_{\rm I}$ commutes with both $\mathbf{L}$ and $\mathbf{S}$, it must then commute also with $\mathbf{J} = \mathbf{L} + \mathbf{S}$,

$$[\mathcal{H}_{\rm I}, \mathbf{L}] = 0, \quad [\mathcal{H}_{\rm I}, \mathbf{S}] = 0, \quad [\mathcal{H}_{\rm I}, \mathbf{J}] = 0, \quad \mathbf{L} = \sum_i \boldsymbol{\ell}_i, \quad \mathbf{S} = \sum_i \mathbf{s}_i, \quad \mathbf{J} = \mathbf{L} + \mathbf{S}. \tag{3.46}$$

Now we inspect $\mathcal{H}_{\rm II}$,

$$\mathcal{H}_{\rm II} \;=\; \sum_{i=1}^{N} \left(\frac{p_i^2}{2m} - \frac{Ze^2}{r_i} \right) \;+\; \sum_{i=1}^{N} \xi(r_i)\, \boldsymbol{\ell}_i \mathbf{s}_i \;=\; \mathcal{H}_0 \;+\; \mathcal{H}_2. \tag{3.47}$$

When studying the spin-orbit coupling we learned that $\boldsymbol{\ell}\mathbf{s}$ does not commute with $\boldsymbol{\ell}$ nor $\mathbf{s}$, but it does commute with $\mathbf{j} = \boldsymbol{\ell} + \mathbf{s}$. Therefore

$$[\mathcal{H}_{\rm II}, \mathbf{L}] \neq 0, \quad [\mathcal{H}_{\rm II}, \mathbf{S}] \neq 0, \quad [\mathcal{H}_{\rm II}, \mathbf{J}] = 0, \quad \mathbf{J} = \sum_i \mathbf{j}_i, \quad \mathbf{j}_i = \boldsymbol{\ell}_i + \mathbf{s}_i. \tag{3.48}$$

We have then two coupling schemes. That of equation (3.46) is called ***LS coupling*** or *Russell–Saunders coupling*, and it is used when the Hamiltonian $\mathcal{H}_{\rm I}$ is the appropriate one. That is, for atoms where the inter–electron Coulomb repulsion energy is far greater than the spin-orbit interaction. In this situation, the orbital and spin angular momentum are conserved separately. The coupling described in equation (3.48) is named ***jj coupling***, to be used with the Hamiltonian $\mathcal{H}_{\rm II}$ when the spin-orbit interaction exceeds the electron repulsion energy.

In the LS–coupling we can take the $\mathcal{H}_{\rm I}$ eigenstates to be also eigenstates of L^2, L_z, S^2, S_z, and denote them as $|LSM_LM_S\rangle$, or $|\alpha LSM_LM_S\rangle$ where α is an index to represent any additional information that might be needed to completely specify the state (the radial part, parity etc.). We can also use eigenstates of L^2, S^2, J^2, J_z, denoted as $|LSJM\rangle$ or $|\alpha LSJM\rangle$. Whether we write the eigenfunctions of $\mathcal{H}_{\rm I}$ as $|LSM_LM_S\rangle$, in the uncoupled representation, or as $|LSJM\rangle$, in the coupled representation, is just a matter of choice and convenience. The fact that $\mathcal{H}_{\rm I}$ commutes with $\mathbf{L}$, $\mathbf{S}$ and $\mathbf{J}$ means that $\mathcal{H}_{\rm I}$ is invariant under a rotation of spatial and spin coordinates. A further property of $\mathcal{H}_{\rm I}$ is that the energy eigenvalues are independent of M_L and M_S. This can be seen qualitatively as a consequence of the fact that the energy of a system described by $\mathcal{H}_{\rm I}$ is independent of external coordinates and the relative orientations of orbital and spin angular momenta.

3.2.2 Central field approximation.

We'll focus here on the Hamiltonian $\mathcal{H}_{\rm I}$, which has a wider applicability than $\mathcal{H}_{\rm II}$. It turns out that, although the inter-electron repulsion is too large to be treated as a perturbation, it contains a large spherically–symmetric component that can be used to enormously simplify the problem. Indeed, in the so-called *central field approximation* a spherically–symmetric one-electron operator $U(r_i)$ is built to approximate the potential energy of the ith electron in the

[6] We will adopt the convention of using lowercase letters do designate one–electron quantities and uppercase letters for those belonging to the entire electronic system.

field of the nucleus and the remaining $N-1$ electrons. Let's write the Hamiltonian as

$$\mathcal{H}_{\rm I} = \sum_{i=1}^{N}\left(\frac{p_i^2}{2m} - \frac{Ze^2}{r_i}\right) + \sum_{i<j}\frac{e^2}{r_{ij}} = \underbrace{\sum_{i=1}^{N}\left(\frac{p_i^2}{2m} - \frac{Ze^2}{r_i}\right) + \left\langle \sum_{i<j}^{N}\frac{e^2}{r_{ij}}\right\rangle}_{\mathcal{H}_0'} + \underbrace{\sum_{i<j}^{N}\frac{e^2}{r_{ij}} - \left\langle \sum_{i<j}^{N}\frac{e^2}{r_{ij}}\right\rangle}_{\mathcal{H}_1'}, \quad (3.49)$$

where $\left\langle \sum_{i<j}^{N} e^2/r_{ij}\right\rangle$ is the average over a sphere of the electron repulsion and is therefore independent of the angular coordinates. The term $\mathcal{H}_0'$ can also be written as

$$\mathcal{H}_0' = \sum_{i=1}^{N}\left(\frac{p_i^2}{2m} + U(r_i)\right), \quad \text{with} \quad \sum_{i=1}^{N}U(r_i) = -\sum_{i=1}^{N}\frac{Ze^2}{r_i} + \left\langle \sum_{i<j}^{N}\frac{e^2}{r_{ij}}\right\rangle. \quad (3.50)$$

The $\mathcal{H}_1'$ term contains the non–spherical part of the electronic repulsion, while $\mathcal{H}_0'$ contains the kinetic energy, potential energy in the field of the nucleus, and the spherical–average electron repulsion energy. The main features of the electronic motion are described by the Hamiltonian $\mathcal{H}_0'$, since it contains most of the inter–electron repulsion. The remaining $\mathcal{H}_1'$ part is small enough to be treated as a perturbation. In the central field approximation the atom is replaced by a system which, in the first approximation, consists of *independent* electrons moving in a spherically symmetric field produced by the nucleus and all the other electrons. Higher approximations take into account the departure from spherical symmetry in the effective field seen by the electrons.

We have then for $\mathcal{H}_0'$ the following Schrödinger equation for the N–electron spatial wave function, which we denote as Φ,

$$\mathcal{H}_0'\Phi = \left[\sum_{i=1}^{N}\left(\frac{p_i^2}{2m} + U(r_i)\right)\right]\Phi = E\Phi. \quad (3.51)$$

Note that, since $U(r_i)$ is a one-electron operator, this equation is separable into N equations of the form

$$\left(\frac{p^2}{2m} + U(r)\right)\phi(\mathbf{r}) = \varepsilon\phi(\mathbf{r}). \quad (3.52)$$

From the study of the hydrogen atom we know that this equation can be further separated into two more, one depending on r and the other on the spherical angles θ and φ. The potential $U(r)$ is for the moment unknown,[7] but the solutions of the angular part can be written in terms of the spherical harmonics. The general solution to (3.52) can then be written as

$$\phi_{n\ell m_\ell}(r,\theta,\varphi) = R_{n\ell}(r)Y_\ell^{m_\ell}(\theta,\varphi) = \frac{1}{r}\,u_{n\ell}(r)Y_\ell^{m_\ell}(\theta,\varphi). \quad (3.53)$$

These $\phi_{n\ell m_\ell}$ functions are called *central field orbitals*. Of course, the radial functions $R_{n\ell}(r)$ and $u_{n\ell}(r)$ are not to be confused with the hydrogen-like functions that we have already studied. They are in general complicated functions that may not be expressible in an analytical form.

The complete wave function follows from multiplying the previous expression by a spin function $\chi(m_s)$, exactly as we did with the hydrogen atom. The one-electron central-field spin orbital is then

$$\psi_{n\ell m_\ell m_s}(\mathbf{r},m_s) = \phi_{n\ell m_\ell}(\mathbf{r})\,\chi(m_s) = \frac{1}{r}\,u_{n\ell}(r)Y_\ell^{m_\ell}(\theta,\varphi)\,\chi(m_s). \quad (3.54)$$

[7] Apart from the obvious physical requirement that $U(r)$ must go to zero as $r\to\infty$.

The four quantum numbers needed to characterize such an orbital take on the usual values,

$$n = 1, 2, \ldots \qquad \ell = 0, 1, 2, \ldots n-1, \qquad m_\ell = -\ell, \ldots, +\ell, \qquad m_s = \pm 1/2. \tag{3.55}$$

When reading bibliography on atomic physics, note that a general spin-orbital does not need to have the form (3.54). Any one-electron function that depends on the space and spin coordinates is a spin-orbital, but the central field spin-orbital (3.54) is the most commonly used. To satisfy the Pauli exclusion principle or, in other words, the Fermi-Dirac statistics, the spin orbitals must be organized into an antisymmetric N–electron wave function. The antisymmetric wave function of two electrons is

$$\Psi(\lambda_1, \lambda_2) = \frac{1}{\sqrt{2}}\Big[\psi_1(\lambda_1)\psi_2(\lambda_2) - \psi_2(\lambda_1)\psi_1(\lambda_2)\Big], \tag{3.56}$$

where each $\psi_i(\lambda_j)$ is a spin orbital. The subscript i identifies a particular choice of the four quantum numbers n, ℓ, m_ℓ, m_s, and λ_j stands for the space and spin coordinates of the jth electron. For a three–electron antisymmetric wave function one could write

$$\begin{aligned}\Psi(\lambda_1, \lambda_2, \lambda_3) = \frac{1}{\sqrt{3!}}\Big[&\psi_1(\lambda_1)\psi_2(\lambda_2)\psi_3(\lambda_3) + \psi_3(\lambda_1)\psi_1(\lambda_2)\psi_2(\lambda_3) + \psi_2(\lambda_1)\psi_3(\lambda_2)\psi_1(\lambda_3) \\ &- \psi_2(\lambda_1)\psi_1(\lambda_2)\psi_3(\lambda_3) - \psi_1(\lambda_1)\psi_3(\lambda_2)\psi_2(\lambda_3) - \psi_3(\lambda_1)\psi_2(\lambda_2)\psi_1(\lambda_3)\Big].\end{aligned} \tag{3.57}$$

And for N electrons,

$$\Psi(\lambda_1, \lambda_2, \ldots, \lambda_N) = \frac{1}{\sqrt{N!}} \sum_P (-1)^{\pi_P} \psi_{p_1}(\lambda_1)\psi_{p_2}(\lambda_2)\ldots\psi_{p_N}(\lambda_N), \tag{3.58}$$

where the sum runs over the set of all permutations P of N elements and π_P is the parity of the permutation. This can also be written using a so-called Slater determinant,

$$\Psi(\lambda_1, \lambda_2, \ldots, \lambda_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\lambda_1) & \psi_2(\lambda_1) & \ldots & \psi_N(\lambda_1) \\ \psi_1(\lambda_2) & \psi_2(\lambda_2) & \ldots & \psi_N(\lambda_2) \\ \ldots & \ldots & \ldots & \ldots \\ \psi_1(\lambda_N) & \psi_2(\lambda_N) & \ldots & \psi_N(\lambda_N) \end{vmatrix}. \tag{3.59}$$

The normalization factor $1/\sqrt{N!}$ arises from the fact that there are $N!$ permutations of the electron coordinates. The parity of a spin orbital $\psi_{n\ell m_\ell m_s}$ is given by the parity of the spherical harmonic $Y_\ell^{m_\ell}$, which is $(-1)^\ell$, therefore the parity of such a product of functions is $(-1)^{\ell_1}(-1)^{\ell_2}\ldots(-1)^{\ell_N}$. Each Slater determinant has then a definite parity. Note that, since we are using the spin orbitals in the uncoupled representation, $\psi_{n\ell m_\ell m_s}$, a Slater determinant is an eigenfunction of the operators L_z and S_z, but not necessarily of $\mathbf{L}^2$ and $\mathbf{S}^2$. To obtain an eigenfunction of $\mathbf{L}^2$, $\mathbf{S}^2$, L_z and S_z we need in general to use linear combinations of Slater determinants.

The Slater determinant incorporates in a natural way the Pauli exclusion principle, because it vanishes if we have two rows or two columns equal, which is to say if two electrons have the same values of the four quantum numbers n, ℓ, m_ℓ, m_s. This, of course, imposes a lot of restrictions on the distribution of electrons in atomic orbitals. In the ground state of an atom the electrons occupy those orbitals allowed by the Pauli principle which result in the lowest energy. The distribution of electrons with respect to n and ℓ is known as a ***configuration***. Electrons with the same value of n and ℓ are said to be ***equivalent***. The maximum number of equivalent electrons for a particular value of ℓ is $2(2\ell + 1)$, given by the product of the orbital and spin degeneracies. Such an assembly of equivalent electrons is known as a ***closed shell***. Electrons outside a closed shell in insufficient numbers to form another closed shell are known as ***valence*** electrons. The chemical properties of an atom are almost entirely determined by

its valence electrons. The most important factors are the number of occupied and unoccupied electron states in the last shell, and the energy interval between this and the next higher (empty) shell. For example, an atom tends to be chemically inert if its highest shell is full and there is an appreciable energy gap to the next higher shell, because in this case the electrons are not easily shared with other atoms to form a molecule. In x-ray spectroscopy it is also common to refer to electrons with principal quantum numbers $n = 1, 2, 3, \ldots$ as residing in the K, L, M, ... shells, respectively, employing a different definition of shell than the one stated before (which, in that context, would be called *subshell*).

The horizontal rows of the periodic table can be qualitatively understood as corresponding to the fillings of each shell according to the Pauli principle. The first shell is completed with helium. Next we have lithium ($Z = 3$), its third electron occupying one of the orbitals with $n = 2$ and $\ell = 0$. If we ignore electronic repulsion all possible values of ℓ would have the same energy, however this is not what happens in the real system. Higher values of angular momentum are associated with larger average distances of the electron to the nucleus. The nucleus charge is *screened* by the inner electrons in such a way that the outer electron sees a lower effective nucleus charge, causing it to be less tightly bound. The larger the distance to the nucleus the larger the screening effect is, therefore lower values of ℓ are first filled in atoms. The third electron of Lithium goes then to the orbital $(2, 0, 0)$. Beryllium ($Z = 4$) makes use of this same state (but with opposite spin). Boron ($Z = 5$) needs to use already the value $\ell = 1$.

We could try to proceed with the filling of electrons in the atom this way, however at a given point the screening effect is going to be so strong that the filling pattern has to be modified. Following argon ($Z = 18$), whose last couple of electrons occupy the state with $n = 3$, $\ell = 1$, the next two atoms (potassium with $Z = 19$ and calcium with $Z = 20$) have their last electron in the state $n = 4$, $\ell = 0$ instead of $n = 3$, $\ell = 2$. In the spectroscopic notation the different values of ℓ are denoted by the letters $s, p, d, f, g, h, i, k, l, \ldots$ (the letter j is not used). The state of a particular electron is represented by the pair $n\ell$, with n (a number) giving the principal quantum number and ℓ (a letter) given the angular orbital momentum. The magnetic/projection quantum number m is not specified, but an exponent is used to indicate the number of electrons that occupy the state in question. For example, the ground-state electronic configuration for carbon ($Z = 6$) is

$$[\mathrm{C}] = (1s)^2(2s)^2(2p)^2, \tag{3.60}$$

which means two electrons in the orbital $(1, 0, 0)$, two in the orbital $(2, 0, 0)$, and two in some combination of the orbitals $(2, 1, 1)$, $(2, 1, 0)$ and $(2, 1, -1)$. In this case we have two electrons out of the last closed shell, with a total angular momentum value which could be $L = 0, 1, 2$. This is specified using the notation[8]

$$^{2\mathrm{S}+1}\mathrm{L}_{\mathrm{J}} \tag{3.61}$$

where S and J (numbers) give the total spin and total angular momentum of the electrons out of the last closed shell.[9] The capitalized letter L specifies the total orbital angular momentum of those electrons. The correspondence between the numerical value of L and the alphabetic letter to denote the term is the following

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
S	P	D	F	G	H	I	K	L	M	N	O	Q	R	T	U	V	W	X	Y	Z

In this notation, the ground state of carbon is $^3\mathrm{P}_0$.

The ground-state configurations of many of the elements can be written down simply from a knowledge of the order in which the energies of the shells increase. This order can be inferred

[8] Remember that we use lowercase letters to designate quantities for a single electron (operators, quantum numbers etc.), and uppercase letters to refer to those quantities for the entire electronic system.

[9] The angular momentum (spin plus orbital) of the electrons belonging to closed shells are combined to produce a zero total angular momentum, therefore the total angular momentum of the electronic system in an atom is completely determined by the electrons in the non-closed shells.

from spectroscopic evidence and is the following one

$$1s, 2s, 2p, 3s, 3p, [4s, 3d], 4p, [5s, 4d], 5p, [6s, 4f, 5d], 6p, [7s, 5f, 6d]. \tag{3.62}$$

It is easy to remember this list by crossing the orbitals of the following table along inverted diagonals ($\swarrow$),

$1s$					
$2s$	$2p$				
$3s$	$3p$	$3d$			
$4s$	$4p$	$4d$	$4f$		
$5s$	$5p$	$5d$	$5f$	$5g$	
$6s$	$6p$	$6d$	$6f$	$6g$	$6h$
$7s$	$\ldots$				

The brackets in (3.62) indicate shells with energies so close that they are not always filled in that order. Shells with different values of n may have similar energies because the increase in n and the decrease in ℓ tend to compensate each other. For example the $4s$ state, which has a higher energy than the $3d$ state in hydrogen, is depressed in N–electron atoms due to the *low screening effect* arising from the penetration caused by its low angular momentum.[10] To be more precise, without higher corrections the energy of a hydrogen-like orbital depends only on the principal quantum number n, and $E_{n+1} > E_n$ always. However, when the Coulomb potential is replaced by the more general central field, $U(r)$, the energy of the orbital $\phi_{n\ell m_\ell}$ depends both on n and ℓ. In general, the energy increases with the sum $n + \ell$. This means that in the central field approximation, characterized by the Hamiltonian $\mathcal{H}_0'$, the energy of an atom is completely determined by the electronic configuration.

When we include the term $\mathcal{H}_1'$,

$$\mathcal{H}_1' = \sum_{i<j}^{N} \frac{e^2}{r_{ij}} - \left\langle \sum_{i<j}^{N} \frac{e^2}{r_{ij}} \right\rangle = \mathcal{H}_1 - \left\langle \sum_{i<j}^{N} \frac{e^2}{r_{ij}} \right\rangle, \tag{3.63}$$

it is possible for electrons in a given configuration to have various energies depending on the mutual orientations of the orbital and spin angular momentum of the individual electrons. We say that the interaction $\mathcal{H}_1'$ produces a splitting in the energy of a configuration. All the splitting is contained in the term $\mathcal{H}_1 = \sum e^2/r_{ij}$, therefore we only need to calculate the matrix elements of that operator alone. The energy of a configuration is then split into ***terms***, with which specific values of L and S are associated. Terms are denoted as $^{2\mathrm{S}+1}\mathrm{L}$. The number $2\mathrm{S}+1$ is called the *multiplicity* of the term. We have a singlet for $\mathrm{S} = 0$, a triplet for $\mathrm{S} = 1$, a quartet for $\mathrm{S} = 2$, and so on.

If, subsequently, the spin–orbit interaction $\mathcal{H}_2 = \sum \xi(r_i)\, \boldsymbol{\ell}_i \mathbf{s}_i$ is applied, there is a further splitting into ***levels*** according to the allowed values of J. Levels are denoted as $^{2\mathrm{S}+1}\mathrm{L}_\mathrm{J}$. The level splitting is also known as the ***fine structure*** and the complete set of levels belonging to a given electronic configuration is a ***multiplet***. The spin-orbit interaction energy increases rapidly with atomic number (goes like Z^4 in the one-electron case) while the inter-electron energy varies much more slowly as a function of Z. Therefore, as we move up the periodic table, the $\mathcal{H}_2$ term eventually becomes the dominant perturbation relative to $\mathcal{H}_1$.

3.2.3 Multiplet wave functions and terms.

The wave functions associated to terms are eigenstates of the operators L^2, S^2, L_z, S_z, and can be written as $|LSM_LM_S\rangle$. The notation $|^{2S+1}LM_LM_S\rangle$ is also used, where $2S+1$ is called

[10] In other words, a $3d$ electron is on average farther away from the nucleus than a $4s$ electron, which sees therefore a smaller nuclear screening. It can be shown that, for all atoms, the radial wave functions of the electronic orbitals behave near the origin like r^ℓ, so that electrons with large ℓ spend little time near the nucleus.

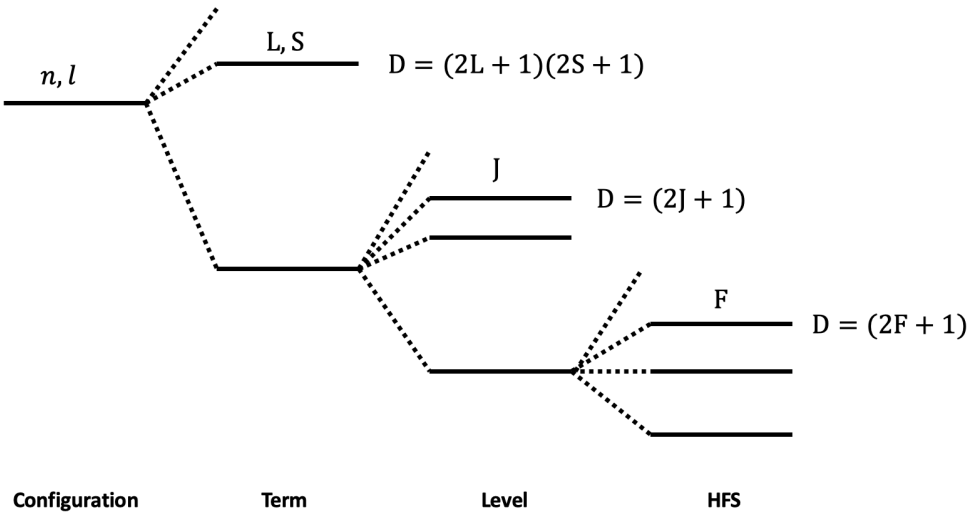


Figure 3.3: Schematic arrangement of the energies of atomic states in LS coupling and their degeneracies, D. The non-spherical part of the Coulomb repulsion splits a configuration into terms. The spin-orbit coupling splits terms into levels. The magnetic hyperfine interaction splits levels into hyperfine structure (HFS). Adapted from Weissbluth, 1978.

the multiplicity of the term. The determination of all possible terms corresponding to a given configuration amounts to find all the possible values of L and S for the electronic system. To do this, all single–electron orbital angular momenta operators $\boldsymbol{\ell}_i$ need to be combined to get the total $\mathbf{L}$, and likewise for $\mathbf{S}$. When pairing up L and S to get the terms, care must be taken to ensure that the overall electronic wave function (space plus spin) is antisymmetric under the interchange of all the coordinates of any two electrons. In other words, the terms must be built in compliance with the Pauli exclusion principle. A particularly simple situation occurs when the last shell is full and therefore contains $2(2\ell+1)$ equivalent electrons. In this case $L = S = 0$ and the only possible term is ^{1}S. For incomplete shells things are more complicated, especially when dealing with equivalent electrons.

The multiplet wave function is built using the one–electron central–field spin orbitals introduced before, and take the form of linear combinations of Slater determinants. The general procedure to construct them can be split in three steps:

1. Find eigenfunctions of S^2 and S_z.
2. Find eigenfunctions of L^2 and L_z.
3. Construct an antisymmetric wave function from the results of the two previous steps.

To illustrate the method let's find the multiplet wave functions for three simple configurations of two electrons.

- Configuration $(1s)^2$. We have $n_1 = n_2 = 1$, $\ell_1 = \ell_2 = 0$ and $s_1 = s_2 = 1/2$. The total spin of the two–electron system can be $S = 0, 1$, and the spin wave functions are organized in a triplet and a singlet as follows

$$\begin{aligned} {}^3\xi_1 &\equiv |11\rangle = \alpha(1)\alpha(2), \\ {}^3\xi_0 &\equiv |10\rangle = \frac{1}{\sqrt{2}}\Big[\alpha(1)\beta(2) + \beta(1)\alpha(2)\Big], \\ {}^3\xi_{-1} &\equiv |1,-1\rangle \equiv |1\bar{1}\rangle = \beta(1)\beta(2), \\ {}^1\xi_0 &\equiv |00\rangle = \frac{1}{\sqrt{2}}\Big[\alpha(1)\beta(2) - \beta(1)\alpha(2)\Big], \end{aligned} \tag{3.64}$$

where the one–electron spin states have been written in a compact notation as $|1/2, 1/2\rangle \equiv \alpha$ and $|1/2, -1/2\rangle \equiv |1/2, \overline{1/2}\rangle \equiv \beta$. The only possible value for L is 0, and the spatial wave function can be written as

$$|LM_L\rangle = |00\rangle = 1s(1)\, 1s(2). \tag{3.65}$$

Since this is a symmetric wave function, it must be combined with the antisymmetric ${}^1\xi$ spin function, so

$$\begin{aligned} |LSM_LM_S\rangle = |0000\rangle = |{}^1S00\rangle = |{}^1SM_LM_S\rangle \\ = \frac{1}{\sqrt{2}}\, 1s(1)\, 1s(2)\Big[\alpha(1)\beta(2) - \beta(1)\alpha(2)\Big] = \frac{1}{\sqrt{2}} \begin{vmatrix} 1s(1)\alpha(1) & 1s(1)\beta(1) \\ 1s(2)\alpha(2) & 1s(2)\beta(2) \end{vmatrix} \\ = \frac{1}{\sqrt{2}} \begin{vmatrix} \psi_1(\lambda_1) & \psi_2(\lambda_1) \\ \psi_1(\lambda_2) & \psi_2(\lambda_2) \end{vmatrix} \equiv |1s^+\, 1s^-\rangle\,. \end{aligned} \tag{3.66}$$

The third line of this equation shows the Slater determinant using the same notation as in (3.59), for easy comparison. The superscript + in $1s^+$ refers to α ($m_s = 1/2$) and $1s^-$ stands for $1s\beta$ ($m_s = -1/2$). For the closed shell configuration $(1s)^2$ the overall wave function is written using only one Slater determinant and the only possible term is ^{1}S.[11] The Pauli exclusion principle is satisfied because the two electrons have opposite values of m_s, as the notation $|1s^+\, 1s^-\rangle$ makes explicitly clear. Note that a triplet state ^{3}S is impossible because the antisymmetric space function is identically zero,

$$\frac{1}{\sqrt{2}}\Big[1s(1)\, 1s(2) \;-\; 1s(1)\, 1s(2)\Big] = 0. \tag{3.67}$$

Incidentally, this is the reason why the ground state of helium is a spin singlet (parahelium).

- Configuration $(1s)(2s)$. We will see that in this case one single Slater determinant is not sufficient to express the wave function. Here $n_1 = 1$, $n_2 = 2$ and $\ell_1 = \ell_2 = 0$. The only possible value for L is 0, so $|LM_L\rangle = 1s(1)\, 2s(2)$, but now we can construct both symmetric and antisymmetric combinations for the space wave function,

$$\begin{aligned} |LM_L\rangle_{\rm S} = |00\rangle_{\rm S} = \psi_{\rm S} = \frac{1}{\sqrt{2}}\Big[1s(1)\, 2s(2) \;+\; 2s(1)\, 1s(2)\Big], \\ |LM_L\rangle_{\rm A} = |00\rangle_{\rm A} = \psi_{\rm A} = \frac{1}{\sqrt{2}}\Big[1s(1)\, 2s(2) \;-\; 2s(1)\, 1s(2)\Big]. \end{aligned} \tag{3.68}$$

The spin wave functions are the same that were found in the configuration $(1s)^2$. The possible multiplet wave functions for the configuration $(1s)(2s)$ are then

$$\begin{aligned} |{}^1S00\rangle = \psi_{\rm S}\, {}^1\xi_0 = \frac{1}{\sqrt{2}}\Big[1s(1)\, 2s(2) + 2s(1)\, 1s(2)\Big] \frac{1}{\sqrt{2}}\Big[\alpha(1)\beta(2) - \beta(1)\alpha(2)\Big] \\ = \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2}} \begin{vmatrix} 1s(1)\alpha(1) & 2s(1)\beta(1) \\ 1s(2)\alpha(2) & 2s(2)\beta(2) \end{vmatrix} + \frac{1}{\sqrt{2}} \begin{vmatrix} 2s(1)\alpha(1) & 1s(1)\beta(1) \\ 2s(2)\alpha(2) & 1s(2)\beta(2) \end{vmatrix} \right) \\ = \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{2}} \begin{vmatrix} 1s(1)\alpha(1) & 2s(1)\beta(1) \\ 1s(2)\alpha(2) & 2s(2)\beta(2) \end{vmatrix} - \frac{1}{\sqrt{2}} \begin{vmatrix} 1s(1)\beta(1) & 2s(1)\alpha(1) \\ 1s(2)\beta(2) & 2s(2)\alpha(2) \end{vmatrix} \right). \end{aligned}$$

[11] This conclusion holds for any closed shell configuration.

In a more compact notation,

$$|^1S00\rangle = \psi_{\rm S}\,{}^1\xi_0 = \frac{1}{\sqrt{2}}\Big(\,|1s^+\,2s^-\rangle - |1s^-\,2s^+\rangle\,\Big), \tag{3.69}$$

$$|^3S01\rangle = \psi_{\rm A}\,{}^3\xi_1 = \frac{1}{\sqrt{2}}\begin{vmatrix} 1s(1)\alpha(1) & 2s(1)\alpha(1) \\ 1s(2)\alpha(2) & 2s(2)\alpha(2) \end{vmatrix} = |1s^+\,2s^+\rangle\,, \tag{3.70}$$

$$|^3S00\rangle = \psi_{\rm A}\,{}^3\xi_0 = \frac{1}{\sqrt{2}}\Big(\,|1s^+\,2s^-\rangle + |1s^-\,2s^+\rangle\,\Big), \tag{3.71}$$

$$|^3S0\bar{1}\rangle = \psi_{\rm A}\,{}^3\xi_{-1} = |1s^-\,2s^-\rangle\,. \tag{3.72}$$

The $(1s)(2s)$ configuration leads then to ^{1}S and ^{3}S terms, and we see that a single Slater determinant is generally not sufficient to represent a state.

- Configuration $(1s)(2p)$. We have $n_1 = 1$, $n_2 = 2$, $\ell_1 = 0$ and $\ell_2 = 1$. In this case $L = 1$ and the possible terms are ^{1}P, ^{3}P. The space wave functions are

$$|LM_L\rangle = |11\rangle = 1s(1)\,2p_{+1}(2), \qquad |LM_L\rangle^{\rm S}_{\rm A} = |11\rangle^{\rm S}_{\rm A} = \frac{1}{\sqrt{2}}\Big[1s(1)\,2p_{+1}(2) \,\pm\, 2p_{+1}(1)\,1s(2)\Big],$$

$$|LM_L\rangle = |10\rangle = 1s(1)\,2p_0(2), \qquad |LM_L\rangle^{\rm S}_{\rm A} = |10\rangle^{\rm S}_{\rm A} = \frac{1}{\sqrt{2}}\Big[1s(1)\,2p_0(2) \,\pm\, 2p_0(1)\,1s(2)\Big],$$

$$|LM_L\rangle = |1\bar{1}\rangle = 1s(1)\,2p_{-1}(2), \qquad |LM_L\rangle^{\rm S}_{\rm A} = |1\bar{1}\rangle^{\rm S}_{\rm A} = \frac{1}{\sqrt{2}}\Big[1s(1)\,2p_{-1}(2) \,\pm\, 2p_{-1}(1)\,1s(2)\Big].$$

The spin states are still the same as before. It is easy to show that the total states for the $(1s)(2p)$ configuration are

$$|^1{\rm P}10\rangle = \frac{1}{\sqrt{2}}\Big(\,|1s^+\,2p^-_{+1}\rangle - |1s^-\,2p^+_{+1}\rangle\,\Big), \qquad |^3{\rm P}01\rangle = |1s^+\,2p^+_0\rangle\,,$$

$$|^1{\rm P}00\rangle = \frac{1}{\sqrt{2}}\Big(\,|1s^+\,2p^-_0\rangle - |1s^-\,2p^+_0\rangle\,\Big), \qquad |^3{\rm P}00\rangle = \frac{1}{\sqrt{2}}\Big(\,|1s^+\,2p^-_0\rangle + |1s^-\,2p^+_0\rangle\,\Big),$$

$$|^1{\rm P}\bar{1}0\rangle = \frac{1}{\sqrt{2}}\Big(\,|1s^+\,2p^-_{-1}\rangle - |1s^-\,2p^+_{-1}\rangle\,\Big), \qquad |^3{\rm P}0\bar{1}\rangle = |1s^-\,2p^-_0\rangle\,,$$

$$|^3{\rm P}11\rangle = |1s^+\,2p^+_{+1}\rangle\,, \qquad |^3{\rm P}\bar{1}1\rangle = |1s^+\,2p^+_{-1}\rangle\,,$$

$$|^3{\rm P}10\rangle = \frac{1}{\sqrt{2}}\Big(\,|1s^+\,2p^-_{+1}\rangle + |1s^-\,2p^+_{+1}\rangle\,\Big), \qquad |^3{\rm P}\bar{1}0\rangle = \frac{1}{\sqrt{2}}\Big(\,|1s^+\,2p^-_{-1}\rangle + |1s^-\,2p^+_{-1}\rangle\,\Big),$$

$$|^3{\rm P}1\bar{1}\rangle = |1s^-\,2p^-_{+1}\rangle\,, \qquad |^3{\rm P}\bar{1}\bar{1}\rangle = |1s^-\,2p^-_{-1}\rangle\,.$$

We deal next with the general problem of finding all possible terms for a given configuration, starting first with inequivalent electrons.

Terms for configurations with only inequivalent electrons.

It is easy to find the Russell–Saunder LS terms for a configuration corresponding to two inequivalent electrons, say $(np)(n'p)$. Since in this case $\ell_1 = \ell_2 = 1$ and $s_1 = s_2 = 1/2$, the list of values for the orbital and spin angular momentum of the two-electron system is $L = 0, 1, 2$ and $S = 0, 1$, respectively. The electron orbitals differ in the principal quantum number, therefore we can freely combine the three values of L with the two values of S without violating the Pauli exclusion principle. The six terms that arise are

$$(np)(n'p): \qquad {}^1{\rm S},\ {}^1{\rm P},\ {}^1{\rm D},\ {}^3{\rm S},\ {}^3{\rm P},\ {}^3{\rm D}. \tag{3.73}$$

As another example consider the configuration $(np)(n'd)$, where $\ell_1 = 1$ and $\ell_2 = 2$. Now we have $L = 1, 2, 3$ and $S = 0, 1$. The possible terms are

$$(np)(n'd): \qquad {}^1{\rm P},\ {}^1{\rm D},\ {}^1{\rm F},\ {}^3{\rm P},\ {}^3{\rm D},\ {}^3{\rm F}. \tag{3.74}$$

If there are more than two electrons, we first combine the angular momenta of two and then add the remaining ones in successive steps. Let's consider the configuration $(np)(n'p)(n''d)$. From the pair $(np)(n'p)$ we get the terms displayed in equation (3.73). Next, we add the third $(n''d)$ electron, which has $\ell = 2$, to each one of those terms. When adding it to the ^{1}S term, we obtain a new term with $L = 2$ and $S = 1/2$, which is ^{2}D. Adding the $(n''d)$ electron to all the terms in (3.73) produces the following list.

- The ^{1}S term gives rise to: ^{2}D.
- The ^{1}P term gives rise to: ^{2}P, ^{2}D, ^{2}F.
- The ^{1}D term gives rise to: ^{2}S, ^{2}P, ^{2}D, ^{2}F, ^{2}G.
- The ^{3}S term gives rise to: ^{2}D, ^{4}D.
- The ^{3}P term gives rise to: ^{2}P, ^{2}D, ^{2}F, ^{4}P, ^{4}D, ^{4}F.
- The ^{3}D term gives rise to: ^{2}S, ^{2}P, ^{2}D, ^{2}F, ^{2}G, ^{4}S, ^{4}P, ^{4}D, ^{4}F, ^{4}G.

This can be written in a compact way such as

$$(np)(n'p)(n''d): \quad {}^2(\underset{2}{\mathrm{S}}\,\underset{4}{\mathrm{P}}\,\underset{6}{\mathrm{D}}\,\underset{4}{\mathrm{F}}\,\underset{2}{\mathrm{G}}) \quad {}^4(\underset{2}{\mathrm{S}}\,\underset{3}{\mathrm{P}}\,\underset{2}{\mathrm{D}}\,\underset{}{\mathrm{F}}\,\mathrm{G}) \tag{3.75}$$

where the number under each symbol indicates the number of identical terms.

Terms for configurations with equivalent electrons.

For equivalent electrons a considerable complication arises: not all combinations of L and S are compatible with the Pauli principle, so we cannot pair all of them up. Let's examine the configuration np^2, for which the possible values of L and S are $L = 0, 1, 2$ and $S = 0, 1$. For $S = 0$ we have an antisymmetric spin wave function, the singlet state, while for $S = 1$ (triplet) the three states are symmetric under the interchange of the two particles. To have an overall antisymmetric wave function we must then combine $L = 0, 2$ (symmetric space wave function) with $S = 0$ only, and $L = 1$ (antisymmetric) with $S = 1$ only. The possible LS terms are

$$(np^2): \qquad {}^1\mathrm{S},\ {}^1\mathrm{D},\ {}^3\mathrm{P}. \tag{3.76}$$

Note that for these three terms, the value of $L + S$ is an even integer. It can be shown that this is a general rule for the terms arising from configurations of two equivalent electrons, namely $L + S$ must be even. However, for more than two electrons there is no comparable rule. Another important point to mention is that the parity of a term *is not* $(-1)^L$. Since the $|LSM_LM_S\rangle$ states are constructed from products of central field orbitals, each of which contains a spherical harmonic, the parity of an N–electron state is given by $(-1)^{\ell_1}(-1)^{\ell_2}\ldots(-1)^{\ell_N}$. For a two-electron state, for example, the parity is odd or even depending on whether $\ell_1 + \ell_2$ is odd or even. In fact, for two equivalent electrons the parity is always even because the sum of any pair of equal integers is even. We must be careful not to associate the parity with the value of L.

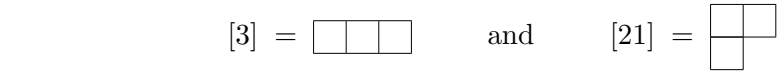
In the np^3 configuration, for example, we can build antisymmetric wave functions with respect to the permutation of only two out of the three electrons. Consequently, we expect in this case wave functions with mixed symmetry: symmetric under the permutation of one pair of electrons and antisymmetric under the permutation of the other pair (we can make two pairs out of three particles). It is of course possible to classify the n–electron wave functions according to their symmetry properties, and find the LS terms, following essentially a *bookkeeping* method, but the complexity of the problem rapidly escalates as n increases.

Classification of atomic terms using Young tableaux.

The general problem of classifying atomic terms according to the symmetry properties of the spin and space wave functions can be handled by group theoretical methods. There is no time to go into that procedure in this course in detail, so we will just summarize here the main results in a very simplified way.

The LS Russell–Saunders terms compatible with the Pauli exclusion principle in a configuration of k equivalent electrons can be found by going through the following four steps.

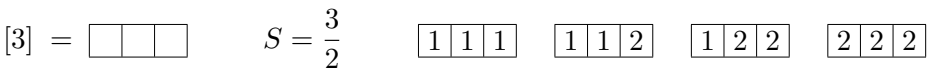
1. Build all the possible Young diagrams for the spin wave functions of k electrons. These spin diagrams have k boxes organized in a maximum of 2 rows. The number of boxes in the second row must be equal or less than that of the first row. As an example, the two possible spin Young diagrams for $k = 3$ electrons are



(3.77)

   The notation [3] means 3 boxes in the (only) first row, while [21] means 2 boxes in the first row and 1 box in the second row. A single column of k boxes will be denoted also as $[1^k]$, so that the notations [111] and $[1^3]$ refer to the same diagram (this would be a space diagram, however, because a spin diagram cannot have more than two rows).

2. Each box can represent one of the two possible spin third–projections of the electrons, which we may denote by 1 and 2 instead of $+1/2$ and $-1/2$. Fill the boxes with the symbols 1 and 2 following the rule that the numbers must increase when going down a column and must not decrease when going from left to right along a row. A diagram filled with a permissible sequence of symbols is called a Young tableau. The number of different Young tableaux associated with a Young diagram is equal to the number of spin states in the corresponding multiplet. For example, the diagram [3] represents a quartet (spin $S = 3/2$) because we can build only $4 = 2S + 1$ different Young tableaux:

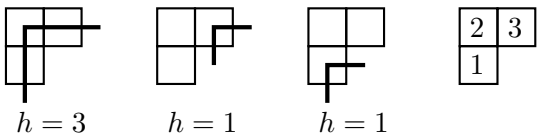


(3.78)

   The diagram [21] represents a doublet (spin $S = 1/2$) because we can build only $2 = 2S+1$ different Young tableaux:

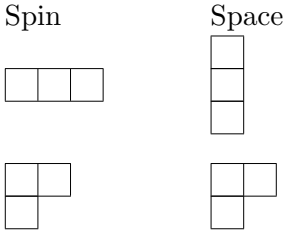
(3.79)

   A more convenient way to find the number of spin Young tableaux associated with a Young diagram is to use the so-called *factors over hooks rule*, which we explain now. Put a *factor* of 2 in the upper-left box. Then put factors in all boxes by adding 1 as you move to the right and subtracting 1 as you move down. The product of all these factors is F. A *hook* is a line passing vertically up through the bottom of a column, making a right-hand turn in a given box, and going out through the row of boxes. The number of boxes that crosses the hook is h. Draw one hook for each box in the Young diagram. Multiply all h-values and call the result H. The number of Young tableaux associated with the Young diagram is given by the ratio F/H. Example:



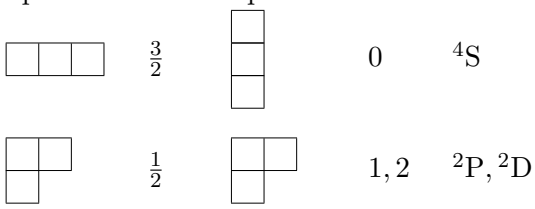
$$\frac{F}{H} = \frac{2 \times 3 \times 1}{3 \times 1 \times 1} = 2.$$

3. Two Young diagrams, Y_1 and Y_2, are said to be adjoint to each other if the number of rows in Y_1 is equal to the number of columns in Y_2. Furthermore, the number of boxes in the first row of Y_1 must be equal to the number of boxes in the first column of Y_2, etc. In other words, you go from Y_1 to Y_2 changing rows by columns. To each spin Young diagram we must associate an adjoint space Young diagram. Thus the space Young diagrams can have a maximum of two columns. For our example of $k = 3$ electrons we would have the following correspondence



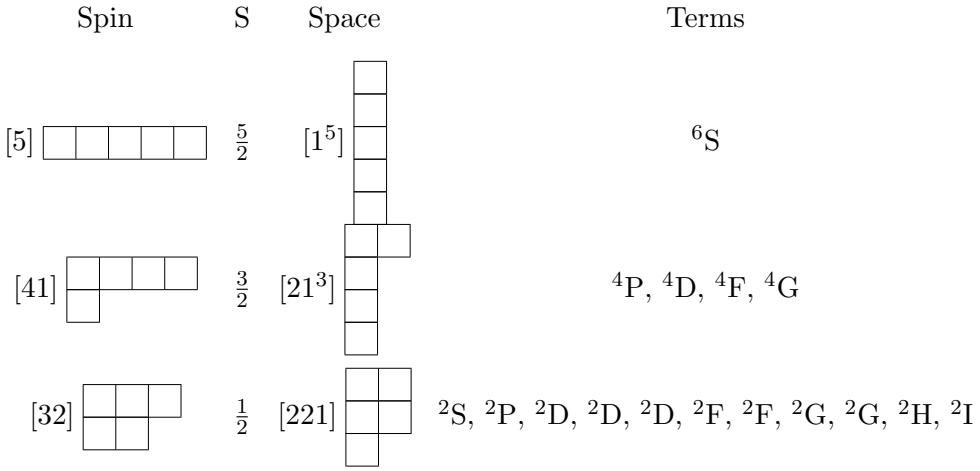
(3.80)

4. It remains now to find the possible values of the orbital angular momentum, L, which correspond to each space Young diagram. This is a bit more complicated and we merely list the main results in table (3.1). Assuming we are dealing with 3 electrons in a p orbital, which is a $(np)^3$ configuration, we finally arrive at the following list



(3.81)

Repeat this procedure for the $(nd)^5$ configuration and check that you obtain



In a slightly shorter notation, the terms for this configuration can be written as

$$(nd)^5: \qquad {}^2(\mathrm{S\,P\,}\underset{3}{\mathrm{D}}\,\underset{2}{\mathrm{F}}\,\underset{2}{\mathrm{G}}\,\mathrm{H\,I})\quad {}^4(\mathrm{P\,D\,FG})\quad {}^6\mathrm{S} \tag{3.82}$$

It can be seen that the configurations $(n\ell)^k$ and $(n\ell)^{2(2\ell+1)-k}$ lead to the same set of terms. In other words, the possible terms corresponding to a configuration in which there are k electrons in a (sub)shell are the same as those in which k electrons are missing (that is, there are k *holes*) in this (sub)shell. Pairs of configurations that give rise to the same terms are, for example, $(nd)^4$ and $(nd)^6$, $(nd)^3$ and $(nd)^7$, $(nd)^2$ and $(nd)^8$, etc. It should be clear also that the only possible term for a closed shell is ^{1}S. If a configuration contains a group of equivalent electrons together with a number of non–equivalent electrons, the possible terms of the group of equivalent

ℓ	k	Young	L
1	1	[1]	1
	2	[2]	0, 2
		$[1^2]$	1
	3	[21]	1, 2
		$[1^3]$	0
2	1	[1]	2
	2	[2]	0
			2, 4
		$[1^2]$	1, 3
	3	[21]	2
			1, 2, 3, 4, 5
		$[1^3]$	1, 3
	4	[22]	0
			2, 4
			0, 2, 3, 4, 6
		[211]	1, 3
			1, 2, 3, 4, 5
		$[1^4]$	2
	5	[221]	2
			1, 2, 3, 4, 5
			0, 2, 3, 4, 6
		$[21^3]$	1, 3
			2, 4
		$[1^5]$	0

Table 3.1: Orbital angular momentum assignments to several Young diagrams for p ($\ell = 1$) and d ($\ell = 2$) orbitals and k electrons. We only need to go up to half-filled shells because the configurations $(n\ell)^k$ and $(n\ell)^{2(2\ell+1)-k}$ lead to the same set of terms. Note that a given value of L may appear multiple times in the same diagram.

electrons must first be determined. The overall possible terms are then obtained by using the ordinary rules of addition of angular momenta to add the non–equivalent electrons. Similarly, if a configuration contains two or more groups of equivalent electrons, the possible terms of each group must first be obtained, and the overall possible terms are then determined by using the usual rules for the addition of angular momenta.

Classification of atomic terms in jj coupling.

For atoms with $Z < 30$ the LS coupling scheme provides a good description of atomic terms. For larger Z, however, the spin-orbit interaction is gradually increasing and eventually becomes dominant over the electron repulsion interaction. For $Z > 50$ the most convenient coupling scheme is certainly the jj-coupling, where the orbital and spin angular momentum are combined first for each electron separately, and then the resulting total momenta are combined for all electrons. Namely

$$\mathbf{j}_i = \boldsymbol{\ell}_i + \mathbf{s}_i, \qquad \mathbf{J} = \sum_i \mathbf{j}_i. \tag{3.83}$$

In the jj-scheme, one–electron states are denoted by ℓ_j. For example,

$$s_{1/2},\ \ p_{1/2},\ \ p_{3/2},\ \ d_{3/2},\ \ d_{1/2}, \ldots \tag{3.84}$$

Each of these states has a degeneracy of $2j+1$. A two-electron state is denoted by $(j_1 j_2)_J$. For inequivalent electrons, the possible values of J are obtained following the usual rule of addition of angular momenta, $J = j_1 + j_2, j_1 + j_2 - 1, \ldots, |j_1 - j_2|$. A three-electron state is denoted by $(j_1 j_2 j_3)_J$, etc. For the pd configuration, for example, the possible jj coupling states are

$$\left(\frac{1}{2}\frac{3}{2}\right)_{1,2}, \quad \left(\frac{1}{2}\frac{5}{2}\right)_{2,3}, \quad \left(\frac{3}{2}\frac{3}{2}\right)_{0,1,2,3}, \quad \left(\frac{3}{2}\frac{5}{2}\right)_{1,2,3,4}. \tag{3.85}$$

When writing the spectral terms it is standard practice to specify the individual quantum numbers (n_i, ℓ_i, j_i) of each electron and the total electronic angular momentum quantum number, J. For example, one of the energy levels of the ground state configuration $(6p)^2$ of Pb, corresponding to the values $j_1 = 1/2$, $j_2 = 3/2$ and $J = 2$, is written as

$$6p^2 \left(\frac{1}{2}\frac{3}{2}\right)_2.$$

Like in LS-coupling, for equivalent electrons not all states which come out of the angular momentum coupling rules are possible. We give two simple rules to find the states of two equivalent electrons.

1. When $j_1 \neq j_2$, the value of J in $(j_1 j_2)_J$ is just given by the ordinary rules of angular momentum coupling, namely $J = j_1 + j_2, \ldots, |j_1 - j_2|$. Since the electrons are equivalent, the states $(j_1 j_2)_J$ are identical to the states $(j_2 j_1)_J$, so we only need to retain, for example, those states for which $j_1 < j_2$.

2. When $j_1 = j_2 = j$, the allowed values of J in $(j_1 j_2)_J$ are given by $J = 2j-1, 2j-3, 2j-5, \ldots$ The values $J = 2j, 2j-2, 2j-4, \ldots$ are forbidden. This result can be derived from the following symmetry property of the Clebsch-Gordan coefficients,

$$\langle jjm_1m_2 \,|\, jjJM_J\rangle = (-1)^{2j-J} \langle jjm_2m_1 \,|\, jjJM_J\rangle . \tag{3.86}$$

   Indeed, since $2j$ is an *odd number*, we need J to be an *even number* to get $(-1)^{2j-J} = -1$ and make sure the wave function is anti-symmetric under the interchange of the two equivalent electrons.

As an example, the possible states in jj-coupling for the $(np)^2$ configuration are

$$\left(\frac{1}{2}\frac{1}{2}\right)_0, \quad \left(\frac{1}{2}\frac{3}{2}\right)_{1,2}, \quad \left(\frac{3}{2}\frac{3}{2}\right)_{0,2}. \tag{3.87}$$

In the LS–coupling scheme, this same configuration gives the following terms (see equation (3.76))

$$^1\mathrm{S}_0, \quad ^3\mathrm{P}_{0,1,2}, \quad ^1\mathrm{D}_2. \tag{3.88}$$

For more than two equivalent electrons is not possible to give such a simple rule to compute the total angular momentum J and one has to perform a (somewhat tedious) analysis using group theoretical methods. In the table (3.2) we list the values of the total angular momenta for the most common $(j)^k$ configurations. Note that configurations with k electrons give the same values of J than configurations with $(2j+1) - k$ electrons.

Since equivalent electrons are in the same orbital, the Coulomb repulsion is large and LS–coupling is expected to be more appropriate. It is for inequivalent electrons that the jj–coupling is more interesting, because in this case the spin-orbit interaction may be larger that the Coulomb repulsion. Nevertheless, there are many situations for which neither LS nor jj coupling are valid approximations. In such cases, it is necessary to diagonalize the Coulomb repulsion and the spin-orbit interaction for each value of J. The coupling is then said to be *intermediate*.

The number of states obtained in LS and jj coupling is of course the same. Figure (3.4) shows, for a ps configuration, the transition (not to scale) from LS to jj coupling and how the four LS terms are related to the four jj terms. The energy dependence of the terms is different for each coupling scheme.

		J	g
$j = 1/2$	$k = 1$	$1/2$	2
	2	0	1
$j = 3/2$	$k = 1, 3$	$3/2$	4
	2	$2, 0$	6
$j = 5/2$	$k = 1, 5$	$5/2$	6
	$2, 4$	$4, 2, 0$	15
	3	$9/2, 5/2, 3/2$	20
$j = 7/2$	$k = 1, 7$	$7/2$	8
	$2, 6$	$6, 4, 2, 0$	28
	$3, 5$	$15/2, 11/2, 9/2, 7/2, 5/2, 3/2$	56
	4	$8, 6, 5, (4)^2, (2)^2, 0$	70

Table 3.2: Total angular momenta and degeneracies in the configuration $(j)^k$.

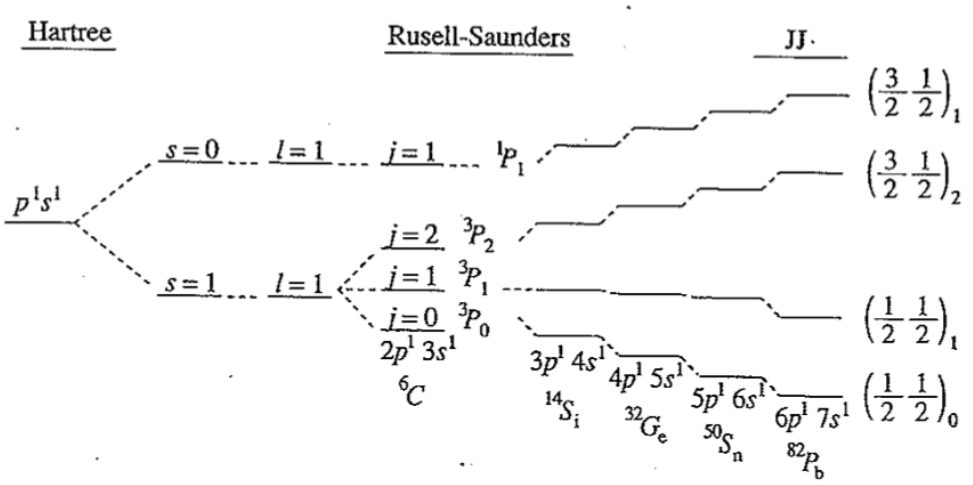


Figure 3.4: Comparison between LS and jj levels for the same configuration in a series of atoms with increasing values of Z (figure adapted from *Sáchez del Río, Física Cuántica, eds. Pirámide, 1997*).

Hund's rules.

The ground state of an atom is determined by minimizing the energy of its valence electrons with the residual electrostatic and spin-orbit interactions included. This is a very complicated calculation. There are, however, a set of so-called Hund's rules (established empirically) to determine which level is the ground state for atoms that have LS coupling.

1. The term with the highest total spin, S, has the lowest energy.
2. For a given spin, S, the term with largest L has the lowest energy.
3. The level with $J = |L - S|$ or $J = L + S$ has the lowest energy depending, respectively, on whether the (sub)shell is less than, or more than, half-full.

A slightly different wording of the rules to find the ground state is

1. Maximize the spin and set $S = \sum m_s$.
2. Maximize the orbital angular momentum, subject to the rule 1, and set $L = \sum m_\ell$.

3. Set $J = |L - S|$ if the shell is less than half–full; otherwise set $J = L + S$.

Note that rules 1 and 2 must be applied taking into account Pauli's exclusion principle to guarantee that no pair of equivalent electrons have the same values of m_s and m_ℓ. Consider, for example, the $(nd)^2$ configuration. The maximum value of the total spin is $S = 1$, corresponding to a value of $m_s = 1/2$ for both electrons. Then, for a state with $S = 1$, the two electrons must have different values of m_ℓ, giving a maximum possible value of the third component of the total orbital angular momentum equal to $M_L = 2+1 = 3$. Therefore, $L = 4$ is not possible for a term with $S = 1$ in the $(nd)^2$ configuration.

Although these rules are empirical and there are exceptions to them, it is convenient to understand their physical origin. First, note that maximizing S makes the spin wave function as symmetric as possible. This tends to make the spatial wave function antisymmetric, and hence reduces the Coulomb repulsion. Second, maximizing L also tends to keep the electrons apart. This is less obvious, although a simple classical picture of electrons rotating round the nucleus in the same or different senses makes it at least plausible. Finally, the separation of energies for states of different J arises from treating as a perturbation the spin-orbit interaction (fine structure). In section 3.2.4 it will be shown that the matrix element of the spin-orbit Hamiltonian satisfies the relation

$$\langle LSJM_J \,|\, \mathcal{H}_{\rm so} \,|\, LSJM_J\rangle = \frac{\hbar^2\zeta(L,S)}{2}\Big[J(J+1) - L(L+1) - S(S+1)\Big], \tag{3.89}$$

where $\zeta(L,S)$ depends on the total L and S values, but not on J. Since it can also be shown that the sign of $\zeta(L,S)$ changes according to whether the subshell is more of less than half-filled, the third Hund's rule is established.

3.2.4 Spin-orbit interaction.

The study of the spin-orbit interaction for an N-electron system is just an extension of what we have already done for a single electron. The Hamiltonian to consider is

$$\mathcal{H}_{\rm so} = \sum_{i=1}^{N} \xi(r_i)\boldsymbol{\ell}_i\boldsymbol{s}_i. \tag{3.90}$$

We will assume LS coupling, which implies that the electronic repulsion energy is much greater than the spin-orbit energy. The wave function terms $|LSM_LM_S\rangle$ have energies that differ from one another as a consequence of the Coulomb repulsion among electrons. As in the one-electron case, it is convenient to work in the representation $|LSJM_J\rangle$. To obtain the energies associated with the spin-orbit interaction we need to evaluate the matrix elements of the operator products $\boldsymbol{\ell}_i\boldsymbol{s}_i$. This can be done using the Wigner-Eckart theorem.[12] The result is

$$\begin{aligned}\big\langle LSJM_J \,\big|\, \boldsymbol{\ell}_i\boldsymbol{s}_i \,\big|\, L'S'J'M'_J\big\rangle \\ = (-1)^{J+S+L'}\delta(J,J')\delta(M_J,M'_J)\begin{Bmatrix} L' & S' & J \\ S & L & 1\end{Bmatrix}\big\langle L \,\big\|\, \boldsymbol{\ell}_i \,\big\|\, L'\big\rangle\big\langle S \,\big\|\, \boldsymbol{s}_i \,\big\|\, S'\big\rangle.\end{aligned} \tag{3.91}$$

The $6j$ symbol vanishes unless the triangle conditions $\triangle(L'L1)$ and $\triangle(S'S1)$ are satisfied (see appendix B.5.3). Hence, the matrix element of $\boldsymbol{\ell}_i\boldsymbol{s}_i$ vanishes unless

$$\begin{aligned} \Delta S = 0,\pm1, \qquad & S' + S \geq 1, \\ \Delta L = 0,\pm1, \qquad & L' + L \geq 1, \\ \Delta J = 0, \qquad & \Delta M_J = 0. \end{aligned} \tag{3.92}$$

[12] For more details see section 21.4 of [16]

It is important to note that the spin-orbit interaction has non-vanishing matrix elements between states that differ in spin by one unit. This implies that the spin-orbit interaction can produce an admixture of singlet and triplet states. Nevertheless, the diagonal elements are of special importance because they represent the spin-orbit interaction within a particular term, which in most cases is well separated from other terms. Setting $L' = L$ and $S' = S$, the $6j$ symbol becomes

$$\left\{ \begin{array}{ccc} L & S & J \\ S & L & 1 \end{array} \right\} = (-1)^{L+S+J} \frac{J(J+1) - L(L+1) - S(S+1)}{\sqrt{L(2L+1)(2L+2)S(2S+1)(2S+2)}}. \tag{3.93}$$

So we may write

$$\begin{aligned} \langle \alpha LSJM_J \,|\, \mathcal{H}_{\rm so} \,|\, \alpha LSJM_J \rangle &= \left\langle \alpha LSJM_J \left| \sum_{i=1}^{N} \xi(r_i) \boldsymbol{\ell}_i \boldsymbol{s}_i \right| \alpha LSJM_J \right\rangle \\ &= \frac{\hbar^2}{2} \zeta(\alpha, L, S) \Big[J(J+1) - L(L+1) - S(S+1) \Big] = E(J), \end{aligned} \tag{3.94}$$

where α represents the dependence of the wave function on the radial coordinates and $\zeta(\alpha, L, S)$ is a proportionality factor which depends on α, L and S, but not on J. The reduced matrix elements, phase factors and the dependence on $\xi(r_i)$, are all contained in $\zeta(\alpha, L, S)$. We have written the spin-orbit matrix element in such a way because $\mathbf{LS} = (J^2 - L^2 - S^2)/2$, and

$$\begin{aligned} \langle \alpha LSJM_J \,|\, \mathbf{LS} \,|\, \alpha LSJM_J \rangle &= \frac{1}{2} \left\langle \alpha LSJM_J \,\middle|\, (J^2 - L^2 - S^2) \,\middle|\, \alpha LSJM_J \right\rangle \\ &= \frac{\hbar^2}{2} \Big[J(J+1) - L(L+1) - S(S+1) \Big], \end{aligned} \tag{3.95}$$

which then permits the identification

$$\langle \alpha LSJM_J \,|\, \mathcal{H}_{\rm so} \,|\, \alpha LSJM_J \rangle = \zeta(\alpha, L, S) \, \langle \alpha LSJM_J \,|\, \mathbf{LS} \,|\, \alpha LSJM_J \rangle \,. \tag{3.96}$$

This shows that, within a multiplet of states comprising a given LS term, the matrix elements of $\mathcal{H}_{\rm so}$ are proportional to the matrix elements of $\mathbf{LS}$. Therefore, for this subspace of states having the same L and S quantum numbers, we can write

$$\mathcal{H}_{\rm so} \;=\; \zeta \, \mathbf{LS}, \qquad\qquad \text{(for a given LS term)}. \tag{3.97}$$

Equation (3.94) contains the so-called *Landé interval rule*, which states that the energy differences between adjacent levels of a multiplet are proportional to J,

$$\Delta E(J) = E(J) - E(J-1) = \hbar^2 \zeta(\alpha, L, S) \, J. \tag{3.98}$$

For a configuration of equivalent electrons it is easy to give a more explicit interpretation of the constant $\zeta(\alpha, L, S)$. For this purpose we work in the $|LSM_L M_S\rangle$ basis and note that

$$\begin{aligned} \langle LSM_L M_S \,|\, \mathbf{LS} \,|\, LSM_L M_S \rangle &= \langle LSM_L M_S \,|\, -L_{+1} S_{-1} + L_0 S_0 - L_{-1} S_{+1} \,|\, LSM_L M_S \rangle \\ &= \langle LSM_L M_S \,|\, L_0 S_0 \,|\, LSM_L M_S \rangle = \hbar^2 M_L M_S. \end{aligned} \tag{3.99}$$

Since

$$M_L = \sum_{i=1}^{N} m_\ell^i, \qquad \text{and} \qquad S_L = \sum_{i=1}^{N} s_\ell^i, \tag{3.100}$$

the matrix element of $\mathcal{H}_{\rm so}$ becomes

$$\begin{aligned}\langle \alpha LSJM_J \,|\, \mathcal{H}_{\rm so} \,|\, \alpha LSJM_J\rangle &= \left\langle \alpha LSJM_J \,\middle|\, \sum_{i=1}^{N} \xi(r)\boldsymbol{\ell}_i\boldsymbol{s}_i \,\middle|\, \alpha LSJM_J \right\rangle \\ &= \sum_{i=1}^{N} \left\langle LSm_\ell^i m_s^i \,\middle|\, \boldsymbol{\ell}_i\boldsymbol{s}_i \,\middle|\, LSm_\ell^i m_s^i \right\rangle \langle \xi(r)\rangle_{n\ell} \\ &= \sum_{i=1}^{N} \left\langle LSm_\ell^i m_s^i \,\middle|\, -\ell_{+1}^i s_{-1}^i + \ell_0^i s_0^i - \ell_{-1}^i s_{+1}^i \,\middle|\, LSm_\ell^i m_s^i \right\rangle \langle \xi(r)\rangle_{n\ell} \\ &= \sum_{i=1}^{N} \left\langle LSm_\ell^i m_s^i \,\middle|\, \ell_0^i s_0^i \,\middle|\, LSm_\ell^i m_s^i \right\rangle \langle \xi(r)\rangle_{n\ell} = \hbar^2 \langle \xi(r)\rangle_{n\ell} \sum_{i=1}^{N} m_\ell^i m_s^i. \end{aligned} \tag{3.101}$$

Here $\langle \xi(r)\rangle_{n\ell}$ is the expectation value of $\xi(r)$ over the radial part of the wave function which, for equivalent electrons, is the same for all electrons. Now, the result in (3.96) must hold also in the $\langle LSM_LM_S\rangle$ basis, therefore we have

$$\hbar^2 \langle \xi(r)\rangle_{n\ell} \sum_{i=1}^{N} m_\ell^i m_s^i = \zeta(\alpha, L, S) \langle \alpha LSJM_J \,|\, \mathbf{LS} \,|\, \alpha LSJM_J\rangle = \hbar^2 \zeta(\alpha, L, S) M_L M_S.$$

This leads to

$$\zeta(\alpha, L, S) = \frac{\langle \xi(r)\rangle_{n\ell}}{M_L M_S} \sum_{i=1}^{N} m_\ell^i m_s^i. \tag{3.102}$$

Since $\zeta(\alpha, L, S)$ is independent of M_L and M_S, these may be chosen arbitrarily, the most convenient choice being the maximum values of M_L and M_S. In this way we find, for example, that

$$\zeta(\mathrm{p}^2, {}^3\mathrm{P}) = \frac{1}{2} \langle \xi(r)\rangle_{\rm p}, \qquad \zeta(\mathrm{d}^2, {}^3\mathrm{F}) = \frac{1}{2} \langle \xi(r)\rangle_{\rm d}. \tag{3.103}$$

Note that for a close shell there is no spin-orbit splitting.

It is possible to evaluate the reduced matrix elements in (3.91) for k equivalent electrons in a shell ℓ, thereby obtaining an explicit expression for $\zeta(\ell^k, \alpha, L, S)$. Here we just quote the following properties

Multiplet	$\zeta(\ell^k, \alpha, L, S)$	Shell
Regular	> 0	less than half-full, $k < 2\ell + 1$
Inverted	< 0	more than half-full, $k > 2\ell + 1$
No splitting	0	half-full, $k = 2\ell + 1$

For a configuration ℓ^k, the constant $\zeta(\ell^k, \alpha, L, S)$ is positive for $k < 2\ell+1$, negative for $k > 2\ell+1$, and zero for $k = 2\ell + 1$. Note that, according to the Landé interval rule (3.98), the sign of ζ determines whether the energy of the levels in a multiplet increases or decreases with the value of J. It is customary to define a *regular* multiplet as one in which the energy increases with increasing values of J, and an *inverted* multiplet as one in which the energy increases with decreasing values of J. The absence of any multiplet splitting in a shell that is half-full is correct to first order in the spin-orbit interaction, that is, within an LS term. In higher order, when spin-orbit interaction between terms is taken into account, a half-filled shell will generally exhibit some multiplet splitting.

3.2.5 The Thomas-Fermi model.

For the determination of the potential energy $U(r)$ of the central field approximation we will explore two methods. In this section we discuss the first one, due to Thomas and Fermi, and

the second, due mainly to Hartree, is taken up in the next section. The Thomas–Fermi model is statistical in nature, it assumes that the variation of $U(r)$ is so slow that many electrons can be localized within a volume where the potential is nearly constant. The electrons can then be treated with the methods of statistical mechanics, obeying the Fermi–Dirac statistics.

Consider a cubic box of side L and impose boundary conditions for the electronic wave functions at each face of the cube. Since the potential is assumed to be nearly constant in the volume of our box, the wave functions of the electrons can be taken to be approximately plane waves (only in the box). The periodic boundary conditions we should impose are then

$$e^{ik_x x} = e^{ik_x(x+L)}, \qquad e^{ik_y y} = e^{ik_y(y+L)}, \qquad e^{ik_z z} = e^{ik_z(z+L)}, \tag{3.104}$$

therefore

$$k_x = \frac{2\pi N_x}{L}, \qquad k_y = \frac{2\pi N_y}{L}, \qquad k_z = \frac{2\pi N_z}{L}, \qquad N_x, N_y, N_z = 0, \pm 1, \pm 2, \ldots \tag{3.105}$$

These are the components of the vector $\mathbf{k} = \mathbf{p}/\hbar$. The periodic conditions give rise to an infinite (discrete) set of sates in the cavity, each one characterized by the set of integers (N_x, N_y, N_z). The number of states contained in an interval specified by ΔN_x, ΔN_y and ΔN_z is just given by

$$\Delta N = \Delta N_x \Delta N_y \Delta N_z = \left(\frac{L}{2\pi}\right)^3 \Delta k_x \Delta k_y \Delta k_z. \tag{3.106}$$

For a box with dimensions large compared to the electron wavelength we may assume the states to form a continuous distribution. We can then write the number of states contained in an interval dk and in the direction defined by the element of solid angle $d\Omega$ as

$$\mathrm{d}N = \left(\frac{L}{2\pi}\right)^3 \mathrm{d}k_x\, \mathrm{d}k_y\, \mathrm{d}k_z \;= \left(\frac{L}{2\pi}\right)^3 k^2 \sin\theta\, \mathrm{d}k\, \mathrm{d}\theta\, \mathrm{d}\varphi. \tag{3.107}$$

This should be multiplied by 2 to take into account the two possible spin states of each electron. The number of states with momentum $\mathbf{p} = \hbar\mathbf{k}$ less than p_0 is then just

$$2\left(\frac{L}{2\pi}\right)^3 \int_0^{p_0/\hbar} \int_0^{\pi} \int_0^{2\pi} k^2 \,\sin\theta\, \mathrm{d}k\, \mathrm{d}\theta\, \mathrm{d}\varphi \;=\; 2\left(\frac{L}{2\pi}\right)^3 4\pi \int_0^{p_0/\hbar} k^2\, \mathrm{d}k \;=\; \frac{p_0^3 L^3}{3\pi^2\hbar^3}. \tag{3.108}$$

If all these states are occupied, the number of electrons per unit volume whose kinetic energy does not exceed $\dfrac{p_0^2}{2m}$ is

$$\rho(r) = \frac{p_0^3}{3\pi^2\hbar^3}. \tag{3.109}$$

But the maximum kinetic energy at a distance r from the nucleus must be $-U(r)$, since otherwise the electron would escape from the atom. We can then take $p_0 = \sqrt{-2mU(r)}$ and write the relation between the volume density of electrons, $\rho(r)$, and the potential energy as

$$\rho(r) = \frac{\big[-2mU(r)\big]^{3/2}}{3\pi^2\hbar^3}. \tag{3.110}$$

The electrostatic potential, $-U(r)/e$, can also be determined using the Poisson's equation[13] in terms of the charge density, $-e\rho(r)$,

$$-\frac{1}{e}\nabla^2 U = -\frac{1}{er^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}U}{\mathrm{d}r}\right) = 4\pi e\rho(r). \tag{3.111}$$

[13]Remember that the Poisson equation in Gaussian units is $\nabla^2 V = -4\pi\rho$, while in the SI system is $\nabla^2 V = -\rho/\varepsilon$, where ρ is the electric charge density (in the text above we are denoting the electric charge density as $-e\rho$, and the electrostatic potential as $-U/e$).

Equations (3.110) and (3.111) are two simultaneous equations for ρ and U. Appropriate boundary conditions should be provided to solve them. For a neutral atom of atomic number Z, the potential should approach the form $-Ze^2/r$ as $r \to 0$, since the dominant contribution at short distances should be the potential created by the nucleus. As $r \to \infty$ it is clear that no net charge remains in the sphere of radius r, so $U(r)$ must fall off more rapidly than $1/r$ and we should have $rU(r) \to 0$.

The elimination of $\rho(r)$ in equations (3.110) and (3.111) leads to the following equation for $-U(r)$,

$$\frac{1}{r^2}\frac{\mathrm{d}}{\mathrm{d}r}\left(r^2\frac{\mathrm{d}(-U)}{\mathrm{d}r}\right) = \frac{4e^2\big[-2mU(r)\big]^{3/2}}{3\pi\hbar^3}. \tag{3.112}$$

This equation, and the boundary conditions stated before, can be conveniently expressed in a dimensionless form in which Z, E, m and $\hbar$ appear only in scale factors. Let us write

$$U(r) = -\frac{Ze^2}{r}\chi, \qquad r = bx, \qquad b = \frac{1}{2}\left(\frac{3\pi}{4}\right)^{2/3}\frac{\hbar^2}{me^2Z^{1/3}} = \frac{0.885a_0}{Z^{1/3}}, \tag{3.113}$$

where $a_0 = \hbar^2/me^2$ is the Bohr radius. After some algebraic manipulation (see problem 14) we find that, with these substitutions, equation (3.112) takes the form

$$x^{1/2}\frac{\mathrm{d}^2\chi}{\mathrm{d}x^2} = \chi^{3/2}, \qquad \chi \geq 0. \tag{3.114}$$

From our previous discussion it is clear that, for neutral atoms, the appropriate boundary conditions to find acceptable solutions are

$$\chi \to 1 \quad \text{as} \quad x \to 0, \qquad \text{and} \qquad \chi \to 0 \quad \text{as} \quad x \to \infty. \tag{3.115}$$

The Thomas-Fermi equation (3.114) is a *universal* equation that does not depend on any physical constants, because they have been absorbed in the change of variables performed. It is a second-order non-linear differential equation that is solved by numerical methods. This provides a *universal function*, $\chi(x)$, for all neutral atoms. $\chi(x)$ is monotonically decreasing and it can be shown that its asymptotic form at large x is given by the function $144x^{-3}$. At $x = 0$ the slope is $\chi'(0) = -1.588$, so in the neighborhood of the origin it can be written as

$$\chi(x) \approx 1 - 1.588x. \tag{3.116}$$

The potential at low values of r can be computed using this $\chi(x)$,

$$U(r) = -\frac{Ze^2}{r}\chi \approx -\frac{Ze^2}{r}(1 - 1.588x) = e^2\left(-\frac{Z}{r} + \frac{1.588Z}{b}\right) = e^2\left(-\frac{Z}{r} + 1.794\frac{Z^{4/3}}{a_0}\right). \tag{3.117}$$

The first term of this equation is just the pure nuclear attraction while the second one, which is repulsive, is due to the contribution of the electrons.

We can compute also the electronic density, $\rho(r)$, which is given by (see problem 14)

$$\rho(x) = \frac{Z}{4\pi b^3}\left(\frac{\chi}{x}\right)^{3/2}. \tag{3.118}$$

Since $\chi(x) \to 1$ as $x \to 0$ and $\chi(x) \to x^{-3}$ as $x \to \infty$, the electron density diverges as $r^{-3/2}$ when $r \to 0$ and goes like r^{-6} when $r \to \infty$. The correct electron density, however, should remain finite at $r = 0$ and decrease exponentially at large r. Obviously, a real atom does not fulfill the assumptions of the Thomas-Fermi model for all values of r, and that's why we obtain such conclusions. This model is inaccurate at both small r and large r, where the assumptions in which it rests are not valid. It is only applicable for intermediate distances, where in any

case most of the electrons of complex atoms are found. The Thomas–Fermi approximation improves with increasing Z, and might be useful to compute quantities that depend on average electron properties, such as the total energy of the atom, but cannot give accurate predictions for quantities that rely on properties of the outer electrons, such as the ionization potential, for example.

Although in this model the electronic density falls off so slowly at large r that it is not possible to really define a *boundary* for the neutral atom, we can still talk about an atomic radius, $R = r(\alpha)$, as the radius of a sphere centered at the origin and enclosing a given fraction, $(1-\alpha)$, of the Z electrons. This means that

$$4\pi \int_0^R \rho(r)\, r^2 \, \mathrm{d}r = (1-\alpha)Z. \tag{3.119}$$

Applying here the change of variable (3.113), $r = bx$, $r(\alpha) = bx(\alpha)$, and using the Thomas–Fermi equation $x^{1/2}\chi'' = \chi^{3/2}$, together with the boundary condition $\chi(0) = 1$, we can derive the following equation for $x_\alpha \equiv x(\alpha)$ (see problem 15)

$$\chi(x_\alpha) - x_\alpha \chi'(x_\alpha) = \alpha. \tag{3.120}$$

This equation has to be solved using numerical methods. If we adopt the same value of α for all atoms, this becomes a *universal* equation and the value of x_α is the same for all atoms. It can be seen that the atomic radius $R = r(\alpha)$ is proportional to $Z^{-1/3}$.

3.2.6 Hartree–Fock equations.

We consider now the second procedure which will lead us to a determination of the central potentials $U(r_i)$. Let's start by trying to find an approximate solution to the Schrödinger equation with the Hamiltonian $\mathcal{H}_{\mathrm{I}}$

$$\mathcal{H}_{\mathrm{I}} = \sum_{i=1}^{N} \mathcal{H}_0(i) + \sum_{i<j}^{N} \frac{e^2}{r_{ij}}, \qquad \text{where} \qquad \mathcal{H}_0(i) = \frac{p_i^2}{2m} - \frac{Ze^2}{r_i}. \tag{3.121}$$

For the ground state we take as a first approximation a single Slater determinant, written in a shortened notation as

$$\Psi(\lambda_1, \lambda_2, \ldots, \lambda_N) = \frac{1}{\sqrt{N!}} \det\Big[\psi_1(\lambda_1)\psi_2(\lambda_2)\ldots\psi_N(\lambda_N)\Big], \tag{3.122}$$

with the spin orbitals[14] satisfying the orthonormality condition

$$\langle \psi_i(\lambda_j) \,|\, \psi_k(\lambda_j) \rangle = \delta_{ik}, \tag{3.123}$$

and chosen to satisfy the variational principle

$$\delta \langle \Psi \,|\, \mathcal{H} \,|\, \Psi \rangle = 0. \tag{3.124}$$

In other words, we wish to find the wave functions ψ_i that make the functional $\langle \Psi \,|\, \mathcal{H} \,|\, \Psi \rangle$ stationary, under the constraints (3.123). This will be done using the Lagrange multipliers method. We are assuming that the N–electron wave function Ψ, represented by a single Slater determinant and satisfying conditions (3.123) and (3.124), is an approximate solution to the N–electron Schrödinger equation. This is known as the *Hartree–Fock* method. Note that an arbitrary, normalizable, antisymmetric N–electron wave function can only be represented by an infinite sum of Slater determinants, each constructed from a complete orthonormal set of one–electron functions and each multiplied by an appropriate coefficient. Even recognizing

[14] At this point the spin orbitals are not necessarily of the central field type, they could be more general.

this limitation, the Hartree–Fock method turns out to be a very good first step towards the construction of atomic wave functions.

To illustrate the mechanics of the method it is convenient to examine first a two-electron atom. The Hamiltonian is

$$\mathcal{H} = \mathcal{H}_0(1) + \mathcal{H}_0(2) + \frac{e^2}{r_{12}}, \tag{3.125}$$

where

$$\mathcal{H}_0(1) = \frac{p_1^2}{2m} - \frac{Ze^2}{r_1}, \qquad \mathcal{H}_0(2) = \frac{p_2^2}{2m} - \frac{Ze^2}{r_2}. \tag{3.126}$$

Let us write the Slater determinant for the ground state as

$$\Psi = \frac{1}{\sqrt{2}} \begin{vmatrix} a(\lambda_1) & b(\lambda_1) \\ a(\lambda_2) & b(\lambda_2) \end{vmatrix} = \frac{1}{\sqrt{2}} \Big[a(\lambda_1)b(\lambda_2) - b(\lambda_1)a(\lambda_2) \Big], \tag{3.127}$$

where $a(\lambda)$ and $b(\lambda)$ are spin orbitals, while λ_1 and λ_2 stand for the coordinates (space and spin) of electrons 1 and 2, respectively. We will often abbreviate the notation $a(\lambda_1)$ to $a(1)$ etc., so the previous expression is converted into

$$\Psi = \frac{1}{\sqrt{2}} \Big[a(1)b(2) - b(1)a(2) \Big]. \tag{3.128}$$

If we are referring to an arbitrary electron, the spin orbitals may be written just as a and b, without arguments. The orthonormality condition (3.123) can then be expressed as

$$\langle a(\lambda_j) \,|\, b(\lambda_j) \rangle \equiv \langle a(j) \,|\, b(j) \rangle = \langle a \,|\, b \rangle = \delta_{ab}, \tag{3.129}$$

which is equivalent to the following three independent conditions

$$\langle a \,|\, a \rangle = 1, \qquad \langle b \,|\, b \rangle = 1, \qquad \langle a \,|\, b \rangle = \langle b \,|\, a \rangle = 0. \tag{3.130}$$

Our next job is to compute the expectation value of the Hamiltonian in the Ψ state. First, we have

$$\begin{aligned} \langle \Psi \,|\, \mathcal{H}_0(1) \,|\, \Psi \rangle &= \frac{1}{2} \langle a(1)b(2) - b(1)a(2) \,|\, \mathcal{H}_0(1) \,|\, a(1)b(2) - b(1)a(2) \rangle \\ &= \frac{1}{2} \langle a(1) \,|\, \mathcal{H}_0(1) \,|\, a(1) \rangle \underbrace{\langle b(2) \,|\, b(2) \rangle}_{1} + \frac{1}{2} \langle b(1) \,|\, \mathcal{H}_0(1) \,|\, b(1) \rangle \underbrace{\langle a(2) \,|\, a(2) \rangle}_{1} \\ &\quad - \frac{1}{2} \langle b(1) \,|\, \mathcal{H}_0(1) \,|\, a(1) \rangle \underbrace{\langle a(2) \,|\, b(2) \rangle}_{0} - \frac{1}{2} \langle a(1) \,|\, \mathcal{H}_0(1) \,|\, b(1) \rangle \underbrace{\langle b(2) \,|\, a(2) \rangle}_{0} \\ &= \frac{1}{2} \langle a(1) \,|\, \mathcal{H}_0(1) \,|\, a(1) \rangle + \frac{1}{2} \langle b(1) \,|\, \mathcal{H}_0(1) \,|\, b(1) \rangle . \end{aligned} \tag{3.131}$$

Similarly

$$\langle \Psi \,|\, \mathcal{H}_0(2) \,|\, \Psi \rangle = \frac{1}{2} \langle a(2) \,|\, \mathcal{H}_0(2) \,|\, a(2) \rangle + \frac{1}{2} \langle b(2) \,|\, \mathcal{H}_0(2) \,|\, b(2) \rangle . \tag{3.132}$$

But of course $\langle a(1) \,|\, \mathcal{H}_0(1) \,|\, a(1) \rangle = \langle a(2) \,|\, \mathcal{H}_0(2) \,|\, a(2) \rangle$, etc. So

$$\Big\langle \Psi \Big| \Big(\mathcal{H}_0(1) + \mathcal{H}_0(2) \Big) \Big| \Psi \Big\rangle = \langle \Psi \,|\, \mathcal{H}_0 \,|\, \Psi \rangle = \langle a \,|\, \mathcal{H}_0 \,|\, a \rangle + \langle b \,|\, \mathcal{H}_0 \,|\, b \rangle \tag{3.133}$$

Now let's go for the expectation value of the $\mathcal{V} \equiv e^2/r_{12}$ term. To shorten even more the notation, we can afford to omit the arguments in the orbitals since they appear always in the same order (first 1 to the left, and then 2 to the right, as in equation (3.128)). Therefore

$$\begin{aligned} \langle \Psi \,|\, \mathcal{V} \,|\, \Psi \rangle &= \frac{1}{2} \langle ab - ba \,|\, \mathcal{V} \,|\, ab - ba \rangle \\ &= \frac{1}{2} \langle ab \,|\, \mathcal{V} \,|\, ab \rangle + \frac{1}{2} \langle ba \,|\, \mathcal{V} \,|\, ba \rangle - \frac{1}{2} \langle ab \,|\, \mathcal{V} \,|\, ba \rangle - \frac{1}{2} \langle ba \,|\, \mathcal{V} \,|\, ab \rangle . \end{aligned} \tag{3.134}$$

But from the symmetry of our problem it is clear that $\langle ab\,|\,\mathcal{V}\,|\,ab\rangle = \langle ba\,|\,\mathcal{V}\,|\,ba\rangle$, because it doesn't really matter if we choose the electron 1 to be in the orbital a and the electron 2 in orbital b, or the other way around. Also, $\langle ab\,|\,\mathcal{V}\,|\,ba\rangle = \langle ba\,|\,\mathcal{V}\,|\,ab\rangle$ for the same reason. Therefore

$$\langle \Psi\,|\,\mathcal{V}\,|\,\Psi\rangle = \langle ab\,|\,\mathcal{V}\,|\,ab\rangle - \langle ab\,|\,\mathcal{V}\,|\,ba\rangle\,. \tag{3.135}$$

Collecting everything together,

$$\langle \Psi\,|\,\mathcal{H}\,|\,\Psi\rangle = \langle a\,|\,\mathcal{H}_0\,|\,a\rangle + \langle b\,|\,\mathcal{H}_0\,|\,b\rangle + \left\langle a(1)b(2)\,\middle|\,\left(\frac{e^2}{r_{12}}\right)\,\middle|\,a(1)b(2)\right\rangle - \left\langle a(1)b(2)\,\middle|\,\left(\frac{e^2}{r_{12}}\right)\,\middle|\,b(1)a(2)\right\rangle. \tag{3.136}$$

The implementation of the variational equation $\delta\,\langle \Psi\,|\,\mathcal{H}\,|\,\Psi\rangle = 0$, with the orthonormality conditions (3.130), can be done using the Lagrange multipliers technique (also called variational parameters). In principle, four such variational parameters would be needed, $\lambda_{aa}, \lambda_{bb}, \lambda_{ab}, \lambda_{ba}$, with $\lambda^*_{ab} = \lambda_{ba}$. However, we will assume that our starting point is already a set of orthogonal spin orbitals,[15] such that $\langle a|b\rangle = 0$ and $\lambda_{ab} = \lambda_{ba} = 0$, therefore the variational equation to be solved is

$$\delta\,\langle \Psi\,|\,\mathcal{H}\,|\,\Psi\rangle = \delta\Bigg[\langle a\,|\,\mathcal{H}_0\,|\,a\rangle + \langle b\,|\,\mathcal{H}_0\,|\,b\rangle + \left\langle a(1)b(2)\,\middle|\,\left(\frac{e^2}{r_{12}}\right)\,\middle|\,a(1)b(2)\right\rangle - \left\langle a(1)b(2)\,\middle|\,\left(\frac{e^2}{r_{12}}\right)\,\middle|\,b(1)a(2)\right\rangle - \lambda_{aa}\,\langle a(1)\,|\,a(1)\rangle - \lambda_{bb}\,\langle b(1)\,|\,b(1)\rangle\Bigg] = 0. \tag{3.137}$$

Let's consider first an arbitrary variation, δa, on a. Using our extra–compact notation, the condition to be satisfied can be written as

$$\langle \delta a\,|\,\mathcal{H}_0\,|\,a\rangle + \langle a\,|\,\mathcal{H}_0\,|\,\delta a\rangle + \langle (\delta a)b\,|\,\mathcal{V}\,|\,ab\rangle + \langle ab\,|\,\mathcal{V}\,|\,(\delta a)b\rangle - \langle (\delta a)b\,|\,\mathcal{V}\,|\,ba\rangle - \langle ab\,|\,\mathcal{V}\,|\,b\delta a\rangle = \lambda_{aa}\,\langle \delta a\,|\,a\rangle + \lambda_{aa}\,\langle a\,|\,\delta a\rangle\,. \tag{3.138}$$

Although the variations $|\delta a\rangle$ and $\langle \delta a|$ are not independent, it can be shown (see appendix B.4) that it is sufficient to impose the condition

$$\langle \delta a\,|\,\mathcal{H}_0\,|\,a\rangle + \langle (\delta a)b\,|\,\mathcal{V}\,|\,ab\rangle - \langle (\delta a)b\,|\,\mathcal{V}\,|\,ba\rangle = \lambda_{aa}\,\langle \delta a\,|\,a\rangle\,. \tag{3.139}$$

Expanding now the notation,

$$\langle \delta a(1)\,|\,\mathcal{H}_0\,|\,a(1)\rangle + \left\langle \delta a(1)b(2)\,\middle|\,\left(\frac{e^2}{r_{12}}\right)\,\middle|\,a(1)b(2)\right\rangle - \left\langle \delta a(1)b(2)\,\middle|\,\left(\frac{e^2}{r_{12}}\right)\,\middle|\,b(1)a(2)\right\rangle = \lambda_{aa}\,\langle \delta a(1)\,|\,a(1)\rangle\,. \tag{3.140}$$

This equation is verified for *every* possible δa only if the following one holds

$$\mathcal{H}_0(1)a(1) + \left[\int b^*(2)\left(\frac{e^2}{r_{12}}\right)b(2)\,d\tau_2\right]a(1) - \left[\int b^*(2)\left(\frac{e^2}{r_{12}}\right)a(2)\,d\tau_2\right]b(1) \;=\; \lambda_{aa}a(1). \tag{3.141}$$

[15] There may be special situations where it is not possible to start with an orthogonal set of spin orbitals, but even in such cases the off-diagonal terms are small and can often be neglected.

A similar equation follows for $b(1)$ when performing an arbitrary variation δb. These expressions can be generalized to an N-electron system, giving the following equation for a single spin orbital

$$\mathcal{H}_0(1)\psi_i(1) + \left[\sum_{j\neq i}\int \psi_j^*(2)\left(\frac{e^2}{r_{12}}\right)\psi_j(2)\,d\tau_2\right]\psi_i(1) - \sum_{j\neq i}\left[\int \psi_j^*(2)\left(\frac{e^2}{r_{12}}\right)\psi_i(2)\,d\tau_2\right]\psi_j(1) = \varepsilon_i\psi_i(1), \quad (3.142)$$

where the summation over j extends to the N occupied spin orbitals and the Lagrange multipliers have been written as ε_i.

Before proceeding, let us make a short digression and write the previous equation *without the second sum*, and only for the spatial part of the orbitals, ϕ, forgetting about the spin. We get

$$\left[-\frac{\hbar^2}{2m}\nabla_i^2 - \frac{Ze^2}{r_i} + \sum_{j\neq i}\int |\phi_j(\mathbf{r}_j)|^2\left(\frac{e^2}{r_{ij}}\right)\mathrm{d}^3\mathbf{r}_j\right]\phi_i(\mathbf{r}_i) = \varepsilon_i\phi_i(\mathbf{r}_i). \quad (3.143)$$

This is the original equation of the Hartree method, before Fock's generalization introducing the Slater determinant to take into account the antisymmetrization of the wavefunctions. There are Z such equations, one for every electron of the atom. As stated before, this equation is the result of assuming that each electron moves in a central field that can be calculated from the nuclear potential and the wave functions of all the other electrons, by taking the charge density associated with an electron to be $-e$ times its position probability density. We have a system of Z simultaneous nonlinear integrodifferential equations for the Z functions $\phi_i(\mathbf{r}_i)$, which is of course impossible to solve directly, so a method of successive approximations is followed. One takes an approximate potential energy to begin with, solves the system for the wavefunctions, new potentials for each electron are computed with these new functions and they are plugged into the system of equations to get new functions and new potentials. This procedure is repeated until the potentials are self-consistent to the desired level of accuracy. This is, in a nutshell, how the so-called self-consistent method works.

We stop the digression here and go back to equation (3.142). Note that the restriction to $j \neq i$ can be dropped because if we set $j = i$ in the first sum on the left we have

$$\left[\int \psi_j^*(2)\left(\frac{e^2}{r_{12}}\right)\psi_j(2)\,d\tau_2\right]\psi_j(1), \quad (3.144)$$

but exactly the same term will appear in the second sum, with a negative sign. This cancellation leads to say that a particle in a spin orbital ψ_i does not *react with itself*, or that the *self energy* of such a particle is zero. We can then write the Hartree-Fock equations as

$$\mathcal{H}_0(1)\psi_i(1) + \left[\sum_{j=1}^{N}\int \psi_j^*(2)\left(\frac{e^2}{r_{12}}\right)\psi_j(2)\,d\tau_2\right]\psi_i(1) - \sum_{j=1}^{N}\left[\int \psi_j^*(2)\left(\frac{e^2}{r_{12}}\right)\psi_i(2)\,d\tau_2\right]\psi_j(1) = \varepsilon_i\psi_i(1). \quad (3.145)$$

The so-called Coulomb and exchange operators, $J_j(1)$ and $K_j(1)$, are usually defined so that we can write the Hartree-Fock equations in this way

$$\left[\mathcal{H}_0(1) + \sum_j\Big[J_j(1) - K_j(1)\Big]\right]\psi_i(1) = \varepsilon_i\psi_i(1), \quad (3.146)$$

$$J_j(1)\psi_i(1) = \left[\int \psi_j^*(2)\left(\frac{e^2}{r_{12}}\right)\psi_j(2)\,d\tau_2\right]\psi_i(1), \tag{3.147}$$

$$K_j(1)\psi_i(1) = \left[\int \psi_j^*(2)\left(\frac{e^2}{r_{12}}\right)\psi_i(2)\,d\tau_2\right]\psi_j(1). \tag{3.148}$$

If we now define the Fock operator $F(1)$ doing

$$F(1) = \mathcal{H}_0(1) + V(1), \qquad \text{with} \qquad V(1) = \sum_j \Big[J_j(1) - K_j(1)\Big], \tag{3.149}$$

then the Hartree-Fock equation takes the form of an eigenvalue equation,

$$F(1)\psi_i(1) = \varepsilon_i\psi_i(1). \tag{3.150}$$

Note that the Lagrange multiplier ε_i has, through the Hartree-Fock equations, acquired the meaning of a one-electron energy eigenvalue.

In the Hartree-Fock equations the spin orbitals depend both on space coordinates and spin coordinates,

$$\psi_{n\ell m_\ell m_s}(\mathbf{r}, m_s) = \phi_{n\ell m_\ell}(\mathbf{r})\,\chi(m_s), \qquad a(\lambda_i) = \phi_a(\mathbf{r}_i)\chi_i^a(m_s), \tag{3.151}$$

$$\chi(m_s = 1/2) = \left|\frac{1}{2}, +\frac{1}{2}\right\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \chi(m_s = -1/2) = \left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{3.152}$$

However, for computational purposes the spin and space parts need to be separated. For a one-electron operator f which acts only on space coordinates we would have

$$\langle a \,|\, f \,|\, b\rangle = \left\langle \phi_a\chi^a \,\middle|\, f \,\middle|\, \phi_b\chi^b \right\rangle = \langle \phi_a \,|\, f \,|\, \phi_b\rangle \left\langle \chi^a \,\middle|\, \chi^b \right\rangle = \langle \phi_a \,|\, f \,|\, \phi_b\rangle\ \delta(m_s^a, m_s^b), \tag{3.153}$$

where, of course, $\delta(m_s^a, m_s^b) = 1$ if $m_s^a = m_s^b$ and $\delta(m_s^a, m_s^b) = 0$ otherwise. For a two–electron operator like $g_{12} = e^2/r_{12}$,

$$\begin{aligned}\langle a(1)b(2) \,|\, g_{12} \,|\, c(1)d(2)\rangle &= \Big(\,\langle \phi_a(1)\phi_b(2) \,|\, g_{12} \,|\, \phi_c(1)\phi_d(2)\rangle\,\Big)\Big(\,\langle \chi_1^a(m_s^a) \,|\, \chi_1^c(m_s^c)\rangle\,\Big)\Big(\,\left\langle \chi_2^b(m_s^b) \,\middle|\, \chi_2^d(m_s^d)\right\rangle\Big) \\ &= \Big(\,\langle \phi_a(1)\phi_b(2) \,|\, g_{12} \,|\, \phi_c(1)\phi_d(2)\rangle\,\Big)\delta(m_s^a, m_s^c)\,\delta(m_s^b, m_s^d).\end{aligned} \tag{3.154}$$

If $a = c$ and $b = d$ this expression reduces to

$$\begin{aligned}\langle a(1)b(2) \,|\, g_{12} \,|\, a(1)b(2)\rangle &= \langle \phi_a(1)\phi_b(2) \,|\, g_{12} \,|\, \phi_a(1)\phi_b(2)\rangle\ \delta(m_s^a, m_s^a)\,\delta(m_s^b, m_s^b) \\ &= \langle \phi_a(1)\phi_b(2) \,|\, g_{12} \,|\, \phi_a(1)\phi_b(2)\rangle\,.\end{aligned} \tag{3.155}$$

Whereas if $a = d$ and $b = c$,

$$\begin{aligned}\langle a(1)b(2) \,|\, g_{12} \,|\, b(1)a(2)\rangle &= \langle \phi_a(1)\phi_b(2) \,|\, g_{12} \,|\, \phi_b(1)\phi_a(2)\rangle\ \delta(m_s^a, m_s^b)\,\delta(m_s^b, m_s^a) \\ &= \langle \phi_a(1)\phi_b(2) \,|\, g_{12} \,|\, \phi_b(1)\phi_a(2)\rangle\ \delta(m_s^a, m_s^b).\end{aligned} \tag{3.156}$$

Therefore the Hartree-Fock equation for a single spatial orbital becomes

$$\begin{aligned}\mathcal{H}_0(1)\phi_i(1) + \left[\sum_{j=1}^{N}\int \phi_j^*(2)\left(\frac{e^2}{r_{12}}\right)\phi_j(2)\,d\mathbf{r_2}\right]\phi_i(1) \\ - \sum_{j=1}^{N}\left[\delta(m_s^i, m_s^j)\int \phi_j^*(2)\left(\frac{e^2}{r_{12}}\right)\phi_i(2)\,d\mathbf{r_2}\right]\phi_j(1) = \varepsilon_i\phi_i(1),\end{aligned} \tag{3.157}$$

where the ϕ are functions of the space coordinates only. The Coulomb term

$$\sum_{j=1}^{N} J_j(1) = \sum_{j=1}^{N} \int \phi_j^*(2) \left(\frac{e^2}{r_{12}}\right) \phi_j(2)\, d\mathbf{r_2} \tag{3.158}$$

admits now an easy interpretation. The quantity $e\phi_j^*(2)\phi_j(2)$ is the charge density of electron 2 *smeared* out over the orbital ϕ_j. The integral is then the electrostatic repulsion potential due to electron 2 when the position of the latter is averaged over the orbital ϕ_j. When this is summed over all electrons we get the total electrostatic repulsion. The exchange term

$$-\sum_{j=1}^{N} \left[\delta(m_s^i, m_s^j) \int \phi_j^*(2) \left(\frac{e^2}{r_{12}}\right) \phi_i(2)\, d\mathbf{r_2}\right] \phi_j(1), \tag{3.159}$$

has no classical analog and is more difficult to interpret. Formally it arises from the antisymmetry requirement on the total wave function Ψ, and demands the spin projections of electrons i and j to be parallel. If Ψ had been written as a simple product of one–electron orbitals (also known as a Hartree product) rather than as a Slater determinant, the exchange term would be absent. Further analysis of this term can show that the total exchange charge density corresponds to the charge of a single electron (a unit charge). Because of the negative sign in front of the exchange term, it is then often said that an *exchange hole* (or *Fermi hole*) is surrounding any electron in an atom. This may be visualized as a spherical region surrounding a given electron, from which electronic charge of the same spin as that of the given electron has been removed. The total charge that has been removed is equivalent to the charge of one electron. This is a manifestation of the Pauli exclusion principle.

In summary, by applying the variational method based on a single Slater determinant, constructed from a set of orthonormal one–electron spin orbitals, we have obtained the Hartree-Fock equations (3.145). These are coupled nonlinear integro-differential equations for the one-electron spin orbitals and consist entirely of one-electron operators. We can regard a Hartree-Fock equation as the wave equation for a single electron moving in the field produced by the nucleus and the average field of the remaining electrons. This is exactly the goal of the central field approximation. The average field of the $N-1$ electrons is represented by the Coulomb and exchange operators, which replace the operator e^2/r_{ij} in the exact Hamiltonian. This central field is made up by the field produced by electrons of opposite spin, and the field produced by electrons of the same spin but outside the Fermi hole. In addition, it should be noted that the Hartree-Fock method only provides an upper bound to the energy of a state of a given symmetry because it is based on a variational principle.

Properties of the Hartree-Fock solutions.

It is not difficult to prove that the spin orbitals that satisfy the Hartree-Fock equations are orthogonal,

$$\langle \psi_k | \psi_i \rangle = 0, \qquad \varepsilon_i \neq \varepsilon_k \tag{3.160}$$

where ψ_i and ψ_k are two solutions of the Hartree-Fock equations,

$$\mathcal{H}_0(1)\psi_i(1) + \left[\sum_{j=1}^{N} \int \psi_j^*(2) \left(\frac{e^2}{r_{12}}\right) \psi_j(2)\, d\tau_2\right] \psi_i(1) - \sum_{j=1}^{N} \left[\int \psi_j^*(2) \left(\frac{e^2}{r_{12}}\right) \psi_i(2)\, d\tau_2\right] \psi_j(1) = \varepsilon_i \psi_i(1), \tag{3.161}$$

and

$$\mathcal{H}_0(1)\psi_k(1) + \left[\sum_{j=1}^{N}\int \psi_j^*(2)\left(\frac{e^2}{r_{12}}\right)\psi_j(2)\,d\tau_2\right]\psi_k(1) - \sum_{j=1}^{N}\left[\int \psi_j^*(2)\left(\frac{e^2}{r_{12}}\right)\psi_k(2)\,d\tau_2\right]\psi_j(1) = \varepsilon_k\psi_k(1). \quad (3.162)$$

The total energy of the N-electron system is

$$E = \langle\Psi|\mathcal{H}|\Psi\rangle\,, \quad (3.163)$$

which can be written as (we leave the algebraic details out),

$$E = \sum_i I_i + \sum_{i<j}\Big[J(i,j) - K(i,j)\Big] = \sum_i I_i + \frac{1}{2}\sum_{i,j}\Big[J(i,j) - K(i,j)\Big], \quad (3.164)$$

where

$$I_i = \langle\psi_i|\mathcal{H}_0|\psi_i\rangle = \langle\phi_i|\mathcal{H}_0|\phi_i\rangle\,, \quad (3.165)$$

$$J(i,j) = \left\langle\psi_i(1)\psi_j(2)\left|\left(\frac{e^2}{r_{12}}\right)\right|\psi_i(1)\psi_j(2)\right\rangle = \left\langle\phi_i(1)\phi_j(2)\left|\left(\frac{e^2}{r_{12}}\right)\right|\phi_i(1)\phi_j(2)\right\rangle, \quad (3.166)$$

$$K(i,j) = \left\langle\psi_i(1)\psi_j(2)\left|\left(\frac{e^2}{r_{12}}\right)\right|\psi_j(1)\psi_i 2)\right\rangle = \delta(m_s^i, m_s^j)\left\langle\phi_i(1)\phi_j(2)\left|\left(\frac{e^2}{r_{12}}\right)\right|\phi_j(1)\phi_i(2)\right\rangle. \quad (3.167)$$

Also, when equation (3.161) is multiplied by ψ_i^* and integrated we obtain the following explicit expression for ε_i,

$$\varepsilon_i = I_i + \sum_j\Big[J(i,j) - K(i,j)\Big]. \quad (3.168)$$

It is possible to sum this over all electrons to get

$$\sum_i \varepsilon_i = \sum_i I_i + \sum_{i,j}\Big[J(i,j) - K(i,j)\Big] = \sum_i I_i + 2\sum_{i<j}\Big[J(i,j) - K(i,j)\Big]. \quad (3.169)$$

Inserting this result into the expression for the total energy we get

$$E = \sum_i \varepsilon_i - \sum_{i<j}\Big[J(i,j) - K(i,j)\Big]. \quad (3.170)$$

This clearly tells us that the total energy is not the sum of all electron energies ε_i. It can be shown that the energy ε_i is to be interpreted as the energy required to remove an electron from the ith orbital (Koopman's theorem). Indeed, let's consider the energy difference between a system of $N+1$ electrons and a system of N electrons. Assuming we have a set of orbitals labeled $j = 1, 2, 3, \ldots, N, i$, the energy of the $N+1$ system is (applying directly equation (3.164))

$$E_{N+1} = \sum_{j=1}^{N} I_j + \sum_{j<k}^{N}\Big[J(j,k) - K(j,k)\Big] + I_i + \sum_{k=1}^{N}\Big[J(i,k) - K(i,k)\Big], \quad (3.171)$$

and

$$E_{N+1} - E_N = I_i + \sum_{k=1}^{N}\Big[J(i,k) - K(i,k)\Big] = \varepsilon_i. \quad (3.172)$$

However, this interpretation is not completely rigorous because we have ignored that the orbitals of the ion cannot be the same as those of the neutral atom. After the removal of an electron, it takes place a readjustment of the electron orbitals (the charge of the nucleus remains the same but the total electronic charge decreases by one unit).

For a closed shell determined by the quantum numbers n and ℓ we have a system of $2(2\ell+1)$ electrons. In this case it can be proved that the Hartree-Fock one-electron orbitals are, rigorously, solutions of a central field problem. In other words, each electron in a closed shell moves in a spherically symmetric potential. Without going into the details of the proof, let's say that this result follows from the fact that in a closed shell all magnetic substates are occupied, enabling to carry out summations over all these substates and getting at the end a spherically symmetric field.

Since the Hartree–Fock equations constitute a system of coupled, nonlinear, integro-differential equations, there is no direct method to solve them and we must use iterative numerical computational methods. In fact, each Hartree–Fock equation requires prior knowledge of all other orbitals. We need therefore to select an initial set of spin orbitals as a first step of a so-called *self–consistent* method. These initial orbitals are used to calculate the Coulomb and exchange operators J_j and K_j and hence the Fock operator F. With this approximate Fock operator a new improved set of orbitals are calculated by solving the eigenvalue equations $F(1)\psi_i(1) = \varepsilon_i\psi_i(1)$. The process is continued until the difference between two sets of orbitals in successive iterations of computation is sufficiently small to achieve the desired accuracy. This is the general method to solve the equations, and is known as the Hartree-Fock-Self-Consistent Field method.

It has to be noted that the Hartree-Fock method does not take completely into account the correlation between the electrons. This is known as the *correlation error*. In a real atom the Coulomb repulsion between electrons, whose spins may be either parallel or anti-parallel, prevents any two electrons from approaching arbitrarily close to one another. This means that the positions of all electrons, regardless of their spin, are correlated, and the N-electron wave function cannot really be written as a product of single-electron wave functions (or linear combinations of them). To give a simple specific case, the relation $\Psi(\lambda_1, \lambda_2) \approx \psi_1(\lambda_1)\psi_2(\lambda_2)$ holds only approximately. Using another terminology, one says that all electrons are surrounded by a *Coulomb hole*. An exact N–electron wave function would automatically give a correct description of the Coulomb hole, but the Hartree-Fock wave function is not exact and the correlation is not completely accounted for. It has to be admitted, however, that the correlation error for electrons with parallel spin is not the same as for electrons with opposite spin. This is due to the antisymmetry property of the Hartree-Fock wave function, which correlates the positions of electrons with parallel spins and prevents them from occupying the same spatial orbital, but has no effect on electrons with opposite spin. This feature was described by the *Fermi hole*, which is almost equivalent to the Coulomb hole for electrons with parallel spin. The largest source of error in the Hartree-Fock method arises from the absence of correlation among electrons with opposite spin. Estimates of correlation errors give $-1.1\,\text{eV}$ for He, $-11\,\text{eV}$ for Ne and $-21.5\,\text{eV}$ for Ar, for example.

To expand an N–electron antisymmetric wave function one needs a complete set of Slater determinants. This set consists of all possible determinants containing N spin orbitals selected from a complete set of spin orbitals. Each Slater determinant constructed in this way corresponds to a single electronic configuration, therefore such an expansion of an antisymmetric wave function is also known as a superposition of configurations. If a single configuration is employed, which means the infinite series of Slater determinants is truncated after the first term, and the spin orbitals are optimized with respect to the total energy, we have the ordinary Hartree–Fock approximation. Improved results can be achieved by increasing the number of configurations in order to get closer to the exact N–electron antisymmetric wave function. This method for obtaining improved eigenfunctions and eigenvalues for the N–electron problem is known as *configuration interaction*. In this method, N–electron wave functions are obtained by

taking linear combinations of Slater determinants,

$$\Psi(1,2,\ldots,N) = \sum_i c_i \psi_i, \tag{3.173}$$

where the c_i are coefficients and the ψ_i are Slater determinants which differ in the choice of spin orbitals that the electrons are assumed to occupy, and hence correspond to different electronic configurations. The coefficients are obtained by minimizing the energy. This leads to a system of linear equations and a secular determinant,

$$\sum_j (H_{ij} - E\delta_{ij})c_j = 0, \qquad H_{ij} = \int \psi_i^* \mathcal{H} \psi_j \, d\tau, \qquad |H_{ij} - E\delta_{ij}| = 0, \tag{3.174}$$

from which the required coefficients are obtained.

3.3 Molecules.

Molecules are considerably more difficult to study from first principles than atoms and a new set of approximations are needed. A key point to note is that electrons have a much smaller mass than nuclei, while the forces to which both electrons and nuclei are submitted are of comparable magnitude. As a result, the motion of the nuclei is much slower than that of the electrons and, to a first good approximation, the nuclei can be considered to occupy nearly fixed positions within the molecule. This is confirmed by observing the diffraction pattern in neutron scattering by molecules. In this type of experiments neutrons do not interact with the atomic electrons, but they do interact strongly with the nuclei within the molecule. In the simplest neutral diatomic molecule, H_2, the equilibrium spacing of the two protons is 0.74 Å. In the O_2 molecule the internuclear spacing is 1.21 Å.

From X-ray diffraction observations and molecular spectra we know that, when atoms associate to form molecules, the tightly bound inner shells of electrons are nearly undisturbed and remain localized about each nucleus. The outer electrons are distributed throughout the molecule and it is the charge distribution of these valence electrons that provides the binding force.

3.3.1 Born–Oppenheimer approximation.

For a system of electrons and nuclei we can build the Hamiltonian

$$\mathcal{H} = \sum_\alpha \frac{P_\alpha^2}{2M_\alpha} + \sum_i \frac{p_i^2}{2m} + V(Q_\alpha, q_i), \tag{3.175}$$

where $P_\alpha^2/2M_\alpha$ is the kinetic energy operator for a nucleus of mass M_α, $P_i^2/2m$ is the kinetic energy operator for the ith electron of mass m, Q_α is a set of nuclear coordinates, q_i is a set of electronic coordinates, and

$$V(Q_\alpha, q_i) = \sum_{i<j} \frac{e^2}{r_{ij}} + \sum_{\alpha<\beta} \frac{(Z_\alpha e)(Z_\beta e)}{r_{\alpha\beta}} - \sum_{i,\alpha} \frac{(Z_\alpha e)e}{r_{i\alpha}} \tag{3.176}$$

is the potential energy of the entire system of electrons and nuclei. The first term is the sum of the Coulomb interactions between all pairs of electrons, the second is the same for all pairs of nuclei, and the third term is the sum of the Coulomb interactions between nuclei and electrons. The Born–Oppenheimer procedure to find an approximate solution to

$$\mathcal{H}\Psi(Q_\alpha, q_i) = E\,\Psi(Q_\alpha, q_i), \tag{3.177}$$

can be summarized in five steps:

1. Assume the nuclei occupy fixed positions described by the coordinates Q'_α. The nuclei kinetic energy in the Hamiltonian (3.175) can then be neglected.

2. Under that assumption, solve the following Schrödinger equation,

$$\left[\sum_i \frac{p_i^2}{2m} + V(Q'_\alpha, q_i)\right] \psi(Q'_\alpha, q_i) = E'(Q'_\alpha)\, \psi(Q'_\alpha, q_i), \tag{3.178}$$

   where $\psi(Q'_\alpha, q_i)$ is an electronic wave function and $E'(Q'_\alpha)$ is the corresponding electronic energy eigenvalue when the nuclei are fixed at coordinates $Q_\alpha = Q'_\alpha$. If the nuclear positions are assigned a different set of values, Q''_α, the potential energy is altered to $V(Q''_\alpha, q_i)$ and a new set of electronic energy eigenvalues and eigenfunctions can be obtained. It is possible, at least in principle, to repeat the computation for a sufficiently large series of fixed nuclear positions to describe the dependence of the wave functions and the energies on both Q_α and q_i.

3. Assume that the total wavefunction can be written as a product of the electronic wavefunction and a so-called *vibrational function* which depends only on the nuclear coordinates,

$$\Psi(Q_\alpha, q_i) = \psi(Q_\alpha, q_i)\, v(Q_\alpha). \tag{3.179}$$

   At this point $v(Q_\alpha)$ is unknown.

4. Insert this assumption in $\mathcal{H}\Psi(Q_\alpha, q_i) = E\,\Psi(Q_\alpha, q_i)$ to get

$$\mathcal{H}\psi(Q_\alpha, q_i)\, v(Q_\alpha) = E\,\psi(Q_\alpha, q_i)\, v(Q_\alpha), \tag{3.180}$$

$$\left[\sum_\alpha \frac{P_\alpha^2}{2M_\alpha} + \sum_i \frac{p_i^2}{2m} + V(Q_\alpha, q_i)\right] \psi(Q_\alpha, q_i)\, v(Q_\alpha) = E\,\psi(Q_\alpha, q_i)\, v(Q_\alpha) \tag{3.181}$$

5. Finally, assume that the electronic wave function, $\psi(Q_\alpha, q_i)$, is at first order insensitive to changes in the nuclear positions and momenta, being able to adjust quasi–statically to the nuclear motion (this is also known as *adiabatic approximation*). This means we can write

$$\frac{P_\alpha^2}{2M_\alpha}\, \psi(Q_\alpha, q_i)\, v(Q_\alpha) = \psi(Q_\alpha, q_i)\, \frac{P_\alpha^2}{2M_\alpha} v(Q_\alpha). \tag{3.182}$$

   To see this in more detail, note that if $\psi(Q_\alpha, q_i)$ does not change with respect to variations of nuclear positions and momenta, then $\boldsymbol{\nabla}_\alpha \psi(Q_\alpha, q_i) \approx 0$ and $\nabla_\alpha^2 \psi(Q_\alpha, q_i) \approx 0$. Therefore[16]

$$\begin{aligned}
&\frac{P_\alpha^2}{2M_\alpha}\, \psi(Q_\alpha, q_i)\, v(Q_\alpha) = -\frac{\hbar^2}{2M_\alpha} \nabla_\alpha^2 \Big(\psi(Q_\alpha, q_i)\, v(Q_\alpha)\Big) \\
&= -\frac{\hbar^2}{2M_\alpha}\left[\psi(Q_\alpha, q_i)\, \nabla_\alpha^2 v(Q_\alpha) + 2\left(\boldsymbol{\nabla}_\alpha \psi(Q_\alpha, q_i)\right) \cdot \left(\boldsymbol{\nabla}_\alpha v(Q_\alpha)\right) + v(Q_\alpha) \nabla_\alpha^2 \psi(Q_\alpha, q_i)\right] \\
&\qquad \approx -\frac{\hbar^2}{2M_\alpha} \psi(Q_\alpha, q_i)\, \nabla_\alpha^2 v(Q_\alpha) = \psi(Q_\alpha, q_i)\, \frac{P_\alpha^2}{2M_\alpha} v(Q_\alpha).
\end{aligned} \tag{3.183}$$

The approximation in step 5 is the central assumption of the Born–Oppenheimer approach, and relies on the fact that $m/M_\alpha \ll 1$. From a classical point of view, we would say that nuclear

[16] Use the vector derivative rule $\nabla^2(fg) = (\nabla^2 f)\, g + 2\,(\boldsymbol{\nabla} f) \cdot (\boldsymbol{\nabla} g) + f\,(\nabla^2 g)$.

velocities are small in comparison with electronic velocities. Using (3.182), equation (3.181) is then transformed into

$$\psi(Q_\alpha,q_i)\sum_\alpha \frac{P_\alpha^2}{2M_\alpha}v(Q_\alpha)+\underbrace{\left[\sum_i \frac{p_i^2}{2m}+V(Q_\alpha,q_i)\right]\psi(Q_\alpha,q_i)}_{E'\psi(Q_\alpha,q_i)}\,v(Q_\alpha)=E\,\psi(Q_\alpha,q_i)\,v(Q_\alpha),$$

$$\psi(Q_\alpha,q_i)\sum_\alpha \frac{P_\alpha^2}{2M_\alpha}v(Q_\alpha)+E'\psi(Q_\alpha,q_i)\,v(Q_\alpha)=E\,\psi(Q_\alpha,q_i)\,v(Q_\alpha), \tag{3.184}$$

where equation (3.178) has been used for the term indicated by the under–brace symbol. We have then

$$\left[\sum_\alpha \frac{P_\alpha^2}{2M_\alpha}+E'(Q_\alpha)\right]v(Q_\alpha) \;=\; E\,v(Q_\alpha). \tag{3.185}$$

This is an important result. It states that the wave function for the nuclear motion is calculated in an effective potential which is obtained from the energy eigenvalues of a particular electronic state as a function of the nuclear position

Note that the form in which the total wave function is written, $\Psi(Q_\alpha,q_i)=\psi(Q_\alpha,q_i)\,v(Q_\alpha)$, implies that electronic and nuclear motions can be treated separately. Since the electronic motion can be regarded as taking place in a potential field created by stationary nuclei, the electronic wave function may be evaluated with the nuclei in their equilibrium positions and the small displacements of the nuclei about their equilibrium positions will have little effect on the electronic wave function. The energy eigenvalues of a particular electronic state, as a function of nuclear position, provide the effective potential for the nuclear motion, which is then solved as a separate problem.

Let's go back to the Schrödinger equation for the electronic wave function in the field of the fixed nuclei whose coordinates are collectively represented by Q_α,

$$\mathcal{H}_e\psi(Q_\alpha,q_i)=E'(Q_\alpha)\,\psi(Q_\alpha,q_i), \qquad \text{with} \qquad \mathcal{H}_e=\sum_i \frac{p_i^2}{2m}+V(Q_\alpha,q_i). \tag{3.186}$$

Multiplying on the left by $\psi^*(Q_\alpha,q_i)$ and integrating over the electronic coordinates gives

$$E'(Q_\alpha) \;=\; \int \psi^*(Q_\alpha,q_i)\,\mathcal{H}_e\,\psi(Q_\alpha,q_i)\,d\tau_q. \tag{3.187}$$

We are interested now in computing $\partial E'(Q_\alpha)/\partial Q_\alpha$. This is

$$\begin{aligned}\frac{\partial E'(Q_\alpha)}{\partial Q_\alpha}&=\frac{\partial}{\partial Q_\alpha}\left(\int \psi^*(Q_\alpha,q_i)\,\mathcal{H}_e\,\psi(Q_\alpha,q_i)\,d\tau_q\right)\\&=\int \frac{\partial\psi^*(Q_\alpha,q_i)}{\partial Q_\alpha}\,\mathcal{H}_e\,\psi(Q_\alpha,q_i)\,d\tau_q+\int \psi^*(Q_\alpha,q_i)\,\frac{\partial\mathcal{H}_e}{\partial Q_\alpha}\,\psi(Q_\alpha,q_i)\,d\tau_q\\&\qquad+\int \psi^*(Q_\alpha,q_i)\,\mathcal{H}_e\,\frac{\partial\psi(Q_\alpha,q_i)}{\partial Q_\alpha}\,d\tau_q.\end{aligned} \tag{3.188}$$

Since $\mathcal{H}_e$ is Hermitian and $\mathcal{H}_e\,\psi=E'\,\psi$, the third integral may be written as

$$\begin{aligned}\int \psi^*(Q_\alpha,q_i)\,\mathcal{H}_e\,\frac{\partial\psi(Q_\alpha,q_i)}{\partial Q_\alpha}\,d\tau_q \;&=\; \left(\int \frac{\partial\psi^*(Q_\alpha,q_i)}{\partial Q_\alpha}\,\mathcal{H}_e\,\psi(Q_\alpha,q_i)\,d\tau_q\right)^*\\&=\int \frac{\partial\psi(Q_\alpha,q_i)}{\partial Q_\alpha}\,\underbrace{\mathcal{H}_e\,\psi^*(Q_\alpha,q_i)}_{E'\,\psi^*(Q_\alpha,q_i)}\,d\tau_q=\int \frac{\partial\psi(Q_\alpha,q_i)}{\partial Q_\alpha}\,E'\,\psi^*(Q_\alpha,q_i)\,d\tau_q.\end{aligned} \tag{3.189}$$

Using this, we can see that the first and third integrals of the right-hand side of equation (3.188) cancel each other out. Indeed

$$\int \frac{\partial \psi^*}{\partial Q_\alpha}\,\mathcal{H}_e\,\psi\,d\tau_q \;+\; \int \psi^*\,\mathcal{H}_e\,\frac{\partial \psi}{\partial Q_\alpha}\,d\tau_q \;=\; E'\int\left(\frac{\partial \psi^*}{\partial Q_\alpha}\,\psi \;+\; \psi^*\,\frac{\partial \psi}{\partial Q_\alpha}\right)d\tau_q$$
$$=\; E'\int \frac{\partial}{\partial Q_\alpha}(\psi^*\psi)\,d\tau_q \;=\; E'\,\frac{\partial}{\partial Q_\alpha}\int \psi^*\psi\,d\tau_q \;=\; 0. \quad (3.190)$$

Therefore

$$\frac{\partial E'(Q_\alpha)}{\partial Q_\alpha} = \int \psi^*(Q_\alpha, q_i)\,\frac{\partial \mathcal{H}_e}{\partial Q_\alpha}\,\psi(Q_\alpha, q_i)\,d\tau_q. \quad (3.191)$$

But the kinetic energy operator does not depend on Q_α,

$$\frac{\partial \mathcal{H}_e}{\partial Q_\alpha} = \frac{\partial}{\partial Q_\alpha}\left(\sum_i \frac{p_i^2}{2m} + V(Q_\alpha, q_i)\right) = \frac{\partial V(Q_\alpha, q_i)}{\partial Q_\alpha}. \quad (3.192)$$

So we conclude that the derivative of $E'(Q_\alpha)$ with respect to Q_α is

$$\frac{\partial E'(Q_\alpha)}{\partial Q_\alpha} = \int \psi^*(Q_\alpha, q_i)\,\frac{\partial V(Q_\alpha, q_i)}{\partial Q_\alpha}\,\psi(Q_\alpha, q_i)\,d\tau_q. \quad (3.193)$$

This is known as the *Hellmann-Feynman theorem*. It implies that, within the Born-Oppenheimer approximation, the forces acting on nuclei are derivable from the classical potential

$$V(Q_\alpha, q_i) = \sum_{i<j}\frac{e^2}{r_{ij}} + \sum_{\alpha<\beta}\frac{Z_\alpha Z_\beta e^2}{r_{\alpha\beta}} - \sum_{i\alpha}\frac{Z_\alpha e^2}{r_{i\alpha}} \quad (3.194)$$

and the electronic wave function. If we write a force component F_α as

$$F_\alpha = -\frac{\partial E'(Q_\alpha)}{\partial Q_\alpha}, \quad (3.195)$$

then F_α can be obtained from the classical expression for

$$-\frac{\partial V(Q_\alpha, q_i)}{\partial Q_\alpha} \quad (3.196)$$

averaged over the electronic wave function, as explicitly shown in equation (3.193).

3.3.2 Molecular orbitals and the self–consistent field method.

The Born–Oppenheimer method leaves us with the problem of finding solutions for the electronic and nuclear wave functions. Remember that, in the atomic case, multi-electron wave functions are constructed from orbitals which are populated by electrons as allowed by the Pauli principle. For molecules, it is possible to adopt a similar approach in which the multi-electron molecular wave functions are constructed from a set of *molecular orbitals* which have the following properties

- A molecular orbital is an eigenfunction of a one-electron Hamiltonian consisting of one-electron operators. A molecular orbital is then a wave function that depends on the coordinates of a single electron.
- A molecular orbital can extend over any number of atoms in a molecule. This implies that an electron in a molecular orbital is not localized on a particular atom, but may have a non–vanishing probability density in various parts of the molecule.

- A molecular orbital can accommodate a maximum of two electrons with opposite spin and reflects the basic symmetry properties of the molecule.

- A molecular orbital multiplied by a spin function is a *molecular spin orbital*.

The basic approach to the calculation of molecular orbitals would be to solve the Hartree-Fock equations, which are formally identical for atoms and molecules. However, in general we don't have now three-dimensional rotational symmetry and we would need to solve multicenter integrals. This makes the problem so complex that finding direct iterative solutions to the Hartree–Fock equations is usually discarded. Instead, one postulates from the beginning that molecular orbitals can be expressed as linear combinations of atomic orbitals (LCAO). This is known as the *Roothaan method*.

Let's restrict the following discussion to closed shells, which is the easiest configuration to handle and corresponds in fact to the ground state configuration of many molecules. Since all electrons in a closed shell have their spins paired (total $S = 0$), the total antisymmetric wave function can be represented by a single Slater determinant,

$$\Psi = \frac{1}{\sqrt{(2N)!}} \begin{vmatrix} \psi_1(1) & \overline{\psi}_1(1) & \dots & \psi_N(1) & \overline{\psi}_N(1) \\ \psi_1(2) & \overline{\psi}_1(2) & \dots & \psi_N(2) & \overline{\psi}_N(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \psi_1(2N) & \overline{\psi}_1(2N) & \dots & \psi_N(2N) & \overline{\psi}_N(2N) \end{vmatrix}, \tag{3.197}$$

where the orthonormal molecular spin orbitals $\psi_j(\lambda)$ and $\overline{\psi}_j(\lambda)$ refer to the same spatial orbital but with $m_s = 1/2$ in the former and $m_s = -1/2$ in the latter. In this way, the $2N$ electrons are distributed in $2N$ spin orbitals or N spatial orbitals. As in the atomic case, our goal is to find the orbitals that minimize the energy $E = \langle \Psi | \mathcal{H} | \Psi \rangle$, where $\mathcal{H}$ is the complete electronic Hamiltonian of the molecule given by

$$\mathcal{H} = \sum_i \frac{p_i^2}{2m} + \sum_{i<j} \frac{e^2}{r_{ij}} - \sum_{i,\alpha} \frac{Z_\alpha e^2}{r_{i\alpha}} + \sum_{\alpha<\beta} \frac{Z_\alpha Z_\beta e^2}{r_{\alpha\beta}}. \tag{3.198}$$

The indices i, j refer to electrons; α, β to nuclei; Z_α is a nuclear charge and r stands for an inter–particle distance. This Hamiltonian contains all Coulombic terms under the assumption the nuclei of the molecule are fixed in their equilibrium configuration (there is no kinetic term for the nuclei).

The molecular orbitals that satisfy the variational principle are given by the solutions of the Hartree–Fock equations. For a closed shell we have seen (equation 3.150) that they take the form

$$F(1)\psi_i(1) = \varepsilon_i \psi_i(1), \tag{3.199}$$

where

$$F(1) = \mathcal{H}_0(1) + \sum_k \Big[2J_k(1) - K_k(1) \Big], \qquad \mathcal{H}_0(1) = \frac{p_1^2}{2m} - \sum_\alpha \frac{Z_\alpha e^2}{r_{1\alpha}}. \tag{3.200}$$

Here J_k and K_k are the Coulomb and exchange operators, respectively, which depend on the molecular orbitals themselves and are defined as in equations (3.147) and (3.148), with the extra factor 2 now

$$J_k(1)\psi_i(1) = \left[\int \psi_k^*(2) \left(\frac{e^2}{r_{12}} \right) \psi_k(2)\, d\tau_2 \right] \psi_i(1), \tag{3.201}$$

$$K_k(1)\psi_i(1) = \left[\int \psi_k^*(2) \left(\frac{e^2}{r_{12}} \right) \psi_i(2)\, d\tau_2 \right] \psi_k(1). \tag{3.202}$$

Also, similarly to equation (3.164), the total energy is given by[17]

$$E = 2\sum_i I_i + \sum_{i,k}\Big[2J(i,k) - K(i,k)\Big]. \tag{3.203}$$

Now is the time to introduce the constraint of the Roothaan method, which is to express the molecular orbitals as linear combinations of atomic orbitals,

$$\psi_i = \sum_\mu c_{\mu i}\phi_\mu, \qquad \mathbf{\Psi} = \mathbf{\Phi C}. \tag{3.204}$$

The second expression is written in matrix form, where $\mathbf{\Psi}$ and $\mathbf{\Phi}$ are row matrices and $\mathbf{C}$ is a square matrix. The ϕ_μ are the atomic orbitals which constitute the basis set. However, it is important to note that they need not be *true* atomic orbitals in the sense that it is not necessary for them to be eigenfunctions of an atomic Hamiltonian. They are *atomic* only to the extent that they are one-electron functions centered on the various nuclei and chosen largely for their ability to provide a flexible starting point for the calculation. The ϕ_μ are not necessarily orthogonal, therefore we define the overlap integral for the orbital ϕ_μ and ϕ_ν as

$$S_{\mu\nu} = \langle\phi_\mu|\phi_\nu\rangle = \int \phi^*_\mu(1)\phi_\nu(1)\, d\tau_1. \tag{3.205}$$

The molecular orbitals, on the contrary, are required to be orthogonal $\langle\psi_i|\psi_j\rangle = \delta_{ij}$. In the LCAO approximation this condition is written as

$$\langle\psi_i|\psi_j\rangle = \sum_{\mu,\nu} c^*_{\mu i}c_{\nu j}\,\langle\phi_\mu|\phi_\nu\rangle = \sum_{\mu,\nu} c^*_{\mu i}c_{\nu j}S_{\mu\nu} = \delta_{ij}, \qquad \mathbf{C}^\dagger\mathbf{SC} = 1. \tag{3.206}$$

The second expression is written in matrix form.

The charge density at a given point $\mathbf{r}$ is defined by

$$\rho(\mathbf{r}) = 2\sum_i \psi^*_i(\mathbf{r})\psi_i(\mathbf{r}). \tag{3.207}$$

The factor 2 is due to the double occupancy of the molecular orbitals by two electrons with opposite spin. In terms of the atomic orbitals this density is

$$\rho(\mathbf{r}) = 2\sum_i\sum_{\mu,\nu} c^*_{\mu i}c_{\nu i}\phi^*_\mu(\mathbf{r})\phi_\nu(\mathbf{r}) = \sum_{\mu,\nu} P_{\mu\nu}\phi^*_\mu(\mathbf{r})\phi_\nu(\mathbf{r}), \tag{3.208}$$

where $P_{\mu\nu}$ is the *density matrix* is defined as

$$P_{\mu\nu} = 2\sum_i c^*_{\mu i}c_{\nu i}. \tag{3.209}$$

If N is the number of occupied (spatial) orbitals, the closed shell will contain $2N$ electrons, therefore the integral of the charge density over all space must be equal to $2N$,

$$\int \rho(\mathbf{r})\, d\mathbf{r} = \sum_{\mu,\nu} P_{\mu\nu}\int \phi^*_\mu(\mathbf{r})\phi_\nu(\mathbf{r})\, d\mathbf{r} = \sum_{\mu,\nu} P_{\mu\nu}S_{\mu\nu} = 2N. \tag{3.210}$$

With these relations at hand and

$$\psi_i = \sum_\nu c_{\nu i}\phi_\nu, \qquad \psi_k = \sum_\sigma c_{\sigma k}\phi_\sigma, \qquad \psi^*_k = \sum_\lambda c^*_{\lambda k}\phi^*_\lambda, \tag{3.211}$$

[17] The extra factor 2 in front of I_i and $J(i,k)$ is due to the way we have built the Slater determinant, with $2N$ spin orbitals.

we can write

$$\mathcal{H}_0(1)\psi_i(1) = \sum_\nu \mathcal{H}_0(1) c_{\nu i}\phi_\nu(1), \tag{3.212}$$

$$J_k(1)\psi_i(1) = \sum_{\lambda,\sigma,\nu} c^*_{\lambda k} c_{\sigma k} c_{\nu i} \left[\int \phi^*_\lambda(2)\left(\frac{e^2}{r_{12}}\right)\phi_\sigma(2)\,d\tau_2\right]\phi_\nu(1), \tag{3.213}$$

$$K_k(1)\psi_i(1) = \sum_{\lambda,\sigma,\nu} c^*_{\lambda k} c_{\sigma k} c_{\nu i} \left[\int \phi^*_\lambda(2)\left(\frac{e^2}{r_{12}}\right)\phi_\nu(2)\,d\tau_2\right]\phi_\sigma(1), \tag{3.214}$$

$$F(1)\psi_i(1) = \sum_\nu \mathcal{H}_0(1)c_{\nu i}\phi_\nu(1) + \sum_{k,\lambda,\sigma,\nu} c^*_{\lambda k}c_{\sigma k}c_{\nu i}\left[2\left[\int \phi^*_\lambda(2)\left(\frac{e^2}{r_{12}}\right)\phi_\sigma(2)\,d\tau_2\right]\phi_\nu(1) - \left[\int \phi^*_\lambda(2)\left(\frac{e^2}{r_{12}}\right)\phi_\nu(2)\,d\tau_2\right]\phi_\sigma(1)\right] = \sum_\nu \varepsilon_i c_{\nu i}\phi_\nu(1). \tag{3.215}$$

This last expression is the Hartree–Fock equation in LCAO form. To proceed, it is useful to shorten the notation by defining

$$\left\langle \phi_\mu(1)\phi_\lambda(2)\left|\left(\frac{e^2}{r_{12}}\right)\right|\phi_\nu(1)\phi_\sigma(2)\right\rangle \equiv \langle \mu\nu|\lambda\sigma\rangle\,, \tag{3.216}$$

where the convention is to put the indices of the functions referring to the same electron in the same half–bracket ($\mu\nu$ for electron 1 and $\lambda\sigma$ for electron 2, in this case), and to put the complex conjugates on the left side of each half–bracket (μ and λ in this case). This definition implies

$$\langle \mu\nu|\lambda\sigma\rangle = \langle \lambda\sigma|\mu\nu\rangle\,, \qquad \langle \mu\nu|\lambda\sigma\rangle = \langle \nu\mu|\sigma\lambda\rangle^*\,. \tag{3.217}$$

More explicitly,

$$\langle \mu\nu|\lambda\sigma\rangle = \langle \phi_\mu(1)\phi_\lambda(2)\,|\,\mathcal{V}\,|\,\phi_\nu(1)\phi_\sigma(2)\rangle\,, \tag{3.218}$$

$$\langle \lambda\sigma|\mu\nu\rangle = \langle \phi_\lambda(1)\phi_\mu(2)\,|\,\mathcal{V}\,|\,\phi_\sigma(1)\phi_\nu(2)\rangle\,, \tag{3.219}$$

$$\langle \nu\mu|\sigma\lambda\rangle^* = \langle \phi_\nu(1)\phi_\sigma(2)\,|\,\mathcal{V}\,|\,\phi_\mu(1)\phi_\lambda(2)\rangle^* = \langle \phi_\mu(1)\phi_\lambda(2)\,|\,\mathcal{V}\,|\,\phi_\nu(1)\phi_\sigma(2)\rangle\,. \tag{3.220}$$

Multiplying both sides of equation (3.215) by $\phi^*_\mu(1)$ we get

$$\sum_\nu c_{\nu i}\,\phi^*_\mu(1)\mathcal{H}_0(1)\phi_\nu(1) + \sum_{k,\lambda,\sigma,\nu} c^*_{\lambda k}c_{\sigma k}c_{\nu i}\left[2\left[\int \phi^*_\mu(1)\phi^*_\lambda(2)\left(\frac{e^2}{r_{12}}\right)\phi_\sigma(2)\,d\tau_2\right]\phi_\nu(1) - \left[\int \phi^*_\mu(1)\phi^*_\lambda(2)\left(\frac{e^2}{r_{12}}\right)\phi_\nu(2)\,d\tau_2\right]\phi_\sigma(1)\right] = \sum_\nu \varepsilon_i c_{\nu i}\,\phi^*_\mu(1)\phi_\nu(1). \tag{3.221}$$

Now we perform an additional integration over the coordinates of electron (1),

$$\sum_\nu c_{\nu i}\int \phi^*_\mu(1)\mathcal{H}_0(1)\phi_\nu(1)\,d\tau_1 + \sum_{k,\lambda,\sigma,\nu} c^*_{\lambda k}c_{\sigma k}c_{\nu i}\left[2\int \phi^*_\mu(1)\phi^*_\lambda(2)\left(\frac{e^2}{r_{12}}\right)\phi_\nu(1)\phi_\sigma(2)\,d\tau_1\,d\tau_2 - \int \phi^*_\mu(1)\phi^*_\lambda(2)\left(\frac{e^2}{r_{12}}\right)\phi_\sigma(1)\phi_\nu(2)\,d\tau_1\,d\tau_2\right] = \sum_\nu \varepsilon_i c_{\nu i}\int \phi^*_\mu(1)\phi_\nu(1)\,d\tau_1. \tag{3.222}$$

Using our compact notation this can be written as

$$\sum_{\nu} c_{\nu i} H_{\mu\nu} + \sum_{k,\lambda,\sigma,\nu} c^*_{\lambda k} c_{\sigma k} c_{\nu i} \Big[2 \langle \mu\nu | \lambda\sigma \rangle - \langle \mu\sigma | \lambda\nu \rangle \Big] = \sum_{\nu} \varepsilon_i c_{\nu i} S_{\mu\nu}, \tag{3.223}$$

where we have defined

$$H_{\mu\nu} = \langle \phi_\mu(1) \,|\, \mathcal{H}_0(1) \,|\, \phi_\nu(1) \rangle \,. \tag{3.224}$$

According to our previous definition of the density matrix, we have $P_{\lambda\sigma} = 2 \sum_k c^*_{\lambda k} c_{\sigma k}$. Therefore

$$\sum_{\nu} c_{\nu i} H_{\mu\nu} + \sum_{\lambda,\sigma,\nu} P_{\lambda\sigma} \left[\langle \mu\nu | \lambda\sigma \rangle - \frac{1}{2} \langle \mu\sigma | \lambda\nu \rangle \right] = \sum_{\nu} \varepsilon_i c_{\nu i} S_{\mu\nu}. \tag{3.225}$$

Finally, defining

$$F_{\mu\nu} = H_{\mu\nu} + \sum_{\lambda,\sigma} P_{\lambda\sigma} \left[\langle \mu\nu | \lambda\sigma \rangle - \frac{1}{2} \langle \mu\sigma | \lambda\nu \rangle \right], \tag{3.226}$$

we can write the Hartree–Fock equation in this compact form

$$\sum_{\nu} (F_{\mu\nu} - \varepsilon_i S_{\mu\nu}) c_{\nu i} = 0, \qquad (\mathbf{F} - \boldsymbol{\varepsilon}\mathbf{S})\mathbf{C} = 0, \tag{3.227}$$

where the last expression is written in matrix notation, with $\boldsymbol{\varepsilon}$ a diagonal matrix whose diagonal elements ε_i are the orbital energies. This set of homogeneous equations will have nontrivial solutions only when

$$\det |F_{\mu\nu} - \varepsilon_i S_{\mu\nu}| = 0 \tag{3.228}$$

At this point the LCAO approximation has converted the original partial differential equations (Hartree-Fock) into algebraic equations (Roothaan). This is a very significant simplification, without which it is extremely difficult to calculate reliable molecular orbitals of any but the simplest molecules. The Roothaan equations, also called LCAO-SCF equations[18], are the basic equations of molecular orbital theory. For molecules with atoms of high enough Z it might be necessary to include also the spin-orbit interaction. This can be done in a way that closely resembles the atomic case.

[18] Also SCF-LCAO-MO, for *Self-Consistent Field-Linear Combination of Atomic Orbitals-Molecular Orbital.*

3.4 Problems.

1. Compute the expectation value of the electronic repulsion term of the helium Hamiltonian using the trial function $\psi_0 = (8/\pi a^3)e^{-2(r_1+r_2)/a}$ and obtain the result shown in equation (3.13).
Hint: fix $\mathbf{r}_1$ and do the integral over $\mathbf{r}_2$ choosing the coordinate system in such a way that the polar axis (z-axis) runs along $\mathbf{r}_1$, so that $|\mathbf{r_1} - \mathbf{r_2}| = \sqrt{r_1^2 + r_2^2 - 2r_1 r_2 \cos\theta_2}$. Then do the integral over $\mathbf{r}_1$.

2. Using the variational method derive a generic expression to estimate the ground state energy of an atom/ion with Z protons in the nucleus and only two bound electrons. You may adapt the expressions we have used to study the helium atom. Give the numerical result in eV for $Z = 3$.

3. Consider a Helium atom with the Hamiltonian

$$\mathcal{H} = \left(\frac{p_1^2}{2m} - \frac{2e^2}{r_1}e^{-\lambda r_1}\right) + \left(\frac{p_2^2}{2m} - \frac{2e^2}{r_2}e^{-\lambda r_2}\right) + \frac{e^2}{r_{12}},$$

where the electron-nucleus interaction is described by a Yukawa potential.

i) Use the trial function

$$\psi_1(\mathbf{r}_1, \mathbf{r}_2) = \frac{\zeta^3}{\pi a^3}e^{-\zeta(r_1+r_2)/a},$$

to show that,

$$E(\zeta) = \langle\psi_1|\mathcal{H}|\psi_1\rangle = \left[-2\zeta^2 - \frac{5\zeta}{4} + \frac{32\zeta^3}{(2\zeta + \lambda a)^2}\right] E_1,$$

where a is the Bohr radius and $E_1 = e^2/(2a) \approx 13.6$ eV.

(**Hint:** you can use much of what was done in class for the helium atom.)

ii) According to the variational principle, an estimate of the ground-state energy of this system can be obtained by minimizing the previous function. However, the minimum condition, $E'(\zeta) = 0$, leads to a transcendental equation that can only be solved numerically. An approximate solution can be found for $\lambda a \ll 2\zeta$. In such case show that

$$E(\zeta) \approx \left[-2\zeta^2 + \frac{27}{4}\zeta - 8\lambda a + 6\frac{(\lambda a)^2}{\zeta}\right] E_1.$$

When $\lambda = 0$, this function has a minimum for $\zeta_0 = 3^3/2^4$ (see the lecture notes), therefore for small $\lambda \neq 0$ the new minimum must be found close to ζ_0. Define then $\epsilon = \zeta - \zeta_0$ and find that

$$\epsilon \approx -\frac{3}{2}\left(\frac{\lambda a}{\zeta_0}\right)^2,$$

concluding that the approximate solution is

$$\zeta \approx \zeta_0 - \frac{3}{2}\left(\frac{\lambda a}{\zeta_0}\right)^2 = \frac{3^3}{2^4} - \frac{2^7}{3^5}(\lambda a)^2,$$

which leads to a minimum energy of

$$E_{\min} \approx \left[-2\zeta_0^2 + \frac{27}{4}\zeta_0 - 8(\lambda a) + \left(\frac{12}{\zeta_0} - \frac{81}{8\zeta_0^2}\right)(\lambda a)^2\right] E_1$$
$$= \left[\frac{3^6}{2^7} - 2^3(\lambda a) + \frac{2^5}{3^2}(\lambda a)^2\right] E_1.$$

Hint: use the Taylor series $1/(1+x)^2 = 1 - 2x + 3x^2 - \ldots$

4. Compute the expectation value known as *overlap integral* in the variational approach to the hydrogen molecule ion H_2^+ and check that you obtain the result shown in equation (3.27).
Hint: choose a coordinate system such that proton 1 is at the origin and proton 2 is on the z-axis at a point $z = R$, so that $r_1 = r$ and $r_2 = \sqrt{r^2 + R^2 - 2rR\cos\theta}$.

5. Compute the expectation values known as *direct integral* and *exchange integral* in the variational approach to the hydrogen molecule ion H_2^+ and check that you obtain the results shown in equations (3.33) and (3.34).
Hint: perform the calculations using the same reference system of the previous problem. If the notation used in the text for the wavefunctions seems confusing when doing the integrals, you may write them as $\psi_0^{(p_1)}(\mathbf{r})$ and $\psi_0^{(p_2)}(\mathbf{r})$, instead of $\psi_0(\mathbf{r}_1)$ and $\psi_0(\mathbf{r}_2)$.

6. In the variational study of the hydrogen molecule ion H_2^+ use as a new trial wave function $\psi = A[\psi_0(\mathbf{r}_1) - \psi_0(\mathbf{r}_2)]$, where $\psi_0(\mathbf{r}_1)$ and $\psi_0(\mathbf{r}_2)$ are hydrogen ground eigenstates corresponding to the electron being bound to one or the other proton. You don't need to do the integrals all over again, just make the necessary changes in the signs to get a new $F(x)$ (see equation (3.40)). Plot the new $F(x)$ function and explain if bonding is possible in this situation.

7. Prove that the total orbital angular momentum of the entire electronic system of an atom with N electrons, $\mathbf{L} = \sum_{i=1}^{N} \boldsymbol{\ell}_i$, commutes with the Hamiltonian $\mathcal{H}_I = \mathcal{H}_0 + \mathcal{H}_1$ given by equation (3.42).
Hints: the commutation with $\mathcal{H}_0$ should be easy to prove. When studying the commutation with $\mathcal{H}_1$ you need to realize that $\boldsymbol{\ell}_k \frac{1}{r_{ij}} \neq \frac{1}{r_{ij}} \boldsymbol{\ell}_k$. Express the orbital angular momentum of the k-th electron as $\boldsymbol{\ell}_k = -i\hbar\, \mathbf{r}_k \times \boldsymbol{\nabla}_k$ and use a dummy wave function to prove that

$$\boldsymbol{\ell}_k \frac{1}{r_{ij}} = \frac{1}{r_{ij}} \boldsymbol{\ell}_k - i\hbar\, \mathbf{r}_k \times \boldsymbol{\nabla}_k \left(\frac{1}{r_{ij}}\right).$$

Also, noting that the operator $\boldsymbol{\nabla}_k$ differentiates with respect to the coordinates of the k-th electron only, you should be able to prove and use the relation

$$\boldsymbol{\nabla}_k \left(\frac{1}{r_{ij}}\right) = -\frac{\mathbf{r}_i - \mathbf{r}_j}{r_{ij}^3}\delta_{ik} - \frac{\mathbf{r}_j - \mathbf{r}_i}{r_{ij}^3}\delta_{jk}.$$

8. In some atoms, the charge-screening effect of other electrons on the motion of each of them can be well approximated by the replacement of the Coulomb potential with

the so-called *Hulthen potential*,

$$V(r) = -\frac{Ze^2}{a}\frac{1}{e^{r/a}-1},$$

where $a > 0$ is the effective screening radius.

- Study the asymptotic behaviour of this potential for the limiting cases $r \ll a$ and $r \gg a$, explaining the corresponding physical meaning.
- Assuming that $r \ll a$, use perturbation theory to calculate the energy spectrum in this model and prove that the ℓ-degeneracy of the pure Coulomb potential is broken.

Hints: use the Taylor series expansion $\frac{x}{e^x-1} = 1 - \frac{x}{2} + \frac{x^2}{12} - \ldots$; use also the expression for $\langle r \rangle_{n\ell}$ computed with hydrogen-like wave functions.

9. Consider the following configurations of two electrons:

   i) One electron in orbital a with spin β ($m_s = -1/2$) and the other in orbital b with spin α ($m_s = 1/2$).

   ii) One electron in orbital a with spin α and the other in orbital b with spin β.

   In those conditions address the following items:

   a) Write the Slater determinants of both configurations.

   b) Calculate their energy by assuming the Hamiltonian

   $$H = \frac{p_1^2}{2m} - \frac{Ze^2}{r_1} + \frac{p_2^2}{2m} - \frac{Ze^2}{r_2} + \frac{e^2}{r_{12}} = h(1) + h(2) + \frac{e^2}{r_{12}}.$$

   Express the result in terms of the integrals $I_a = \langle a|h|a\rangle$, $I_b = \langle b|h|b\rangle$ and $J_{ab} = \langle a(1)b(2)|\frac{e^2}{r_{12}}|a(1)b(2)\rangle$.

   c) Show that i) and ii) are not eigenstates of $\mathbf{S}^2$, where $\mathbf{S} = \mathbf{s}(1) + \mathbf{s}(2)$, while their linear symmetric and antisymmetric combinations are.

   d) Calculate the energy of these two eigenstates, expressing the results as a function of the integrals I_a, I_b, J_{ab}, defined before, and $K_{ab} = \langle a(1)b(2)|\frac{e^2}{r_{12}}|b(1)a(2)\rangle$ (the exchange integral).

10. Compute the degeneracy of the niobium ground state configuration, $[\mathrm{Kr}](5\mathrm{s}^1)(4\mathrm{d}^4)$, (the number of ways the electrons can be placed in the different states of the configuration).

11. Assuming that LS coupling holds, list the possible spectral terms $^{2S+1}\mathrm{L}$ which result from the following electronic configurations
(a): $(np)(n'p)(n''p)$; (b): (nd^4); (c): (nd^8); (d): $(ns)(n'p^2)$.

12. Assuming that jj coupling holds, list the possible spectral terms $(j_1 j_2)_J$ which result from the following electronic configurations
(a): $(ns)(n'p)$; (b): (nd^2).

13. The electronic configuration of the silicon atom is $[\mathrm{Ne}](3s)^2(3p)^2$.

a) Determine, assuming LS coupling, the possible atomic terms for this configuration and explain which one is expected to correspond to the silicon ground state.

One of the possible terms is $^3\mathrm{P}$. For this term:

b) Study the fine structure due to the spin-orbit coupling. Restricted to the subspace of a term of degenerate states having the same L and S quantum numbers, the corresponding Hamiltonian can be written as $\mathcal{H}_{\mathrm{so}} = \zeta\,\mathbf{LS}$. Work out the spin, orbital angular momentum and total angular momentum for each of the fine structure levels. Compute the energy differences between the levels and draw them in a diagram (leave the results in terms of ζ).

c) When an atom is surrounded by a weak magnetic field, $\mathbf{B}$, it takes place an additional energy splitting on top of the fine structure. Assume the magnetic field is oriented along the z axis, $\mathbf{B} = B\hat{\mathbf{z}}$, and draw this further splitting in the previous energy diagram, indicating the necessary quantum numbers to identify each level. Compute also the energy differences between them.

d) Repeat the previous study assuming the magnetic field B is very strong. If the spin-orbit interaction is neglected, how many levels with *different* energies are there? Draw the appropriate diagram.

e) Study now the case in which the spin-orbit interaction is of similar magnitude to the perturbation caused by the external magnetic field.

14. Using the change of variable given in equation (3.113), derive the Thomas–Fermi universal equation, $x^{1/2}\chi'' = \chi^{3/2}$, and prove that it predicts an atomic electron density given by

$$\rho(x) = \frac{Z}{4\pi b^3}\left(\frac{\chi}{x}\right)^{3/2}.$$

15. Suppose that, in the Thomas-Fermi model, we define the radius of an atom as that value, $R = r(\alpha)$, which encloses a fixed fraction of the Z electrons, $(1-\alpha)Z$, with $0 < \alpha < 1$. This means that

$$4\pi \int_0^R \rho(r)\, r^2\, \mathrm{d}r = (1-\alpha)Z.$$

Take the density function of the previous problem, make the change of variable $r = bx$ and arrive at the expression

$$\int_0^{x(\alpha)} x^{1/2}\, \chi^{3/2}\, \mathrm{d}x = 1 - \alpha.$$

Then, using the Thomas–Fermi equation, $x^{1/2}\chi'' = \chi^{3/2}$, and the boundary condition

$\chi(0) = 1$, prove that $x(\alpha) \equiv x_\alpha$ satisfies the equation

$$\chi(x_\alpha) - x_\alpha \chi'(x_\alpha) = \alpha$$

16. Prove that two spin orbitals that satisfy the Hartre-Fock equations, (3.161) and (3.162), are orthogonal: $\langle \psi_k | \psi_i \rangle = 0$, for $\varepsilon_i \neq \varepsilon_k$.
Hint: multiply (3.161) by $\psi_k^*(1)$ and integrate. Then take the complex conjugate of (3.162), multiply by $\psi_i(1)$ and integrate. Finally, compare the two expressions you got.

17. Consider a system of two fermions of spin 1/2 described by the Hamiltonian

$$\mathcal{H} = \mathcal{H}_0(1) + \mathcal{H}_0(2) + \mathcal{V},$$

where

$$\mathcal{H}_0(1) = \frac{p_1^2}{2m} + \frac{1}{2} m\omega^2 x_1^2, \qquad \mathcal{H}_0(2) = \frac{p_2^2}{2m} + \frac{1}{2} m\omega^2 x_2^2, \qquad \mathcal{V} = \frac{1}{2} m\Omega^2 (x_1 - x_2)^2.$$

The spin wave function of the ground state is the anti-symmetric singlet $\frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$ whereas the space wave-function is symmetric under the $x_1 \leftrightarrow x_2$ exchange.

a) Introduce the centre of mass and relative position variables

$$X \equiv \frac{1}{2}(x_1 + x_2), \qquad x \equiv x_1 - x_2,$$

and show, by converting the equation into two uncoupled harmonic oscillators, that the ground state exact space wave-function of this system is

$$\Psi(x_1, x_2) = \frac{\alpha}{\sqrt{\pi}} (1 + 2K)^{1/8} \exp\left[-\frac{\alpha^2}{4} (1 + \sqrt{1 + 2K})(x_1^2 + x_2^2) \right] \exp\left[-\frac{\alpha^2}{2} (1 - \sqrt{1 + 2K}) x_1 x_2 \right],$$

and its energy

$$E_0 = \frac{\hbar^2 \alpha^2}{2m} (1 + \sqrt{1 + 2K}),$$

where $\alpha \equiv (m\omega/\hbar)^{1/2}$ and $K \equiv (\Omega/\omega)^2$. You do not have to solve the harmonic oscillator equation, simply take the ground state solution from the bibliography.

b) Demonstrate that

$$\Psi_{\mathrm{HF}}(x_1, x_2) = \phi(x_1)\phi(x_2),$$

with

$$\phi(x) = \frac{\alpha^{1/2}}{\pi^{1/4}} (1 + K)^{1/8} \exp\left[-\frac{\alpha^2}{2} \sqrt{1 + K} x^2 \right],$$

is the space wave-function solution of the Hartree-Fock equations for this system

$$\mathcal{H}_0(1)a(1) + \left[\int b^*(2) \mathcal{V} b(2)\, d\tau_2 \right] a(1) - \left[\int b^*(2) \mathcal{V} a(2)\, d\tau_2 \right] b(1) = \varepsilon_a a(1),$$

$$\mathcal{H}_0(2)b(2) + \left[\int a^*(1) \mathcal{V} a(1)\, d\tau_1 \right] b(2) - \left[\int a^*(1) \mathcal{V} b(1)\, d\tau_1 \right] a(2) = \varepsilon_b b(2),$$

being $a = |\phi\uparrow\rangle$ and $b = |\phi\downarrow\rangle$, and also show that the estimated energy is

$$E_0^{\rm HF} = \frac{\hbar^2\alpha^2}{2m} 2\sqrt{1+K}.$$

c) Study the ratio

$$R(K) \equiv \frac{E_0^{\rm HF}}{E_0}$$

as a function of K by plotting $R(K)$ versus K in the $[0, 20]$ range and discuss the result.

18. Study the Helium atom described by the Hamiltonian

$$\mathcal{H} = \mathcal{H}_0(1) + \mathcal{H}_0(2) + \frac{e^2}{r_{12}}, \qquad \mathcal{H}_0(i) = \frac{p_i^2}{2m} - \frac{2e^2}{r_i},$$

where $(i = 1, 2)$. Assume that its ground state can be described using the Slater determinant

$$\Psi = \frac{1}{\sqrt{2}}(\psi_1(1)\psi_2(2) - \psi_2(1)\psi_1(2)),$$

where $\psi_1(i) = \phi_{1s}(r_i)\,|\uparrow_i\rangle$ and $\psi_2(i) = \phi_{1s}(r_i)\,|\downarrow_i\rangle$ (the notation $|\uparrow_i\rangle = \alpha(i)$ and $|\downarrow_i\rangle = \beta(i)$ can also be used).

a) Show that the Hartree-Fock equations

$$\mathcal{H}_0(1)\psi_1(1) + \left[\int \psi_2^*(2)\left(\frac{e^2}{r_{12}}\right)\psi_2(2)\,d\tau_2\right]\psi_1(1) - \left[\int \psi_2^*(2)\left(\frac{e^2}{r_{12}}\right)\psi_1(2)\,d\tau_2\right]\psi_2(1) = \varepsilon_1\psi_1(1),$$

$$\mathcal{H}_0(2)\psi_2(2) + \left[\int \psi_1^*(1)\left(\frac{e^2}{r_{12}}\right)\psi_1(1)\,d\tau_1\right]\psi_2(2) - \left[\int \psi_1^*(1)\left(\frac{e^2}{r_{12}}\right)\psi_2(1)\,d\tau_1\right]\psi_1(2) = \varepsilon_2\psi_2(2),$$

lead to the Hartree equation for the spatial orbital $\phi_{1s}(r)$

$$\left[-\frac{\hbar^2}{2m}\nabla^2 - \frac{2e^2}{r} + \int \left|\phi_{1s}(r')\right|^2 \frac{e^2}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}'\right]\phi_{1s}(r) = \varepsilon_{1s}\phi_{1s}(r).$$

b) Use the orbital

$$\phi_{1s}(r) = \sqrt{\frac{\zeta^3}{\pi a^3}}e^{-\zeta r/a},$$

to make an estimate of the potential

$$U(r) = -\frac{2e^2}{r} + \int \left|\phi_{1s}(r')\right|^2 \frac{e^2}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}'.$$

Prepare a plot of the effective charge in units of the proton charge, $Z^{\rm eff}(r) = U(r)/(-e^2/r)$, and interpret the result.

c) For the Helium atom the energy correction due to the Darwin term is given by

$$\mathcal{H}_D = \frac{e\hbar^2}{8m^2c^2}\left(\nabla_1 \mathbf{E}_1 + \nabla_2 \mathbf{E}_2\right) = \frac{\hbar^2}{8m^2c^2}\left(\nabla_1^2 U(r_1) + \nabla_2^2 U(r_2)\right).$$

Show, using perturbation theory, that the energy shift due to the Darwin term, considering the potential obtained in b), is:

$$E_D = \frac{15}{8}\zeta^3\alpha^4 mc^2,$$

and give E_D in eV considering that $\zeta = 1.69$. Discuss the difference with the result that is obtained with two non-interacting electrons, that can be found in equation (1.196) of the handouts:

$$E_D = 2 \times 8\alpha^4 mc^2 = 16\alpha^4 mc^2.$$

Hint: it is very easy to calculate $\nabla^2 U(r)$ considering that:

$$\nabla^2\left(\frac{1}{|\mathbf{r} - \mathbf{r}'|}\right) = -4\pi\delta(\mathbf{r} - \mathbf{r}').$$

19. Consider the Lithium atom described by the Hamiltonian

$$\mathcal{H} = \sum_{i=1}^{3}\mathcal{H}_0(i) + \sum_{i<j}^{3}\frac{e^2}{r_{ij}},$$

where:

$$\mathcal{H}_0(i) = \frac{p_i^2}{2m} - \frac{3e^2}{r_i}.$$

The ground state is given by the $(1s)^2(2s)^1$ configuration. If the orbitals for $1s$ and $2s$ electrons are given by a hydrogen-like $1s$ wave function with charge ζ and a hydrogen-like $2s$ wave function with charge $\zeta/2$:

$$\begin{aligned}\phi_{1s}(r) =& \sqrt{\frac{\zeta^3}{\pi a^3}}e^{-\zeta r/a},\\ \phi_{2s}(r) =& \frac{A}{16}\sqrt{\frac{\zeta^3}{\pi a^3}}\left(2 - c\frac{\zeta r}{2a}\right)e^{-\zeta r/4a},\end{aligned}$$

where c is used to keep ϕ_{1s} and ϕ_{2s} orthogonal and A to normalize ϕ_{2s}.

1) Show that

$$c = \frac{5}{3}, \qquad A = \sqrt{\frac{3}{13}}.$$

Consider the single Slater determinant

$$\Psi = \frac{1}{\sqrt{6}}\begin{vmatrix} \phi_{1s}(r_1)\alpha(1), & \phi_{1s}(r_1)\beta(1), & \phi_{2s}(r_1)\alpha(1) \\ \phi_{1s}(r_2)\alpha(2), & \phi_{1s}(r_2)\beta(2), & \phi_{2s}(r_2)\alpha(2) \\ \phi_{1s}(r_3)\alpha(3), & \phi_{1s}(r_3)\beta(3), & \phi_{2s}(r_3)\alpha(3) \end{vmatrix}.$$

2) Show that

$$\langle\Psi|\mathcal{H}|\Psi\rangle = \langle\phi_{1s}\phi_{1s}\phi_{2s}|\mathcal{H}|\phi_{1s}\phi_{1s}\phi_{2s}\rangle - \langle\phi_{1s}\phi_{1s}\phi_{2s}|\mathcal{H}|\phi_{2s}\phi_{1s}\phi_{1s}\rangle.$$

3) Demonstrate that

$$\langle\phi_{1s}\phi_{1s}\phi_{2s}|\mathcal{H}|\phi_{1s}\phi_{1s}\phi_{2s}\rangle = 2\langle\phi_{1s}|\mathcal{H}_0|\phi_{1s}\rangle + \langle\phi_{2s}|\mathcal{H}_0|\phi_{2s}\rangle + \left\langle \phi_{1s}\phi_{1s}\left|\frac{e^2}{r_{12}}\right|\phi_{1s}\phi_{1s}\right\rangle + 2\left\langle \phi_{1s}\phi_{2s}\left|\frac{e^2}{r_{12}}\right|\phi_{1s}\phi_{2s}\right\rangle.$$

4) Also show that

$$\langle\phi_{1s}\phi_{1s}\phi_{2s}|\mathcal{H}|\phi_{2s}\phi_{1s}\phi_{1s}\rangle = \left\langle \phi_{1s}\phi_{2s}\left|\frac{e^2}{r_{12}}\right|\phi_{2s}\phi_{1s}\right\rangle.$$

5) Use the previous results and the variational principle to estimate ζ and the ground state energy of Lithium.

Chapter 4

Interaction between atoms and radiation.

To study the interaction of atoms with electromagnetic radiation we need to quantize the electromagnetic field, and to do that we'll start with a (very) brief introduction to quantum field theory.

4.1 The electromagnetic radiation field.

Let's start with a pure radiation field without charge and current sources. In this situation, the electromagnetic field can be described with just a vector potential, $\mathbf{A}$, which can be expressed as a linear superposition of plane waves. For normalization purposes, we consider a field in a cubic box of side L, very large compared with the relevant dimensions of our (atomic) problem. Both the magnitude of L and the box shape will be seen to have no impact on the physics results. We impose periodic boundary conditions at each face of the cube, which leads to the requirement that all plane waves must satisfy the relations

$$e^{ik_x x} = e^{ik_x(x+L)}, \qquad e^{ik_y y} = e^{ik_y(y+L)}, \qquad e^{ik_z z} = e^{ik_z(z+L)}. \tag{4.1}$$

Therefore

$$k_x = \frac{2\pi N_x}{L}, \qquad k_y = \frac{2\pi N_y}{L}, \qquad k_z = \frac{2\pi N_z}{L}, \qquad N_x, N_y, N_z = 0, \pm 1, \pm 2, \ldots \tag{4.2}$$

These are the components of the propagation vector or wave vector, which can be normalized to provide a unit vector in the direction of the wave propagation ($k = \omega_{\mathbf{k}}/c = 2\pi\nu/c = 2\pi/\lambda$ for a plane wave of wavelength λ moving at speed c),

$$\mathbf{k} = \frac{2\pi}{L}(N_x, N_y, N_z), \qquad \hat{\mathbf{k}} = \frac{\mathbf{k}}{k}, \quad k = |\mathbf{k}| = \frac{\omega_{\mathbf{k}}}{c}. \tag{4.3}$$

The periodic conditions give rise to an infinite (discrete) set of modes in the cavity, each one characterized by the set of integers (N_x, N_y, N_z), leaving polarization aside for the moment. The number of modes contained in an interval specified by ΔN_x, ΔN_y and ΔN_z is just given by

$$\Delta N = \Delta N_x \Delta N_y \Delta N_z = \left(\frac{L}{2\pi}\right)^3 \Delta k_x \Delta k_y \Delta k_z. \tag{4.4}$$

For a box with dimensions large compared to the radiation wavelength, we may assume the modes to form a continuous distribution. We can then write the number of propagation modes contained in an interval dk and in the direction defined by the element of solid angle $d\Omega$ as

$$dN = \left(\frac{L}{2\pi}\right)^3 dk_x dk_y dk_z = \left(\frac{L}{2\pi}\right)^3 k^2 \sin\theta \, dk \, d\theta \, d\varphi = \frac{V}{(2\pi)^3} \, k^2 \, dk \, d\Omega. \tag{4.5}$$

The (real) vector potential can be expressed as a linear superposition of plane waves,

$$\mathbf{A}(\mathbf{r},t)=\sum_{\mathbf{k},\lambda}\hat{\mathbf{e}}_{\mathbf{k}\lambda}\Big[A_{\mathbf{k}\lambda}\exp\big[i(\mathbf{kr}-\omega_{\mathbf{k}}t)\big]+A^{*}_{\mathbf{k}\lambda}\exp\big[-i(\mathbf{kr}-\omega_{\mathbf{k}}t)\big]\Big], \tag{4.6}$$

where $A_{\mathbf{k}\lambda}$ is a constant amplitude for the mode $(\mathbf{k},\lambda)$ and $\hat{\mathbf{e}}_{\mathbf{k}\lambda}$ is a real unit vector which indicates the linear polarization of the wave. For a given propagation direction, $\mathbf{k}$, there are only two independent polarization vectors, $\hat{\mathbf{e}}_{\mathbf{k}1}$ and $\hat{\mathbf{e}}_{\mathbf{k}2}$, orthogonal to each other

$$\hat{\mathbf{e}}_{\mathbf{k}\lambda}\hat{\mathbf{e}}_{\mathbf{k}\lambda'}=\delta_{\lambda\lambda'},\qquad(\lambda,\lambda'=1,2). \tag{4.7}$$

The label λ in the previous expressions indicates the two independent polarizations of electromagnetic waves, not the wavelength. The Coulomb gauge condition,[1] $\boldsymbol{\nabla}\mathbf{A}=0$, ensures the transversality of the electromagnetic fields, so that the polarization vector and the propagation vector are always orthogonal to each other, $\hat{\mathbf{e}}_{\mathbf{k}\lambda}\hat{\mathbf{k}}=0$. The set $\{\hat{\mathbf{e}}_{\mathbf{k}1},\hat{\mathbf{e}}_{\mathbf{k}2},\hat{\mathbf{k}}\}$ forms a right-handed system of mutually orthogonal unit vectors. The linear polarization vectors can be replaced by helicity polarization vectors in the following way,

$$\hat{\mathbf{e}}_{\mathbf{k}+}=-\frac{1}{\sqrt{2}}\big(\hat{\mathbf{e}}_{\mathbf{k}1}+i\hat{\mathbf{e}}_{\mathbf{k}2}\big),\qquad\hat{\mathbf{e}}_{0}=\hat{\mathbf{k}},\qquad\hat{\mathbf{e}}_{\mathbf{k}-}=\frac{1}{\sqrt{2}}\big(\hat{\mathbf{e}}_{\mathbf{k}1}-i\hat{\mathbf{e}}_{\mathbf{k}2}\big). \tag{4.8}$$

They satisfy

$$\hat{\mathbf{e}}^{*}_{\mathbf{k}q}\times\hat{\mathbf{e}}_{\mathbf{k}q'}=iq\hat{\mathbf{k}}\delta_{qq'},\qquad(q,q'=\pm1). \tag{4.9}$$

For $q=+1$ this cross product gives a unit vector parallel to the direction of propagation, while for $q=-1$ gives one anti-parallel. Therefore $\hat{\mathbf{e}}_{\mathbf{k}+}$ and $\hat{\mathbf{e}}_{\mathbf{k}-}$ are called positive and negative helicity unit vectors, respectively. They correspond to left- and right-handed circularly polarized electromagnetic waves. The rotational properties of a vector field are governed by an orbital angular momentum operator $\mathbf{L}$, and a spin angular momentum operator $\mathbf{S}$ (with $S=1$), which are coupled according to the relation $\mathbf{J}=\mathbf{L}+\mathbf{S}$ (see equation (B.267)). In this case, the unit vectors $\{\hat{\mathbf{e}}_{\mathbf{k}+},\hat{\mathbf{e}}_{0},\hat{\mathbf{e}}_{\mathbf{k}-}\}$ are simultaneous eigenvectors of S^2 and S_z, with the eigenvalues of S_z being $+1$, 0 and -1. Since the electromagnetic field is a vector field one would expect to find also three eigenvectors, but the transversity condition $\hat{\mathbf{e}}_{\mathbf{k}\lambda}\hat{\mathbf{k}}=0$ rules out a polarization parallel to the direction of propagation, leaving only the two helicity vectors $\hat{\mathbf{e}}_{\mathbf{k}+}$ and $\hat{\mathbf{e}}_{\mathbf{k}-}$ (or, alternatively, the two linear polarization vectors $\hat{\mathbf{e}}_{\mathbf{k}1}$ and $\hat{\mathbf{e}}_{\mathbf{k}2}$). In a radiation field, the electric $\mathbf{E}$ and magnetic $\mathbf{B}$ fields are transverse to the direction of propagation.

For a pure radiation field the relation $\mathbf{E}=-\boldsymbol{\nabla}\phi-\frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}$ reduces to $\mathbf{E}=-\frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}$, therefore the electric field can be computed as

$$\mathbf{E}(\mathbf{r},t)=-\frac{1}{c}\frac{\partial}{\partial t}\mathbf{A}(\mathbf{r},t)=\frac{i}{c}\sum_{\mathbf{k},\lambda}\omega_{\mathbf{k}}\hat{\mathbf{e}}_{\mathbf{k}\lambda}\Big[A_{\mathbf{k}\lambda}\exp\big[i(\mathbf{kr}-\omega_{\mathbf{k}}t)\big]-A^{*}_{\mathbf{k}\lambda}\exp\big[-i(\mathbf{kr}-\omega_{\mathbf{k}}t)\big]\Big]. \tag{4.10}$$

To find the magnetic field we need to evaluate $\boldsymbol{\nabla}\times\mathbf{A}$. For any function f we have

$$\boldsymbol{\nabla}\times(f\mathbf{A})=(\boldsymbol{\nabla}f)\times\mathbf{A}+f(\boldsymbol{\nabla}\times\mathbf{A}), \tag{4.11}$$

therefore

$$\boldsymbol{\nabla}\times\left(A_{\mathbf{k}\lambda}\,e^{i\mathbf{kr}}\,\hat{\mathbf{e}}_{\mathbf{k}\lambda}\right)=\boldsymbol{\nabla}\left(A_{\mathbf{k}\lambda}\,e^{i\mathbf{kr}}\right)\times\hat{\mathbf{e}}_{\mathbf{k}\lambda}\;+\;A_{\mathbf{k}\lambda}\,e^{i\mathbf{kr}}\,(\boldsymbol{\nabla}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda})\,. \tag{4.12}$$

But $(\boldsymbol{\nabla}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda})=0$ because $\hat{\mathbf{e}}_{\mathbf{k}\lambda}$ is constant. Since $A_{\mathbf{k}\lambda}$ is also a constant, we only need to evaluate

$$\boldsymbol{\nabla}e^{i\mathbf{kr}}=i\mathbf{k}e^{i\mathbf{kr}}=i|\mathbf{k}|\hat{\mathbf{k}}e^{i\mathbf{kr}}=\frac{i}{c}\omega_{\mathbf{k}}\hat{\mathbf{k}}e^{i\mathbf{kr}}, \tag{4.13}$$

[1] The canonical quantization of the electromagnetic field in a relativistic invariant way is a difficult subject, essentially because of gauge invariance. For our purposes in this chapter it is sufficient to work just in the Coulomb gauge, where $\boldsymbol{\nabla}\mathbf{A}=0$ and A^0 (the scalar potential, ϕ) is a constrained variable, rather than a dynamical variable. This allows us to simplify the quantization procedure at the price of not having manifest Lorentz invariance.

so that

$$\boldsymbol{\nabla} \times \left(A_{\mathbf{k}\lambda}\, e^{i\mathbf{kr}}\, \hat{\mathbf{e}}_{\mathbf{k}\lambda} \right) = \frac{i}{c} \omega_{\mathbf{k}} A_{\mathbf{k}\lambda} e^{i\mathbf{kr}} \left(\hat{\mathbf{k}} \times \hat{\mathbf{e}}_{\mathbf{k}\lambda} \right). \tag{4.14}$$

Therefore the magnetic field is

$$\mathbf{B}(\mathbf{r},t) = \boldsymbol{\nabla} \times \mathbf{A}(\mathbf{r},t) = \frac{i}{c} \sum_{\mathbf{k},\lambda} \omega_{\mathbf{k}} \left(\hat{\mathbf{k}} \times \hat{\mathbf{e}}_{\mathbf{k}\lambda} \right) \Big[A_{\mathbf{k}\lambda} \exp\left[i(\mathbf{kr} - \omega_{\mathbf{k}} t) \right] - A^*_{\mathbf{k}\lambda} \exp\left[-i(\mathbf{kr} - \omega_{\mathbf{k}} t) \right] \Big]. \tag{4.15}$$

To get the Hamiltonian we need to find the field energy W in the cavity of volume V in terms of the electric and magnetic fields. Electromagnetic theory tells us that W is given by

$$W = \frac{1}{8\pi} \int_V (\mathbf{E}^2 + \mathbf{B}^2)\, \mathrm{d}V. \tag{4.16}$$

Let's first perform the integral of the $\mathbf{E}^2$ term. Denoting $\mathrm{A}_{\mathbf{k}\lambda} = A_{\mathbf{k}\lambda} e^{-i\omega_{\mathbf{k}} t}$, where $A_{\mathbf{k}\lambda}$ is a constant amplitude for the mode $(\mathbf{k}, \lambda)$, the expression for $\mathbf{E}$ can be written as

$$\mathbf{E}(\mathbf{r},t) = \frac{i}{c} \sum_{\mathbf{k},\lambda} \omega_{\mathbf{k}} \hat{\mathbf{e}}_{\mathbf{k}\lambda} \Big[\mathrm{A}_{\mathbf{k}\lambda}\, e^{i\mathbf{kr}} - \mathrm{A}^*_{\mathbf{k}\lambda}\, e^{-i\mathbf{kr}} \Big]. \tag{4.17}$$

When squaring this sum we get

$$\begin{aligned} \mathbf{E} \cdot \mathbf{E} &= -\frac{1}{c^2} \sum_{\mathbf{k},\lambda} \sum_{\mathbf{k}',\lambda'} \omega_{\mathbf{k}} \omega_{\mathbf{k}'} \left(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{e}}_{\mathbf{k}',\lambda'} \right) \Big[\mathrm{A}_{\mathbf{k}\lambda} e^{i\mathbf{kr}} - \mathrm{A}^*_{\mathbf{k}\lambda} e^{-i\mathbf{kr}} \Big] \Big[\mathrm{A}_{\mathbf{k}'\lambda'} e^{i\mathbf{k}'\mathbf{r}} - \mathrm{A}^*_{\mathbf{k}'\lambda'} e^{-i\mathbf{k}'\mathbf{r}} \Big] \\ &= -\frac{1}{c^2} \sum_{\mathbf{k},\lambda} \sum_{\mathbf{k}',\lambda'} \omega_{\mathbf{k}} \omega_{\mathbf{k}'} \left(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{e}}_{\mathbf{k}',\lambda'} \right) \Big[\mathrm{A}_{\mathbf{k}\lambda} \mathrm{A}_{\mathbf{k}'\lambda'} e^{i(\mathbf{k}+\mathbf{k}')\mathbf{r}} - \mathrm{A}_{\mathbf{k}\lambda} \mathrm{A}^*_{\mathbf{k}'\lambda'} e^{i(\mathbf{k}-\mathbf{k}')\mathbf{r}} + \mathrm{cc} \Big]. \end{aligned} \tag{4.18}$$

where cc stands for the complex conjugate of the terms within the square brackets. The volume integrals can be evaluated using the boundary conditions (4.1). The results are just Kronecker deltas multiplied by the volume of the box,

$$\int_V e^{\pm i(\mathbf{k}+\mathbf{k}')\mathbf{r}}\, \mathrm{d}V = \delta_{\mathbf{k}',-\mathbf{k}} V, \qquad \int_V e^{\pm i(\mathbf{k}-\mathbf{k}')\mathbf{r}}\, \mathrm{d}V = \delta_{\mathbf{k}'\mathbf{k}} V. \tag{4.19}$$

A further simplification can be made by taking into account the orthogonality condition on the polarization vectors,

$$\hat{\mathbf{e}}_{\mathbf{k}\lambda} \hat{\mathbf{e}}_{\mathbf{k}\lambda'} = \delta_{\lambda\lambda'}, \qquad \hat{\mathbf{e}}_{\mathbf{k}\lambda} \hat{\mathbf{e}}_{-\mathbf{k}\lambda'} = -\delta_{\lambda\lambda'}, \tag{4.20}$$

together with the relation $\omega_{\mathbf{k}} = \omega_{-\mathbf{k}}$. Note that we have implicitly used $\hat{\mathbf{e}}_{-\mathbf{k}\lambda'} = -\hat{\mathbf{e}}_{\mathbf{k}\lambda'}$ to derive the last relation in (4.20). Taking all this into account we have

$$\begin{aligned} \int_V \mathbf{E} \cdot \mathbf{E}\, \mathrm{d}V &= -\frac{1}{c^2} \int_V \sum_{\mathbf{k},\lambda} \sum_{\mathbf{k}',\lambda'} \omega_{\mathbf{k}} \omega_{\mathbf{k}'} \left(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{e}}_{\mathbf{k}',\lambda'} \right) \Big[\mathrm{A}_{\mathbf{k}\lambda} \mathrm{A}_{\mathbf{k}'\lambda'} e^{i(\mathbf{k}+\mathbf{k}')\mathbf{r}} - \mathrm{A}_{\mathbf{k}\lambda} \mathrm{A}^*_{\mathbf{k}'\lambda'} e^{i(\mathbf{k}-\mathbf{k}')\mathbf{r}} + c.c \Big]. \\ &= -\frac{V}{c^2} \sum_{\mathbf{k},\lambda} \sum_{\mathbf{k}',\lambda'} \omega_{\mathbf{k}} \omega_{\mathbf{k}'} \left(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{e}}_{\mathbf{k}',\lambda'} \right) \Big[\mathrm{A}_{\mathbf{k}\lambda} \mathrm{A}_{\mathbf{k}'\lambda'} \delta_{\mathbf{k},-\mathbf{k}'} - \mathrm{A}_{\mathbf{k}\lambda} \mathrm{A}^*_{\mathbf{k}'\lambda'} \delta_{\mathbf{k},\mathbf{k}'} + c.c \Big] \\ &= -\frac{V}{c^2} \sum_{\mathbf{k},\lambda} \sum_{\lambda'} \omega_{\mathbf{k}}^2 \Big[\left(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{e}}_{-\mathbf{k},\lambda'} \right) \mathrm{A}_{\mathbf{k}\lambda} \mathrm{A}_{-\mathbf{k}\lambda'} - \left(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{e}}_{\mathbf{k},\lambda'} \right) \mathrm{A}_{\mathbf{k}\lambda} \mathrm{A}^*_{\mathbf{k}\lambda'} + c.c \Big] \\ &= -\frac{V}{c^2} \sum_{\mathbf{k},\lambda} \sum_{\lambda'} \omega_{\mathbf{k}}^2 \Big[-\delta_{\lambda\lambda'} \mathrm{A}_{\mathbf{k}\lambda} \mathrm{A}_{-\mathbf{k}\lambda'} - \delta_{\lambda\lambda'} \mathrm{A}_{\mathbf{k}\lambda} \mathrm{A}^*_{\mathbf{k}\lambda'} + c.c \Big] \\ &= \frac{V}{c^2} \sum_{\mathbf{k},\lambda} \omega_{\mathbf{k}}^2 \Big[\mathrm{A}_{\mathbf{k}\lambda} \mathrm{A}^*_{\mathbf{k}\lambda} + \mathrm{A}_{\mathbf{k}\lambda} \mathrm{A}_{-\mathbf{k},\lambda} + c.c. \Big]. \end{aligned} \tag{4.21}$$

The first term is equal to its conjugate and will have a double contribution. When making explicit the complex conjugates we have:

$$\int_V \mathbf{E}\cdot\mathbf{E}\,\mathrm{d}V = \frac{V}{c^2}\sum_{\mathbf{k},\lambda}\omega_{\mathbf{k}}^2\left[2\mathrm{A}_{\mathbf{k}\lambda}\mathrm{A}^*_{\mathbf{k}\lambda} + \mathrm{A}_{\mathbf{k}\lambda}\mathrm{A}_{-\mathbf{k},\lambda} + \mathrm{A}^*_{\mathbf{k}\lambda}\mathrm{A}^*_{-\mathbf{k},\lambda}\right] \tag{4.22}$$

We need now to perform the integral of the $\mathbf{B}^2$ term, which is given by

$$\begin{aligned}\mathbf{B}\cdot\mathbf{B} &= -\frac{1}{c^2}\sum_{\mathbf{k},\lambda}\sum_{\mathbf{k}',\lambda'}\omega_{\mathbf{k}}\omega_{\mathbf{k}'}\left((\hat{\mathbf{k}}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda})\cdot(\hat{\mathbf{k}}'\times\hat{\mathbf{e}}_{\mathbf{k}',\lambda'})\right)\left[\mathrm{A}_{\mathbf{k}\lambda}e^{i\mathbf{k}\mathbf{r}} - \mathrm{A}^*_{\mathbf{k}\lambda}e^{-i\mathbf{k}\mathbf{r}}\right]\left[\mathrm{A}_{\mathbf{k}'\lambda'}e^{i\mathbf{k}'\mathbf{r}} - \mathrm{A}^*_{\mathbf{k}'\lambda'}e^{-i\mathbf{k}'\mathbf{r}}\right]\\ &= -\frac{1}{c^2}\sum_{\mathbf{k},\lambda}\sum_{\mathbf{k}',\lambda'}\omega_{\mathbf{k}}\omega_{\mathbf{k}'}\left((\hat{\mathbf{k}}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda})\cdot(\hat{\mathbf{k}}'\times\hat{\mathbf{e}}_{\mathbf{k}',\lambda'})\right)\left[\mathrm{A}_{\mathbf{k}\lambda}\mathrm{A}_{\mathbf{k}'\lambda'}e^{i(\mathbf{k}+\mathbf{k}')\mathbf{r}} - \mathrm{A}_{\mathbf{k}\lambda}\mathrm{A}^*_{\mathbf{k}'\lambda'}e^{i(\mathbf{k}-\mathbf{k}')\mathbf{r}} + c.c\right].\end{aligned} \tag{4.23}$$

The vector identity

$$(\mathbf{A}\times\mathbf{B})(\mathbf{C}\times\mathbf{D}) = (\mathbf{AC})(\mathbf{BD}) - (\mathbf{AD})(\mathbf{BC}), \tag{4.24}$$

together with the orthogonality condition, $\hat{\mathbf{e}}_{\mathbf{k}\lambda}\hat{\mathbf{k}} = 0$, allow to find that:

$$\left(\hat{\mathbf{k}}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda}\right)\left(\hat{\mathbf{k}}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda'}\right) = \left(\hat{\mathbf{k}}\hat{\mathbf{k}}\right)\left(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\hat{\mathbf{e}}_{\mathbf{k}\lambda'}\right) - \left(\hat{\mathbf{k}}\hat{\mathbf{e}}_{\mathbf{k}\lambda'}\right)\left(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\hat{\mathbf{k}}\right) = \delta_{\lambda\lambda'}. \tag{4.25}$$

The magnetic field contribution to the energy can be computed following the same steps which lead us to equation (4.21). The result is

$$\int_V \mathbf{B}\cdot\mathbf{B}\,\mathrm{d}V = \frac{V}{c^2}\sum_{\mathbf{k},\lambda}\omega_{\mathbf{k}}^2\left[2\mathrm{A}_{\mathbf{k}\lambda}\mathrm{A}^*_{\mathbf{k}\lambda} - \mathrm{A}_{\mathbf{k}\lambda}\mathrm{A}_{-\mathbf{k},\lambda} - \mathrm{A}^*_{\mathbf{k}\lambda}\mathrm{A}^*_{-\mathbf{k},\lambda}\right] \tag{4.26}$$

We see that the cross terms cancel out those of the electric field contribution. We are left then with the following expression for the field energy,

$$W = \frac{V}{2\pi c^2}\sum_{\mathbf{k},\lambda}\omega_{\mathbf{k}}^2\mathrm{A}_{\mathbf{k}\lambda}\mathrm{A}^*_{\mathbf{k}\lambda} = \frac{V}{2\pi c^2}\sum_{\mathbf{k},\lambda}\omega_{\mathbf{k}}^2 A_{\mathbf{k}\lambda}A^*_{\mathbf{k}\lambda}. \tag{4.27}$$

In the last step we have replaced the time-dependent $\mathrm{A}_{\mathbf{k}\lambda} = A_{\mathbf{k}\lambda}e^{-i\omega_{\mathbf{k}}t}$ amplitudes by the constant $A_{\mathbf{k}\lambda}$ amplitudes because the time dependency cancels out in the product, thus $\mathrm{A}_{\mathbf{k}\lambda}\mathrm{A}^*_{\mathbf{k}\lambda} = A_{\mathbf{k}\lambda}A^*_{\mathbf{k}\lambda}$. The energy of an electromagnetic radiation field is just the sum of the energies of the individual modes, **does not depend on time and is equally shared by the electric and magnetic fields**.

We introduce now a new set of variables through the relations

$$\mathrm{A}_{\mathbf{k}\lambda} = \frac{c}{\omega_{\mathbf{k}}}\sqrt{\frac{\pi}{V}}\left(\omega_{\mathbf{k}}Q_{\mathbf{k}\lambda} + iP_{\mathbf{k}\lambda}\right), \qquad \mathrm{A}^*_{\mathbf{k}\lambda} = \frac{c}{\omega_{\mathbf{k}}}\sqrt{\frac{\pi}{V}}\left(\omega_{\mathbf{k}}Q_{\mathbf{k}\lambda} - iP_{\mathbf{k}\lambda}\right), \tag{4.28}$$

$$Q_{\mathbf{k}\lambda} = \frac{1}{2c}\sqrt{\frac{V}{\pi}}\left(\mathrm{A}_{\mathbf{k}\lambda} + \mathrm{A}^*_{\mathbf{k}\lambda}\right), \qquad P_{\mathbf{k}\lambda} = -i\frac{\omega_{\mathbf{k}}}{2c}\sqrt{\frac{V}{\pi}}\left(\mathrm{A}_{\mathbf{k}\lambda} - \mathrm{A}^*_{\mathbf{k}\lambda}\right). \tag{4.29}$$

The electromagnetic energy can then be written as

$$W = \frac{1}{2}\sum_{\mathbf{k},\lambda}\left(P^2_{\mathbf{k}\lambda} + \omega^2_{\mathbf{k}}Q^2_{\mathbf{k}\lambda}\right). \tag{4.30}$$

We'll see now that the variables $P_{\mathbf{k}\lambda}$ and $Q_{\mathbf{k}\lambda}$ satisfy the Hamilton equations, which means that they are canonical variables and the expression above can be interpreted as the Hamiltonian of the electromagnetic field. Indeed, note that

$$\dot{Q}_{\mathbf{k}\lambda} = P_{\mathbf{k}\lambda}, \qquad \dot{P}_{\mathbf{k}\lambda} = -\omega^2_{\mathbf{k}}Q_{\mathbf{k}\lambda}, \tag{4.31}$$

therefore,

$$\frac{\partial W}{\partial Q_{\mathbf{k}\lambda}} = \omega_{\mathbf{k}}^2 Q_{\mathbf{k}\lambda} = -\dot{P}_{\mathbf{k}\lambda}, \qquad \frac{\partial W}{\partial P_{\mathbf{k}\lambda}} = P_{\mathbf{k}\lambda} = \dot{Q}_{\mathbf{k}\lambda}. \tag{4.32}$$

It can also be verified that the Maxwell equations in free space

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \qquad \boldsymbol{\nabla} \times \mathbf{B} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}, \tag{4.33}$$

are in fact the Hamilton equations (4.32) .

The Hamiltonian (4.30) is actually identical to the Hamiltonian of a collection of simple (independent) harmonic oscillators. Indeed, remember that the Hamiltonian of a one-dimensional harmonic oscillator of mass m and frequency ω is

$$\mathcal{H} = \frac{p^2}{2m} + \frac{m}{2}\omega^2 q^2, \tag{4.34}$$

where p is the momentum and q is the displacement. If we now let

$$P^2 = \frac{p^2}{m}, \qquad Q^2 = mq^2, \tag{4.35}$$

the Hamiltonian becomes

$$\mathcal{H} = \frac{1}{2}\Big(P^2 + \omega^2 Q^2\Big). \tag{4.36}$$

Comparing equations (4.36) and (4.30) we see that each mode of the radiation field is formally equivalent to a single harmonic oscillator.

4.2 Quantization of the radiation field.

We now apply all we know about the quantum mechanical treatment of the harmonic oscillator. The quantities $P_{\mathbf{k}\lambda}$ and $Q_{\mathbf{k}\lambda}$ are postulated to be operators that satisfy the following commutation relations,

$$[Q_{\mathbf{k}\lambda}, P_{\mathbf{k}'\lambda'}] = i\hbar\,\delta_{\mathbf{k}\mathbf{k}'}\delta_{\lambda\lambda'}, \qquad [Q_{\mathbf{k}\lambda}, Q_{\mathbf{k}'\lambda'}] = [P_{\mathbf{k}\lambda}, P_{\mathbf{k}'\lambda'}] = 0. \tag{4.37}$$

The annihilation and creation operators are defined as

$$a_{\mathbf{k}\lambda} = \frac{1}{\sqrt{2\hbar\omega_{\mathbf{k}}}}(\omega_{\mathbf{k}} Q_{\mathbf{k}\lambda} + iP_{\mathbf{k}\lambda}), \qquad a_{\mathbf{k}\lambda}^\dagger = \frac{1}{\sqrt{2\hbar\omega_{\mathbf{k}}}}(\omega_{\mathbf{k}} Q_{\mathbf{k}\lambda} - iP_{\mathbf{k}\lambda}), \tag{4.38}$$

$$Q_{\mathbf{k}\lambda} = \sqrt{\frac{\hbar}{2\omega_{\mathbf{k}}}}\left(a_{\mathbf{k}\lambda} + a_{\mathbf{k}\lambda}^\dagger\right), \qquad P_{\mathbf{k}\lambda} = \frac{1}{2i}\sqrt{2\hbar\omega_{\mathbf{k}}}\left(a_{\mathbf{k}\lambda} - a_{\mathbf{k}\lambda}^\dagger\right). \tag{4.39}$$

Satisfying the commutation rules[2]

$$\left[a_{\mathbf{k}\lambda}, a_{\mathbf{k}'\lambda'}^\dagger\right] = \delta_{\mathbf{k}\mathbf{k}'}\delta_{\lambda\lambda'}, \qquad [a_{\mathbf{k}\lambda}, a_{\mathbf{k}'\lambda'}] = \left[a_{\mathbf{k}\lambda}^\dagger, a_{\mathbf{k}'\lambda'}^\dagger\right] = 0. \tag{4.40}$$

The Hamiltonian (4.30) can be written in terms of these operators as

$$\mathcal{H} = \sum_{\mathbf{k},\lambda} \mathcal{H}_{\mathbf{k}\lambda} = \sum_{\mathbf{k},\lambda} \hbar\omega_{\mathbf{k}}\left(a_{\mathbf{k}\lambda}^\dagger a_{\mathbf{k}\lambda} + \frac{1}{2}\right) = \sum_{\mathbf{k},\lambda} \hbar\omega_{\mathbf{k}}\left(N_{\mathbf{k}\lambda} + \frac{1}{2}\right), \tag{4.41}$$

where $N_{\mathbf{k}\lambda} = a_{\mathbf{k}\lambda}^\dagger a_{\mathbf{k}\lambda}$ is the number operator for the mode $\mathbf{k}$ and polarization λ. The eigenvalues of $N_{\mathbf{k}\lambda}$ are $n_{\mathbf{k}\lambda} = 0, 1, 2, \ldots$ Therefore it satisfies the following eigenvalue equation,

$$N_{\mathbf{k}\lambda}\,|n_{\mathbf{k}\lambda}\rangle = n_{\mathbf{k}\lambda}\,|n_{\mathbf{k}\lambda}\rangle\,, \qquad n_{\mathbf{k}\lambda} = 0, 1, 2, \ldots \tag{4.42}$$

[2]These commutation relations are characteristic of boson operators, but a discussion of this matter here would take us too far afield.

where $|n_{\mathbf{k}\lambda}\rangle$ represents an eigenstate of $N_{\mathbf{k}\lambda}$. The possible energies of the radiation field are given by the eigenvalues of $\mathcal{H}$, which are

$$E = \sum_{\mathbf{k},\lambda} E_{\mathbf{k}\lambda} = \sum_{\mathbf{k},\lambda} \hbar\omega_{\mathbf{k}} \left(n_{\mathbf{k}\lambda} + \frac{1}{2} \right). \tag{4.43}$$

The annihilation and creation operators act on the Hamiltonian eigenstates according to

$$a_{\mathbf{k}\lambda} |n_{\mathbf{k}\lambda}\rangle = \sqrt{n_{\mathbf{k}\lambda}}\, |n_{\mathbf{k}\lambda} - 1\rangle\,, \qquad a^\dagger_{\mathbf{k}\lambda} |n_{\mathbf{k}\lambda}\rangle = \sqrt{n_{\mathbf{k}\lambda} + 1}\, |n_{\mathbf{k}\lambda} + 1\rangle\,. \tag{4.44}$$

The operator $a^\dagger_{\mathbf{k}\lambda}$ creates a photon with propagation vector $\mathbf{k}$, polarization vector $\hat{\mathbf{e}}_{\mathbf{k}\lambda}$, frequency $\omega_{\mathbf{k}}$, momentum $\hbar\mathbf{k}$ and energy $\hbar\omega_{\mathbf{k}}$. The operator $a_{\mathbf{k}\lambda}$ destroys a photon of the same type. The quantities $n_{\mathbf{k}\lambda}$, called occupation numbers, are the eigenvalues of the number operator $N_{\mathbf{k}\lambda}$ and give the number of photons in the radiation field in the mode $\mathbf{k}$ and polarization λ. A complete description of the radiation field can be given by specifying the occupation numbers $n_{\mathbf{k}\lambda}$. Since the radiation field is a collection of independent modes, the tensor product of all eigenstates $|n_{\mathbf{k}\lambda}\rangle$ is an eigenstate of the total Hamiltonian.

By performing the tensor product of single-particle Hilbert space states we build the states of the Fock space (also a Hilbert space). For two single-particle states such as $|\psi\rangle$ and $|\phi\rangle$, for example, we would have

$$|\psi\rangle \otimes |\phi\rangle \equiv |\psi, \phi\rangle\,. \tag{4.45}$$

Omitting possible internal degrees of freedom, in the position-space representation we would have just the product of the two wave functions,

$$\langle \mathbf{r}_1, \mathbf{r}_2 \,|\, \psi, \phi\rangle = (\langle \mathbf{r}_1| \otimes \langle \mathbf{r_2}|)\,(|\psi\rangle \otimes |\phi\rangle) = \langle \mathbf{r}_1 \,|\, \psi\rangle\, \langle \mathbf{r}_2 \,|\, \phi\rangle = \psi(\mathbf{r}_1)\phi(\mathbf{r_2}). \tag{4.46}$$

For identical particles, we should only work with the symmetric subspace (for bosons) or the antisymmetric subspace (for fermions). Since photons are bosons, we will only consider symmetric combinations of states containing N photons, $n_{\mathbf{k}\lambda}$ for each propagation and polarization vector. For the sake of illustration let's consider a 3-photon state, with $n_{\mathbf{k}_1\lambda_1} = 2$ and $n_{\mathbf{k}_2\lambda_2} = 1$. If photons were *distinguishable* particles we would build the state

$$\left|\psi^1_{\mathbf{k}_1\lambda_1}, \psi^2_{\mathbf{k}_1\lambda_1}, \psi^3_{\mathbf{k}_2\lambda_2}\right\rangle. \tag{4.47}$$

But since they are *not distinguishable*, the symmetric state to use is

$$\begin{aligned} &|n_{\mathbf{k}_1\lambda_1} = 2, n_{\mathbf{k}_2\lambda_2} = 1\rangle \\ &\quad = \frac{1}{\sqrt{3}} \Big(\left|\psi^1_{\mathbf{k}_1\lambda_1}, \psi^2_{\mathbf{k}_1\lambda_1}, \psi^3_{\mathbf{k}_2\lambda_2}\right\rangle + \left|\psi^1_{\mathbf{k}_1\lambda_1}, \psi^2_{\mathbf{k}_2\lambda_2}, \psi^3_{\mathbf{k}_1\lambda_1}\right\rangle + \left|\psi^1_{\mathbf{k}_2\lambda_2}, \psi^2_{\mathbf{k}_1\lambda_1}, \psi^3_{\mathbf{k}_1\lambda_1}\right\rangle \Big). \end{aligned} \tag{4.48}$$

Fortunately, we will not need to do all this symmetrization by hand, since the commutation relations (4.40) automatically take care of this when applying the corresponding creation operators on the vacuum state.[3]

We will write a general state of the radiation field as

$$|n_{\mathbf{k}_1\lambda_1}\rangle \otimes |n_{\mathbf{k}_2\lambda_2}\rangle \otimes \ldots \otimes |n_{\mathbf{k}_i\lambda_i}\rangle \otimes \ldots \equiv |n_{\mathbf{k}_1\lambda_1}, n_{\mathbf{k}_2\lambda_2}, \ldots, n_{\mathbf{k}_i\lambda_i}, \ldots\rangle \tag{4.49}$$

To simplify the notation we can make the replacement $n_{\mathbf{k}_i\lambda_i} \equiv n_i$ and write the eigenstate as

$$|n_1, n_2, \ldots, n_i, \ldots\rangle\,. \tag{4.50}$$

The eigenstates satisfy the orthogonality relation,

$$\langle n'_1, n'_2, \ldots, n'_i, \ldots \,|\, n_1, n_2, \ldots, n_i, \ldots\rangle = \delta_{n'_1 n_1} \delta_{n'_2 n_2} \ldots \delta_{n'_i n_i} \ldots \tag{4.51}$$

[3] See the first two problems at the end of this chapter.

Replacing the indices $\mathbf{k}_i\lambda_i$ by a single index i in the operators, the obvious generalization of (4.44) for a general many–photon state is

$$\begin{aligned} a_i \left|n_1, n_2, \ldots, n_i, \ldots\right\rangle &= \sqrt{n_i}\left|n_1, n_2, \ldots, n_i - 1, \ldots\right\rangle, \\ a_i^\dagger \left|n_1, n_2, \ldots, n_i, \ldots\right\rangle &= \sqrt{n_i + 1}\left|n_1, n_2, \ldots, n_i + 1, \ldots\right\rangle, \\ a_i \left|n_1, n_2, \ldots, 0_i, \ldots\right\rangle &= 0. \end{aligned} \tag{4.52}$$

For the number operators and the total energy we have

$$N_i \left|n_1, n_2, \ldots, n_i, \ldots\right\rangle = n_i \left|n_1, n_2, \ldots, n_i, \ldots\right\rangle, \qquad E = \sum_i \hbar\omega_i \left(n_i + \frac{1}{2}\right). \tag{4.53}$$

Note that even when the number of photons for a given mode is zero, $n_{\mathbf{k}\lambda} = 0$, the energy associated with that mode is not zero, but $\hbar\omega_{\mathbf{k}}/2$. When all possible modes are taken into account this leads to an *infinite zero-point energy*. A proper discussion of this matter goes beyond the scope of this course. Let us simply say here that a consistent description of all physical processes in atomic physics can be obtained by ignoring this infinite zero-point energy of the radiation field. Only energy differences relative to a conveniently defined *ground state* really matter.

The rotational properties of the classical radiation field still apply also in the quantum domain. The photon has an intrinsic spin $S = 1$ with only two possible values for the third component, $m_S = \pm 1$, due to the transversality condition. The spin of a photon can be regarded as a vector pointing parallel (positive helicity) or anti-parallel (negative helicity) to its momentum vector $\hbar\mathbf{k}$. This can be shown to be associated to the fact that photons are massless particles.

The vector potential $\mathbf{A}(\mathbf{r}, t)$ given by equation (4.6) must be now interpreted as an operator. According to the definitions of the creation and annihilation operators, comparing equations (4.29) and (4.39) we see that the following replacements need to be done

$$A_{\mathbf{k}\lambda} \longrightarrow \sqrt{\frac{2\pi\hbar c^2}{V\omega_{\mathbf{k}}}}\, a_{\mathbf{k}\lambda}, \qquad A^*_{\mathbf{k}\lambda} \longrightarrow \sqrt{\frac{2\pi\hbar c^2}{V\omega_{\mathbf{k}}}}\, a^\dagger_{\mathbf{k}\lambda}, \tag{4.54}$$

which leads to

$$\mathbf{A}_{\mathrm{H}}(\mathbf{r}, t) = \sum_{\mathbf{k},\lambda} \sqrt{\frac{2\pi\hbar c^2}{V\omega_{\mathbf{k}}}}\, \hat{\mathbf{e}}_{\mathbf{k}\lambda}\Big[a_{\mathbf{k}\lambda} \exp\left[i(\mathbf{kr} - \omega_{\mathbf{k}} t)\right] + a^\dagger_{\mathbf{k}\lambda} \exp\left[-i(\mathbf{kr} - \omega_{\mathbf{k}} t)\right]\Big]. \tag{4.55}$$

The subscript H in $\mathbf{A}_{\mathrm{H}}(\mathbf{r}, t)$ indicates that we are working in the Heisenberg representation (time-dependent operators, time-independent states). From the study of the harmonic oscillator you may recall that operators of the form $ae^{-i\omega t}$ and $a^\dagger e^{i\omega t}$ refer to the Heisenberg representation, while the operators a and $a^\dagger$ are referred to the Schrödinger representation. The vector potential in the Schrödinger representation is

$$\mathbf{A}(\mathbf{r}) = \sum_{\mathbf{k},\lambda} \sqrt{\frac{2\pi\hbar c^2}{V\omega_{\mathbf{k}}}}\, \hat{\mathbf{e}}_{\mathbf{k}\lambda}\Big[a_{\mathbf{k}\lambda} e^{i\mathbf{kr}} + a^\dagger_{\mathbf{k}\lambda} e^{-i\mathbf{kr}}\Big]. \tag{4.56}$$

Note that the vector potential operator is Hermitian, although the creation and annihilation operators themselves are not. The electric and magnetic field operators in terms of $a_{\mathbf{k}\lambda}$ and $a^\dagger_{\mathbf{k}\lambda}$ are, in the Schrödinger representation

$$\mathbf{E}(\mathbf{r}) = i \sum_{\mathbf{k},\lambda} \sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}}{V}}\, \hat{\mathbf{e}}_{\mathbf{k}\lambda}\Big[a_{\mathbf{k}\lambda} e^{i\mathbf{kr}} - a^\dagger_{\mathbf{k}\lambda} e^{-i\mathbf{kr}}\Big]. \tag{4.57}$$

$$\mathbf{B}(\mathbf{r}) = i \sum_{\mathbf{k},\lambda} \sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}}{V}} \left(\hat{\mathbf{k}} \times \hat{\mathbf{e}}_{\mathbf{k}\lambda}\right) \left[a_{\mathbf{k}\lambda} e^{i\mathbf{k}\mathbf{r}} - a^{\dagger}_{\mathbf{k}\lambda} e^{-i\mathbf{k}\mathbf{r}}\right]. \tag{4.58}$$

From these expressions, it is clear that the electric and magnetic fields are orthogonal to each other and orthogonal to the propagation vector.

To see how the ordinary wave functions of quantum mechanics arise from the quantum field theory formalism let us consider the meaning of the state $\mathbf{A}(\mathbf{r})\,|0\rangle$, where $|0\rangle$ represents the state with no photons (the vacuum). Noting that $a_{\mathbf{k}\lambda}\,|0\rangle = 0$, when applying the field operator (4.56) to the vacuum we obtain

$$\mathbf{A}(\mathbf{r})\,|0\rangle = \sum_{\mathbf{k},\lambda} \sqrt{\frac{2\pi\hbar c^2}{V\omega_{\mathbf{k}}}}\, \hat{\mathbf{e}}_{\mathbf{k}\lambda} a^{\dagger}_{\mathbf{k}\lambda} e^{-i\mathbf{k}\mathbf{r}}\ |0\rangle \ = \ \sum_{\mathbf{k},\lambda} \sqrt{\frac{2\pi\hbar c^2}{V\omega_{\mathbf{k}}}}\, \hat{\mathbf{e}}_{\mathbf{k}\lambda} e^{-i\mathbf{k}\mathbf{r}}\ |\mathbf{k}\lambda\rangle\,. \tag{4.59}$$

This is a linear superposition of single-photon states with momentum $\hbar\mathbf{k}$ and polarization λ. The sum over $\mathbf{k}$ would turn into an integral if we let the size of the normalization box go to infinity. In fact, this superposition is very similar to the familiar non-relativistic expression for the eigenstate of position as a linear combination of momentum eigenstates, which in one dimension is[4]

$$|x\rangle = \int \mathrm{d}p\ |p\rangle\langle p|x\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int \mathrm{d}p\, e^{-ipx}\,|p\rangle\,. \tag{4.60}$$

It seems then natural to claim that the field operator $\mathbf{A}(\mathbf{r})$ acting on the vacuum *creates a particle* (a photon) at position $\mathbf{r}$. This interpretation can be confirmed by computing the matrix element $\langle 0\,|\,\mathbf{A}(\mathbf{r})\,|\,\mathbf{k}\lambda\rangle$. This is

$$\langle 0\,|\,\mathbf{A}(\mathbf{r})\,|\,\mathbf{k}\lambda\rangle \ = \ \left\langle 0 \left| \sum_{\mathbf{k}',\lambda'} \sqrt{\frac{2\pi\hbar c^2}{V\omega_{\mathbf{k}'}}}\, \hat{\mathbf{e}}_{\mathbf{k}'\lambda'} \left[a_{\mathbf{k}'\lambda'} e^{i\mathbf{k}'\mathbf{r}} + a^{\dagger}_{\mathbf{k}'\lambda'} e^{-i\mathbf{k}'\mathbf{r}}\right] a^{\dagger}_{\mathbf{k}\lambda} \right| 0 \right\rangle. \tag{4.61}$$

Now, the term containing the operator product $a^{\dagger}_{\mathbf{k}'\lambda'} a^{\dagger}_{\mathbf{k}\lambda}$ will give zero, since $\langle 0|\, a^{\dagger}_{\mathbf{k}'\lambda'} = 0$. In addition, the commutation relations (4.40) allows us to write $a_{\mathbf{k}'\lambda'} a^{\dagger}_{\mathbf{k}\lambda} = a^{\dagger}_{\mathbf{k}\lambda} a_{\mathbf{k}'\lambda'} + \delta_{\mathbf{k}\mathbf{k}'}\delta_{\lambda\lambda'}$. Therefore

$$\begin{aligned}
\langle 0\,|\,\mathbf{A}(\mathbf{r})\,|\,\mathbf{k}\lambda\rangle &= \left\langle 0 \left| \sum_{\mathbf{k}',\lambda'} \sqrt{\frac{2\pi\hbar c^2}{V\omega_{\mathbf{k}'}}}\, \hat{\mathbf{e}}_{\mathbf{k}'\lambda'}\, e^{i\mathbf{k}'\mathbf{r}}\, a_{\mathbf{k}'\lambda'} a^{\dagger}_{\mathbf{k}\lambda} \right| 0 \right\rangle \\
&= \left\langle 0 \left| \sum_{\mathbf{k}',\lambda'} \sqrt{\frac{2\pi\hbar c^2}{V\omega_{\mathbf{k}'}}}\, \hat{\mathbf{e}}_{\mathbf{k}'\lambda'}\, e^{i\mathbf{k}'\mathbf{r}} \left(a^{\dagger}_{\mathbf{k}\lambda} a_{\mathbf{k}'\lambda'} + \delta_{\mathbf{k}\mathbf{k}'}\delta_{\lambda\lambda'}\right) \right| 0 \right\rangle.
\end{aligned} \tag{4.62}$$

Using again the vacuum condition $\langle 0|\, a^{\dagger}_{\mathbf{k}\lambda} = 0$, performing the sum with the Kronecker-delta constraints, and using the vacuum normalization $\langle 0|0\rangle = 1$, we get

$$\langle 0\,|\,\mathbf{A}(\mathbf{r})\,|\,\mathbf{k}\lambda\rangle \ = \ \sqrt{\frac{2\pi\hbar c^2}{V\omega_{\mathbf{k}}}}\, \hat{\mathbf{e}}_{\mathbf{k}\lambda}\, e^{i\mathbf{k}\mathbf{r}}. \tag{4.63}$$

This is the plane-wave function for a photon with momentum $\hbar\mathbf{k}$ and polarization λ. In other words, this is the position-space representation of the single-particle wave function of the state $|\mathbf{k}\lambda\rangle$. Remember that in non-relativistic quantum mechanics $\langle\mathbf{r}|\mathbf{p}\rangle \propto e^{i\mathbf{p}\mathbf{r}/\hbar}$ is the wave function of the state $|\mathbf{p}\rangle$, and compare $\langle\mathbf{r}|$ with $\langle 0|\mathbf{A}(\mathbf{r})$. This confirms the interpretation that the field operator $\mathbf{A}(\mathbf{r})$ acting on the vacuum creates a particle at position $\mathbf{r}$. In the quantum field theory formalism, the vacuum to one-particle matrix elements of the field operators are just the familiar wave functions of single-particle quantum mechanics.

[4] Remember that $\int \mathrm{d}p\ |p\rangle\langle p| = 1$ and $\langle p|x\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{-ipx}$.

4.3 Interaction Hamiltonian and matrix elements.

The total Hamiltonian describing the interaction of an atom with a radiation field can be written as a sum of three terms

$$\mathcal{H} = \mathcal{H}_{\rm rad} + \mathcal{H}_{\rm atom} + \mathcal{H}_{\rm int}, \tag{4.64}$$

where

$$\mathcal{H}_{\rm rad} = \sum_{\mathbf{k},\lambda} \hbar\omega_{\mathbf{k}} \left(a^\dagger_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} + \frac{1}{2} \right), \tag{4.65}$$

is the Hamiltonian of the radiation field and

$$\mathcal{H}_{\rm atom} = \sum_i \frac{p_i^2}{2m} + U, \tag{4.66}$$

is the atomic Hamiltonian. Here the sum runs over the atomic electrons and U contains all the relevant terms to define the atomic state (Coulomb interaction with the nucleus, Coulomb repulsion among the electrons, spin-orbit interaction, external fields, etc.). To write an expression for $\mathcal{H}_{\rm int}$ we go back to the Schrödinger equation (1.87) and pay attention to the terms containing the vector potential,

$$\begin{aligned}(E' + e\phi)\,\psi \;=\; \Bigg[& \frac{1}{2m}\left(\mathbf{p} + \frac{e}{c}\mathbf{A}\right)^2 + \frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A}) \\ & - \frac{p^4}{8m^3c^2} - \frac{\hbar^2 e}{8m^2c^2}\boldsymbol{\nabla}^2\phi - \frac{\hbar e}{4m^2c^2}\boldsymbol{\sigma}\left[(\boldsymbol{\nabla}\phi)\times\mathbf{p}\right] \Bigg]\,\psi. \end{aligned} \tag{4.67}$$

Expanding the square and retaining only the terms with the vector potential,

$$\frac{e}{2mc}\mathbf{p}\mathbf{A} + \frac{e}{2mc}\mathbf{A}\mathbf{p} + \frac{e^2}{2mc^2}A^2 + \frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A}). \tag{4.68}$$

Since we are working in the Coulomb gauge ($\boldsymbol{\nabla}\mathbf{A} = 0$), we have $\mathbf{p}\mathbf{A} = \mathbf{A}\mathbf{p}$ for the radiation field. Thus

$$\mathcal{H}_{\rm int} = \frac{e}{mc}\mathbf{p}\mathbf{A} + \frac{e^2}{2mc^2}A^2 + \frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A}). \tag{4.69}$$

The total Hamiltonian $\mathcal{H} = \mathcal{H}_{\rm rad} + \mathcal{H}_{\rm atom} + \mathcal{H}_{\rm int}$ may then be written as $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_{\rm int}$, where $\mathcal{H}_{\rm int}$ is given by the previous displayed equation and $\mathcal{H}_0 = \mathcal{H}_{\rm rad} + \mathcal{H}_{\rm atom}$. We will treat $\mathcal{H}_{\rm int}$ as a perturbation to $\mathcal{H}_0$ and concentrate first only on the term

$$\mathcal{H}_1 = \frac{e}{mc}\mathbf{p}\mathbf{A}, \tag{4.70}$$

which turns out to be the dominant one in electromagnetic interactions. Using the vector potential in the Schrödinger representation we can write

$$\mathcal{H}_1 = \sum_{\mathbf{k},\lambda} \frac{e}{m}\sqrt{\frac{2\pi\hbar}{V\omega_{\mathbf{k}}}}(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p})\left[a_{\mathbf{k}\lambda}e^{i\mathbf{k}\mathbf{r}} + a^\dagger_{\mathbf{k}\lambda}e^{-i\mathbf{k}\mathbf{r}}\right] = \mathcal{H}_1^{(-)} + \mathcal{H}_1^{(+)}, \tag{4.71}$$

where

$$\mathcal{H}_1^{(-)} = \sum_{\mathbf{k},\lambda} \frac{e}{m}\sqrt{\frac{2\pi\hbar}{V\omega_{\mathbf{k}}}}(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p})\, a_{\mathbf{k}\lambda}e^{i\mathbf{k}\mathbf{r}}, \qquad \mathcal{H}_1^{(+)} = \sum_{\mathbf{k},\lambda} \frac{e}{m}\sqrt{\frac{2\pi\hbar}{V\omega_{\mathbf{k}}}}(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p})\, a^\dagger_{\mathbf{k}\lambda}e^{-i\mathbf{k}\mathbf{r}}. \tag{4.72}$$

In the absence of any interaction (zero order) the wave function of the total system (atom plus radiation field) is a product of an atomic wave function, ψ_a, and a wave function describing the radiation field,

$$\psi_a\,|n_1, n_2, \ldots, n_i, \ldots\rangle \equiv |\psi_a; n_1, n_2, \ldots, n_i, \ldots\rangle\,. \tag{4.73}$$

Here $|n_1, n_2, \ldots, n_i, \ldots\rangle$ is the radiation state in the occupation number representation (or Fock space), as we have already written in the previous section. Consider now the following matrix elements

$$\left\langle \psi_b; n_1', n_2', \ldots, n_i', \ldots \right| \mathcal{H}_1^{(-)} \left| \psi_a; n_1, n_2, \ldots, n_i, \ldots \right\rangle, \tag{4.74}$$

$$\left\langle \psi_b; n_1', n_2', \ldots, n_i', \ldots \right| \mathcal{H}_1^{(+)} \left| \psi_a; n_1, n_2, \ldots, n_i, \ldots \right\rangle. \tag{4.75}$$

According to the orthogonality relations (4.51) and the effect of the creation and annihilation operators on Fock space states (4.52), for a non-zero result of the first matrix element, we need the number occupations for at least one of the modes to verify the relation $n_i' = n_i - 1$. Similarly, for the second matrix element to be non zero we need that $n_i' = n_i + 1$ for at least one of the modes. Let's focus now on a single mode characterized by $\mathbf{k}\lambda$, that is, for photons of momentum $\hbar\mathbf{k}$ and polarization λ, and simplify the notation so that (remember that $a_i |n_i\rangle = \sqrt{n_i}\,|n_i - 1\rangle$).

$$\begin{aligned}\left\langle b; n_{\mathbf{k}\lambda} - 1 \right| \mathcal{H}_1^{(-)} \left| a; n_{\mathbf{k}\lambda} \right\rangle &= \frac{e}{m}\sqrt{\frac{2\pi\hbar}{V\omega_{\mathbf{k}}}} \left\langle b; n_{\mathbf{k}\lambda} - 1 \right| (\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{p}) a_{\mathbf{k}\lambda} e^{i\mathbf{kr}} \left| a; n_{\mathbf{k}\lambda} \right\rangle \\ &= \frac{e}{m}\sqrt{\frac{2\pi\hbar n_{\mathbf{k}\lambda}}{V\omega_{\mathbf{k}}}} \left\langle b \right| (\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{p}) e^{i\mathbf{kr}} \left| a \right\rangle, \end{aligned} \tag{4.76}$$

where we have shortened $|\psi_a\rangle \equiv |a\rangle$ and $|\psi_b\rangle \equiv |b\rangle$ to represent the initial and final atomic state. Similarly,

$$\begin{aligned}\left\langle b; n_{\mathbf{k}\lambda} + 1 \right| \mathcal{H}_1^{(+)} \left| a; n_{\mathbf{k}\lambda} \right\rangle &= \frac{e}{m}\sqrt{\frac{2\pi\hbar}{V\omega_{\mathbf{k}}}} \left\langle b; n_{\mathbf{k}\lambda} + 1 \right| (\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{p}) a_{\mathbf{k}\lambda}^\dagger e^{-i\mathbf{kr}} \left| a; n_{\mathbf{k}\lambda} \right\rangle \\ &= \frac{e}{m}\sqrt{\frac{2\pi\hbar(n_{\mathbf{k}\lambda} + 1)}{V\omega_{\mathbf{k}}}} \left\langle b \right| (\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{p}) e^{-i\mathbf{kr}} \left| a \right\rangle, \end{aligned} \tag{4.77}$$

We will now examine the matrix element

$$\left\langle b \right| (\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{p}) e^{i\mathbf{kr}} \left| a \right\rangle \equiv \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \left\langle b \right| \mathbf{p} e^{i\mathbf{kr}} \left| a \right\rangle, \tag{4.78}$$

in the simplified case that $\mathbf{kr} \ll 1$, so that $e^{i\mathbf{kr}} \approx 1$.[5] The atomic Hamiltonian and the commutation relations[6]

$$\mathcal{H}_{\text{atom}} = \sum_i \frac{p_i^2}{2m} + U, \qquad [r_k, p^2] = 2i\hbar p_k \quad (r_k = x, y, z), \tag{4.79}$$

can be used to write ($\mathbf{r}$ is the position vector of an atomic electron referred to the nucleus),

$$[\mathbf{r}, \mathcal{H}_{\text{atom}}] = \sum_i \frac{[\mathbf{r}, p_i^2]}{2m} + [\mathbf{r}, U] = \frac{[\mathbf{r}, p^2]}{2m} = \frac{i\hbar}{m}\mathbf{p}. \tag{4.80}$$

Therefore, since $\mathcal{H}_{\text{atom}}|a\rangle = E_a|a\rangle$ and $\mathcal{H}_{\text{atom}}|b\rangle = E_b|b\rangle$, where E_a and E_b are the eigenvalues of the atomic Hamiltonian corresponding to the eigenstate $|a\rangle$ and $|b\rangle$, we have

$$\begin{aligned}\langle b|\mathbf{p}|a\rangle &= \frac{m}{i\hbar}\langle b|\,[\mathbf{r}, \mathcal{H}_{\text{atom}}]\,|a\rangle = \frac{m}{i\hbar}\Big(\langle b|\,\mathbf{r}\mathcal{H}_{\text{atom}}\,|a\rangle - \langle b|\,\mathcal{H}_{\text{atom}}\mathbf{r}\,|a\rangle\Big) \\ &= \frac{m}{i\hbar}\Big(E_a\langle b|\,\mathbf{r}\,|a\rangle - E_b\langle b|\,\mathbf{r}\,|a\rangle\Big) = \frac{im}{\hbar}(E_b - E_a)\langle b|\mathbf{r}|a\rangle = im\omega_{ba}\langle b|\mathbf{r}|a\rangle = im\omega_{\mathbf{k}}\langle b|\mathbf{r}|a\rangle. \end{aligned}$$

[5] The wavelength of photons emitted/absorbed by atoms is of the order of 10^3 Å, so $k \sim 10^{-3}$Å^{-1}. The size of an atom is about 1 Å, so we can take $r \sim 1$Å and $kr \sim 10^{-3}$.

[6] Remember that, for example, $[x, p_x^2] = [x, p_x]p_x + p_x[x, p_x] = 2i\hbar p_x$.

In the last step we have used the relation $\omega_{ab} = \omega_{\mathbf{k}}$, as required by energy conservation.

It can be seen that the approximation $e^{i\mathbf{kr}} \approx 1$ in the vector potential neglects all the terms in the multipole expansion of the electromagnetic field except the electric dipole (E1) term.[7] Therefore, within the electric dipole approximation we have

$$\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \left\langle b \left| \mathbf{p} e^{i\mathbf{kr}} \right| a \right\rangle \approx \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \langle b \,|\, \mathbf{p} \,|\, a \rangle = im\omega_{\mathbf{k}}\, \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \langle b \,|\, \mathbf{r} \,|\, a \rangle \,, \tag{4.81}$$

and

$$\left\langle b; n_{\mathbf{k}\lambda} - 1 \left| \mathcal{H}_1^{(-)} \right| a; n_{\mathbf{k}\lambda} \right\rangle = \frac{e}{m} \sqrt{\frac{2\pi\hbar n_{\mathbf{k}\lambda}}{V\omega_{\mathbf{k}}}} \left\langle b \left| (\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{p}) e^{i\mathbf{kr}} \right| a \right\rangle \approx ie \sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}} n_{\mathbf{k}\lambda}}{V}}\, \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \langle b \,|\, \mathbf{r} \,|\, a \rangle \,. \tag{4.82}$$

Consequently,

$$\left\langle b; n_{\mathbf{k}\lambda} - 1 \left| \mathcal{H}_1^{(-)} \right| a; n_{\mathbf{k}\lambda} \right\rangle_{\mathrm{E1}} = ie \sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}} n_{\mathbf{k}\lambda}}{V}}\, \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \langle b \,|\, \mathbf{r} \,|\, a \rangle \,. \tag{4.83}$$

A similar reasoning would conclude that

$$\left\langle b; n_{\mathbf{k}\lambda} + 1 \left| \mathcal{H}_1^{(+)} \right| a; n_{\mathbf{k}\lambda} \right\rangle_{\mathrm{E1}} = ie \sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}} (n_{\mathbf{k}\lambda} + 1)}{V}}\, \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \langle b \,|\, \mathbf{r} \,|\, a \rangle \,. \tag{4.84}$$

The next higher approximation, which is $e^{i\mathbf{kr}} \approx 1 + i\mathbf{kr}$, will lead us to magnetic dipole (M1) and electric quadrupole (E2) matrix elements. The matrix element containing the operator $\mathbf{k} \cdot \mathbf{r}$ (which is equal[8] to $\mathbf{r} \cdot \mathbf{k}$) is proportional to $\langle b \,|\, i(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{p})(\mathbf{k} \cdot \mathbf{r}) \,|\, a \rangle$. To analyze it, note first that we can write

$$\begin{aligned} i(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{p})(\mathbf{k} \cdot \mathbf{r}) = \frac{1}{2} i(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{p})(\mathbf{k} \cdot \mathbf{r}) - \frac{1}{2} i(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{r})(\mathbf{k} \cdot \mathbf{p}) \\ + \frac{1}{2} i(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{p})(\mathbf{k} \cdot \mathbf{r}) + \frac{1}{2} i(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{r})(\mathbf{k} \cdot \mathbf{p}) \equiv \mathcal{O}_A + \mathcal{O}_S, \end{aligned} \tag{4.85}$$

where

$$\mathcal{O}_A = \frac{1}{2} i(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{p})(\mathbf{k} \cdot \mathbf{r}) - \frac{1}{2} i(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{r})(\mathbf{k} \cdot \mathbf{p}), \tag{4.86}$$

and

$$\mathcal{O}_S = \frac{1}{2} i(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{p})(\mathbf{k} \cdot \mathbf{r}) + \frac{1}{2} i(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{r})(\mathbf{k} \cdot \mathbf{p}) \tag{4.87}$$

are the antisymmetric and symmetric operator combinations, respectively. Now we use the vector identity $(\mathbf{AC})(\mathbf{BD}) - (\mathbf{AD})(\mathbf{BC}) = (\mathbf{A} \times \mathbf{B})(\mathbf{C} \times \mathbf{D})$ to transform the $\mathcal{O}_A$ operator into

$$\begin{aligned} \mathcal{O}_A = \frac{1}{2} i(\hat{\mathbf{e}}_{\mathbf{k}\lambda} \times \mathbf{k})(\mathbf{p} \times \mathbf{r}) = \frac{1}{2} i(\mathbf{k} \times \hat{\mathbf{e}}_{\mathbf{k}\lambda})(\mathbf{r} \times \mathbf{p}) = \frac{i}{2} (\mathbf{k} \times \hat{\mathbf{e}}_{\mathbf{k}\lambda}) \cdot \mathbf{L} = \frac{i\omega_{\mathbf{k}}}{2c} (\hat{\mathbf{k}} \times \hat{\mathbf{e}}_{\mathbf{k}\lambda}) \cdot \mathbf{L} \\ = i\mu_{\mathrm{B}} \frac{m\omega_{\mathbf{k}}}{e} (\hat{\mathbf{k}} \times \hat{\mathbf{e}}_{\mathbf{k}\lambda}) \cdot \frac{\mathbf{L}}{\hbar} = -i \frac{m\omega_{\mathbf{k}}}{e} (\hat{\mathbf{k}} \times \hat{\mathbf{e}}_{\mathbf{k}\lambda}) \cdot \boldsymbol{\mu}_{\mathrm{L}}, \end{aligned} \tag{4.88}$$

[7] For a brief review of the multipole expansion of the classical electromagnetic field see appendix (B.7.2).

[8] Note that $\mathbf{k}$ is just a constant vector here, with a magnitude proportional to the energy of the photon and a direction coincident with that of the propagation of the plane wave, while $\mathbf{r}$ is the position vector operator of an electron. Note also that the operator $\mathbf{p}$ acts on everything that stands to its right according to the derivative chain rule, so

$$\mathbf{p}(\mathbf{kr})\psi = (\mathbf{kr})(\mathbf{p}\psi) + [\mathbf{p}(\mathbf{kr})]\psi = (\mathbf{kr})(\mathbf{p}\psi) + [-i\hbar\boldsymbol{\nabla}(k_x x + k_y y + k_z z)]\psi = (\mathbf{kr})(\mathbf{p}\psi) - i\hbar\mathbf{k}\psi,$$

because $\boldsymbol{\nabla}(k_x x + k_y y + k_z z) = \mathbf{k}$. Furthermore, since $\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{k} = 0$, we have

$$\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{p}(\mathbf{kr}) = \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \Big[(\mathbf{kr})\mathbf{p} - i\hbar\mathbf{k} \Big] = \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot (\mathbf{kr})\mathbf{p}.$$

where $\mu_{\rm B} = e\hbar/(2mc)$ is the Bohr magneton and $\boldsymbol{\mu}_{\rm L} = -\mu_{\rm B}\mathbf{L}/\hbar$ is the magnetic moment operator associated with the orbital angular momentum. The vector $\hat{\mathbf{k}} \times \hat{\mathbf{e}}_{\mathbf{k}\lambda}$ lies in the direction of the magnetic field, as it is clear from equation (4.58). The connection with the magnetic field can still be made more explicit by writing the $\mathcal{H}_1^{(-)}$ Hamiltonian as (only for one mode $\mathbf{k}\lambda$),

$$\begin{aligned}
\mathcal{H}_1^{(-)} &= \frac{e}{m}\sqrt{\frac{2\pi\hbar}{V\omega_{\mathbf{k}}}}(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p})\,a_{\mathbf{k}\lambda}e^{i\mathbf{kr}} \;\approx\; \frac{e}{m}\sqrt{\frac{2\pi\hbar}{V\omega_{\mathbf{k}}}}(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p})\,a_{\mathbf{k}\lambda}(1+i\mathbf{kr}) \\
&= \left[\mathcal{H}_1^{(-)}\right]_{\rm E1} + i\frac{e}{m}\sqrt{\frac{2\pi\hbar}{V\omega_{\mathbf{k}}}}\,a_{\mathbf{k}\lambda}\,(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p})(\mathbf{kr}) = \left[\mathcal{H}_1^{(-)}\right]_{\rm E1} + \frac{e}{m}\sqrt{\frac{2\pi\hbar}{V\omega_{\mathbf{k}}}}\,a_{\mathbf{k}\lambda}\,(\mathcal{O}_A+\mathcal{O}_S) \\
&= \left[\mathcal{H}_1^{(-)}\right]_{\rm E1} + \frac{e}{m}\sqrt{\frac{2\pi\hbar}{V\omega_{\mathbf{k}}}}\,a_{\mathbf{k}\lambda}\,\left(-i\frac{m\omega_{\mathbf{k}}}{e}(\hat{\mathbf{k}}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda})\cdot\boldsymbol{\mu}_{\rm L}+\mathcal{O}_S\right) \\
&= \left[\mathcal{H}_1^{(-)}\right]_{\rm E1} - i\sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}}{V}}\,a_{\mathbf{k}\lambda}\,(\hat{\mathbf{k}}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda})\cdot\boldsymbol{\mu}_{\rm L} + \frac{e}{m}\sqrt{\frac{2\pi\hbar}{V\omega_{\mathbf{k}}}}\,a_{\mathbf{k}\lambda}\,\mathcal{O}_S.
\end{aligned}$$

The interaction associated with $\mathcal{O}_A$ is then

$$\left[\mathcal{H}_1^{(-)}\right]_{\rm M1} = -i\sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}}{V}}a_{\mathbf{k}\lambda}(\hat{\mathbf{k}}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda})\cdot\boldsymbol{\mu}_L = -\mathbf{B}_{\mathbf{k}\lambda}^{(-)}\cdot\boldsymbol{\mu}_{\rm L} \tag{4.89}$$

where $\mathbf{B}_{\mathbf{k}\lambda}^{(-)}$ is the term containing the $a_{\mathbf{k}\lambda}$ operator in (4.58). Although it is not long, we will not reproduce here the detailed derivation of the multipole character of this interaction,[9] but merely state that the $\mathcal{O}_A$ operator, or $-\mathbf{B}_{\mathbf{k}\lambda}^{(-)}\cdot\boldsymbol{\mu}_{\rm L}$, or any of the equivalent forms of these operators, are magnetic dipole (M1) operators.

In addition to the magnetic moment operator associated with the orbital angular momentum we can also include that of the spin. This is done by taking into account the term $\frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A})$ of equation (4.69), which just amounts to add $2\mathbf{S}$ to $\mathbf{L}$ in (4.88).[10] Denoting by $\boldsymbol{\mu}$ the total magnetic moment operator associated with both spin and orbital angular momenta, the matrix elements for magnetic dipole (M1) transitions can then be written as

$$\begin{aligned}
&\left\langle b;n_{\mathbf{k}\lambda}-1\left|\,\mathcal{H}_1^{(-)}\,\right|a;n_{\mathbf{k}\lambda}\right\rangle_{\rm M1} \;=\; \left\langle b;n_{\mathbf{k}\lambda}-1\left|\,-\boldsymbol{\mu}\cdot\mathbf{B}_{\mathbf{k}\lambda}^{(-)}\,\right|a;n_{\mathbf{k}\lambda}\right\rangle_{\rm M1} \\
&= -i\sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}n_{\mathbf{k}\lambda}}{V}}\,(\hat{\mathbf{k}}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda})\cdot\langle b|\boldsymbol{\mu}|a\rangle \;=\; i\sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}n_{\mathbf{k}\lambda}}{V}}\,\frac{\mu_{\rm B}}{\hbar}(\hat{\mathbf{k}}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda})\cdot\langle b|\mathbf{L}+\mathbf{2S}|a\rangle\,,
\end{aligned} \tag{4.90}$$

and

$$\begin{aligned}
&\left\langle b;n_{\mathbf{k}\lambda}+1\left|\,\mathcal{H}_1^{(+)}\,\right|a;n_{\mathbf{k}\lambda}\right\rangle_{\rm M1} \;=\; \left\langle b;n_{\mathbf{k}\lambda}+1\left|\,-\boldsymbol{\mu}\cdot\mathbf{B}_{\mathbf{k}\lambda}^{(+)}\,\right|a;n_{\mathbf{k}\lambda}\right\rangle_{\rm M1} \\
&= i\sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}(n_{\mathbf{k}\lambda}+1)}{V}}\,(\hat{\mathbf{k}}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda})\cdot\langle b|\boldsymbol{\mu}|a\rangle \;=\; -i\sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}(n_{\mathbf{k}\lambda}+1)}{V}}\,\frac{\mu_{\rm B}}{\hbar}(\hat{\mathbf{k}}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda})\cdot\langle b|\mathbf{L}+\mathbf{2S}|a\rangle\,.
\end{aligned} \tag{4.91}$$

Let's go back now to the symmetric operator. Since $\mathbf{k}\cdot\mathbf{r} = \mathbf{r}\cdot\mathbf{k}$ and $\mathbf{k}\cdot\mathbf{p} = \mathbf{p}\cdot\mathbf{k}$, the $\mathcal{O}_S$ operator can also be written as

$$\mathcal{O}_S = \frac{1}{2}i(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p})(\mathbf{r}\cdot\mathbf{k}) + \frac{1}{2}i(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{r})(\mathbf{p}\cdot\mathbf{k}) = \frac{1}{2}i\,\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot(\mathbf{pr}+\mathbf{rp})\cdot\mathbf{k}. \tag{4.92}$$

[9] The interested reader can see this in the book of *Weissbluth*.

[10] Note that to derive the orbital angular momentum contribution to the magnetic dipole transitions, given by equation (4.89), we needed to go up to the second term in the approximation $e^{i\mathbf{kr}} \approx 1+i\mathbf{kr}$. The spin contribution, however, is already present explicitly in equation (4.69). It is a big success of Dirac's equation, which is at the origin of the $\frac{e\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A})$ term, to be able to account in a natural way for the spin contribution to the magnetic dipole transitions.

As before, we make the replacement $\mathbf{p} = \frac{im}{\hbar}[\mathcal{H}_{\rm atom}, \mathbf{r}]$ to get

$$\mathcal{O}_S = -\frac{m}{2\hbar}\,\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\Big([\mathcal{H}_{\rm atom},\mathbf{r}]\mathbf{r}+\mathbf{r}[\mathcal{H}_{\rm atom},\mathbf{r}]\Big)\cdot\mathbf{k} = \frac{m}{2\hbar}\,\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot[\mathbf{rr},\mathcal{H}_{\rm atom}]\cdot\mathbf{k}, \tag{4.93}$$

where $\mathbf{rr}$ is a symmetric Cartesian tensor operator of rank 2, with components

$$(\mathbf{rr})_{ik} = \begin{pmatrix} xx & xy & xz \\ yx & yy & yz \\ zx & zy & zz \end{pmatrix}. \tag{4.94}$$

We have then

$$\langle b\,|\,\mathcal{O}_S\,|\,a\rangle = -\frac{m}{2\hbar}\,(E_b-E_a)\,\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\langle b\,|\,\mathbf{rr}\,|\,a\rangle\cdot\mathbf{k} = -\frac{1}{2}m\omega_{ba}\,\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\langle b\,|\,\mathbf{rr}\,|\,a\rangle\cdot\mathbf{k}. \tag{4.95}$$

The Cartesian operator $\mathbf{rr}$ is not an irreducible tensor with respect to orthogonal transformations, but the use of the Wigner-Eckart theorem requires an operator that transforms as an irreducible tensor. However, due to the transversality condition we can write

$$\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\left\langle b\,\middle|\,r^2\delta_{ij}\,\middle|\,a\right\rangle\cdot\mathbf{k} = \left\langle b\,\middle|\,r^2\delta_{ij}\,\middle|\,a\right\rangle\,\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{k} = 0. \tag{4.96}$$

Therefore we can replace $\mathbf{rr}$ by the operator

$$\mathcal{Q}_{ij} = (\mathbf{rr})_{ij} - \frac{1}{3}r^2\delta_{ij}, \tag{4.97}$$

which can be show to be an irreducible tensor of rank 2. In terms of $\mathcal{Q}$ we have now

$$\langle b\,|\,\mathcal{O}_S\,|\,a\rangle = -\frac{m\omega_{ba}}{2\hbar}\,\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\langle b\,|\,\mathcal{Q}\,|\,a\rangle\cdot\mathbf{k} = -\frac{m\omega_{\mathbf{k}}}{2\hbar}\,\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\langle b\,|\,\mathcal{Q}\,|\,a\rangle\cdot\mathbf{k} = -\frac{m\omega_{\mathbf{k}}^2}{2\hbar}\,\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\langle b\,|\,\mathcal{Q}\,|\,a\rangle\cdot\hat{\mathbf{k}},$$

where we have set $\omega_{ba} = \omega_{\mathbf{k}}$ due to energy conservation. Further study of the the symmetric operator $\mathcal{O}_S$ would reveal that it leads to electric quadrupole (E2) matrix elements, but we will not pursue this here.

Up to now, we have been considering the interaction of the radiation field with a single atomic electron. For a system containing N electrons we define the operators

$$\mathbf{R} = \sum_{j=1}^{N}\mathbf{r}_j, \qquad\qquad (\mathcal{Q}_j)_{ik} = (\mathbf{r}_j\mathbf{r}_j)_{ik} - \frac{1}{3}r^2\delta_{ik}, \tag{4.98}$$

where $\mathbf{r}_j$ is the position vector of the j-th electron. The reference frame will be centered in the nucleus or, if molecules are discussed, the center of gravity of the positive charge. Considering only a single radiation mode characterized by $\mathbf{k}\lambda$, the three matrix elements corresponding to the lowest multipoles are

- Electric Dipole, E1:

$$\left\langle b; n_{\mathbf{k}\lambda}-1\,\middle|\,\mathcal{H}_1^{(-)}\,\middle|\,a; n_{\mathbf{k}\lambda}\right\rangle_{\rm E1} = ie\sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}n_{\mathbf{k}\lambda}}{V}}\;\;\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\langle b\,|\,\mathbf{R}\,|\,a\rangle \tag{4.99}$$

$$\left\langle b; n_{\mathbf{k}\lambda}+1\,\middle|\,\mathcal{H}_1^{(+)}\,\middle|\,a; n_{\mathbf{k}\lambda}\right\rangle_{\rm E1} = ie\sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}(n_{\mathbf{k}\lambda}+1)}{V}}\;\;\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\langle b\,|\,\mathbf{R}\,|\,a\rangle \tag{4.100}$$

- Magnetic Dipole, M1:

$$\left\langle b; n_{\mathbf{k}\lambda}-1\,\middle|\,\mathcal{H}_1^{(-)}\,\middle|\,a; n_{\mathbf{k}\lambda}\right\rangle_{\rm M1} = -i\sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}n_{\mathbf{k}\lambda}}{V}}\;\;(\hat{\mathbf{k}}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda})\cdot\left\langle b\,\middle|\,\sum_j\boldsymbol{\mu}_j\,\middle|\,a\right\rangle \tag{4.101}$$

$$\left\langle b; n_{\mathbf{k}\lambda}+1\,\middle|\,\mathcal{H}_1^{(+)}\,\middle|\,a; n_{\mathbf{k}\lambda}\right\rangle_{\rm M1} = i\sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}(n_{\mathbf{k}\lambda}+1)}{V}}\;\;(\hat{\mathbf{k}}\times\hat{\mathbf{e}}_{\mathbf{k}\lambda})\cdot\left\langle b\,\middle|\,\sum_j\boldsymbol{\mu}_j\,\middle|\,a\right\rangle \tag{4.102}$$

- Electric Quadrupole, E2:

$$\left\langle b; n_{\mathbf{k}\lambda} - 1 \left| \mathcal{H}_1^{(-)} \right| a; n_{\mathbf{k}\lambda} \right\rangle_{\text{E2}} = -\frac{e}{2c}\sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}^3 n_{\mathbf{k}\lambda}}{V}}\ \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \left\langle b \left| \sum_j \mathcal{Q}_j \right| a \right\rangle \cdot \hat{\mathbf{k}} \tag{4.103}$$

$$\left\langle b; n_{\mathbf{k}\lambda} + 1 \left| \mathcal{H}_1^{(+)} \right| a; n_{\mathbf{k}\lambda} \right\rangle_{\text{E2}} = \frac{e}{2c}\sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}^3 (n_{\mathbf{k}\lambda}+1)}{V}}\ \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \left\langle b \left| \sum_j \mathcal{Q}_j \right| a \right\rangle \cdot \hat{\mathbf{k}} \tag{4.104}$$

In the formulas above, $\boldsymbol{\mu}_j$ is the total magnetic moment operator associated to the j-th electron, which is given by $\mu_{\text{B}}(\boldsymbol{\ell}_j + 2\mathbf{s}_j)/\hbar$ (see (2.30)), and the $\mathcal{Q}_j$ are the quadrupole operators for each electron. Note that, since $\langle b \,|\, \mathcal{Q} \,|\, a \rangle$ is a 3×3 matrix, the vector operator $\langle b \,|\, \mathcal{Q} \,|\, a \rangle \cdot \hat{\mathbf{k}}$ is not necessarily proportional to $\hat{\mathbf{k}}$, otherwise the matrix elements of (4.103) and (4.104) would always be zero because $\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{k}} = 0$.

The matrix elements corresponding to E1 transitions change sign under parity transformations because they are vectors. The $\boldsymbol{\mu}_j$ operators are axial vectors, therefore the M1 matrix elements do not change under parity. From equation (4.98) it is clear that the quadrupole operators are also parity-invariant, so the E2 matrix operators are parity-invariant too.

4.4 Selection rules.

We have seen that, at least in the first two orders of a multipole expansion, the atomic transitions mediated by the interaction with the electromagnetic radiation field are proportional to the matrix elements of irreducible tensor operators of rank 1 (vectors), for the E1 and M1 contributions, and of rank 2 for the E2 contributions. We can then use the Wigner-Eckart theorem, proved in the appendix (B.5.5), which establishes that, given an irreducible tensor operator[11] $T_q^{(k)}$ and two states of angular momenta J_a and J_b, there exists a constant independent of M_a, M_b, and q (the so-called reduced matrix element), such that the following equation is satisfied:

$$\begin{aligned}\left\langle \alpha_b J_b M_b \left| T_q^{(k)} \right| \alpha_a J_a M_a \right\rangle &= \frac{1}{2J_b+1} \langle J_a k M_a q \,|\, J_a k J_b M_b \rangle \left\langle \alpha_b J_b \left\| \mathbf{T}^{(k)} \right\| \alpha_a J_a \right\rangle \\ &= (-1)^{J_b - M_b} \begin{pmatrix} J_b & k & J_a \\ -M_b & q & M_a \end{pmatrix} \left\langle \alpha_b J_b \left\| \mathbf{T}^{(k)} \right\| \alpha_a J_a \right\rangle . \end{aligned} \tag{4.105}$$

The symbols α_a and α_b stand for all other quantum numbers required to identify the states. They have a degeneracy of $g_a = 2J_a + 1$ and $g_b = 2J_b + 1$. The selection rules follow directly from the properties of the Clebsch-Gordan coefficients in equation (4.105). These coefficients are different from zero if J_a, J_b, k, M_a, M_b and q satisfy the angular momentum sum rules,

$$\langle J_a k M_a q \,|\, J_a k J_b M_b \rangle \neq 0 \iff |J_a - k| \leq J_b \leq J_a + k, \text{ and } M_b = M_a + q.$$

The relation $|J_a - k| \leq J_b \leq J_a + k$ implies $J_a + J_b \geq k$.[12] For E1 and M1 transitions the operators are irreducible tensors of rank 1 (vectors, $k = 1$), therefore the corresponding selection rules are

$$\text{E1 and M1:} \qquad \Delta J = 0, \pm 1, \qquad \Delta M = 0, \pm 1, \qquad J_a + J_b \geq 1. \tag{4.106}$$

[11] It is common terminology to say that $T_q^{(k)}$ is an operator of spin k. This means that the $T_q^{(k)}$ are a set of $2k+1$ operators which have the following commutation rules with the generators of rotations:

$$[J_3, T_q^{(k)}] = \hbar q\, T_q^{(k)}, \qquad [J_1 \pm i J_2, T_q^{(k)}] = \hbar\sqrt{k(k+1) - q(q \pm 1)}\, T_{q\pm 1}^{(k)}.$$

For more details see appendix (B.5).

[12] The minimum value of J_b is $|J_a - k|$, therefore $J_a + J_b \geq J_a + |J_a - k|$. When $k \geq J_a$, this relation leads directly to $J_a + J_b \geq k$. When $J_a > k$ we have $J_a + J_b \geq 2J_a - k > k$, therefore $J_a + J_b \geq k$ always.

The condition $J_a + J_b \geq 1$ implies that transitions $0 \leftrightarrow 0$ cannot take place. This is, of course, a direct consequence of angular momentum conservation, since a photon (spin 1) and an electronic system with $J = 0$ cannot be coupled to produce another system with $J = 0$. For E2 transitions the operator is an irreducible tensor of rank 2 ($k = 2$), so

$$\text{E2:} \qquad \Delta J = 0, \pm 1, \pm 2, \qquad \Delta M = 0, \pm 1, \pm 2, \qquad J_a + J_b \geq 2. \tag{4.107}$$

The condition $J_a + J_b \geq 2$ implies that E2 transitions $0 \leftrightarrow 0$, $1/2 \leftrightarrow 1/2$ or $0 \leftrightarrow 1$ cannot take place.

The previous selection rules follow directly from the Wigner-Eckart theorem. However, there are more constraints arising from the fact that electromagnetic interactions conserve parity.[13] From classical electromagnetic theory, we know that the field associated with an E1 multipole has odd parity ($\pi_{\text{em}} = -1$), while the fields associated with M1 and E2 multipoles have even parity ($\pi_{\text{em}} = +1$). The same is true for the quantum-mechanical operators that describe these transitions, as it is clear from equations (4.99) to (4.104). Therefore we have $\pi_{\text{E1}} = -1$ and $\pi_{\text{M1}} = \pi_{\text{E2}} = +1$. The conservation of parity requires $\pi_a \pi_b \pi_{\text{op}} = +1$, where π_a and π_b are the parities of the initial and final states and π_{op} is the parity of the operator for the particular multipole transition. If the product of parities $\pi_a \pi_b \pi_{\text{op}}$ is -1, the integrand of the transition matrix element is odd and the integral vanishes. We have then the following parity selection rules,

$$\begin{aligned} \text{E1:} \qquad & \pi_a = -\pi_b, \qquad (-1)^{\sum_i \Delta \ell_i} = -1, \\ \text{M1 and E2:} \qquad & \pi_a = \pi_b. \end{aligned} \tag{4.108}$$

The set of selection rules just discussed follows from general symmetry considerations and are independent of the angular momenta coupling scheme, but additional constraints appear when we take into consideration different coupling schemes. Let us discuss briefly the selection rules in the LS coupling scheme. Remember that in this case one forms linear combinations of central-field product-type wave functions belonging to a given configuration, with specific values of the quantum numbers (n_i, ℓ_i) for each of the N electrons. For these wave functions, L, S and J are good quantum numbers, as well as n_i and ℓ_i.

For LS coupling the orbital and spin angular momenta are independent, and since the E1 and E2 operators are independent of spin, the condition $\Delta S = 0$ must hold. We have then

$$\text{E1:} \qquad \Delta S = 0, \qquad \Delta L = 0, \pm 1, \qquad L_a + L_b \geq 1, \qquad \Delta M_L = 0, \pm 1, \tag{4.109}$$

In LS coupling the most usual case is one in which the orbital of only one electron in the atom changes (one electron jump).[14] If the electron concerned has quantum numbers (ℓ, m_ℓ), then for this single electron the following relations must be satisfied

$$\Delta \ell = \pm 1 \quad \text{and} \quad \Delta m_\ell = 0, \pm 1.$$

The case $\Delta \ell = 0$ is forbidden because the initial and final states must have opposite parities. Note that $\Delta L = \Delta \ell$ for atoms with only one valence electron of orbital angular momentum ℓ, therefore $\Delta L = 0$ transitions are forbidden in such cases. However, in atoms with several valence electrons it is possible to get E1 transitions between different configurations with the same value of L. An example is the $\Delta L = 0$ transition seen at 250.7 nm in silicon between the $^3\text{P}_2$ level of an excited configuration and the $^3\text{P}_1$ level of the ground-state configuration,

$$[\text{Ne}](3s)^2(3p)(4s)\,{}^3\text{P}_2 \quad \longrightarrow \quad [\text{Ne}](3s)^2(3p)^2\,{}^3\text{P}_1. \tag{4.110}$$

[13] This implies that atomic states have well-defined parities: odd or even, but not a mixture.

[14] If two or more configurations are strongly mixed in a particular state, then more than one electron can make a transition simultaneously even though only one photon is emitted or absorbed, but this is a rather uncommon event.

The selection rules for electric quadrupole transitions in LS coupling are

$$\text{E2:} \qquad \Delta S = 0, \quad \Delta L = 0, \pm 1, \pm 2 \quad L_a + L_b \geq 2, \quad \Delta M_L = 0, \pm 1, \pm 2. \tag{4.111}$$

For a transition of a single electron we have $\Delta \ell = 0, \pm 2$ only (due to parity conservation), $\ell_a + \ell_b \geq 2$ and $\Delta m_\ell = 0, \pm 1, \pm 2$.

In exact LS coupling, M1 transitions cannot take place. To see why, we should examine the following matrix elements of the operator for M1 transitions,

$$\left\langle L'S'M'_L M'_S \left| \mathbf{L} + \mathbf{2S} \right| LSM_L M_S \right\rangle . \tag{4.112}$$

First, note that $\mathbf{L} + \mathbf{2S}$ does not act on the radial part of the wave functions, and because two radial wave functions with the same ℓ, but different n (principal quantum number), are orthogonal, one has the requirement that $\Delta n_i = 0$ for each electron $i = 1, \ldots N$. Furthermore, since the operator $\mathbf{L} + \mathbf{2S}$ commutes with L^2, S^2 and ℓ_i^2, it follows that the previous matrix element vanishes unless $\Delta L = 0$, $\Delta \ell_i = 0$ and $\Delta S = 0$.[15] But, if this happens, the energies are identical because in LS coupling the states are degenerate with respect to M_L and M_S, and the photon energy in the transition would be zero. If there is some non-negligible spin-orbit interaction, the coupling of $\mathbf{L}$ and $\mathbf{S}$ gives rise to states whose energies depend on J, therefore departing from exact LS coupling. In such a case the matrix element

$$\left\langle LSJ'_L M'_J \left| \mathbf{L} + \mathbf{2S} \right| LSJM_J \right\rangle \tag{4.113}$$

can have off diagonal elements with $\Delta J = \pm 1$. In any case, magnetic dipole transitions between states belonging to the same fine-structure multiplet have a very small probability to occur.

Table (4.1) shows the selection rules for the most relevant discrete transitions. It is common terminology to call a transition *allowed* if it verifies the selection rules of electric-dipole transitions, and *forbidden* otherwise. The use of the word *forbidden* is a bit misleading, because the transition may still take place through higher-order processes, like the M1 and E2 transitions, but at a much slower rate. In this context, the word *forbidden* should really be understood as *electric-dipole forbidden*.

4.5 Absorption and emission of radiation.

4.5.1 Transition probabilities.

We will consider only two discrete atomic states between which the transition occurs. This is appropriate if there are no other states close within the energy of the photon to be absorbed or emitted, $\hbar\omega$. Let $|a\rangle$ be the initial state of the atom and $|n_{\mathbf{k}\lambda}\rangle$ the initial state of the radiation field. The initial state of the total system, atom plus field, is $|A\rangle = |a; n_{\mathbf{k}\lambda}\rangle$. If absorption takes place, the atom makes a transition to another state $|b\rangle$ and the radiation field has one photon less. The final state of the total system would be $|B\rangle = |b; n_{\mathbf{k}\lambda} - 1\rangle$. We have the following equations

$$|A\rangle = |a; n_{\mathbf{k}\lambda}\rangle , \qquad E_A = E_a + \left(n_{\mathbf{k}\lambda} + \frac{1}{2} \right) \hbar\omega_{\mathbf{k}}, \tag{4.114}$$

$$|B\rangle = |b; n_{\mathbf{k}\lambda} - 1\rangle , \qquad E_B = E_b + \left(n_{\mathbf{k}\lambda} - \frac{1}{2} \right) \hbar\omega_{\mathbf{k}}, \tag{4.115}$$

$$E_B - E_A = \underbrace{E_b - E_a}_{\hbar\omega_{ba}} - \hbar\omega_{\mathbf{k}} = \hbar(\omega_{ba} - \omega_{\mathbf{k}}), \tag{4.116}$$

[15] In other words, the components of $\mathbf{L} + \mathbf{2S}$ operating on $|LSM_L M_S\rangle$ do not change L or S, although they can modify the values of M_L and M_S. This implies, due to the orthogonality relations of the states, that the quoted matrix element will vanish unless $L' = L$ and $S' = S$.

		Electric dipole (E1) ("allowed")	**Mag. dipole (M1) ("forbidden")**	**Elec. quadr. (E2) ("forbidden")**
Rigorous rules	1.	$\Delta J = 0, \pm 1$. Except $0 \leftrightarrow 0$.	$\Delta J = 0, \pm 1$. Except $0 \leftrightarrow 0$.	$\Delta J = 0, \pm 1, \pm 2$. Except $0 \leftrightarrow 0$, $\frac{1}{2} \leftrightarrow \frac{1}{2}$, $0 \leftrightarrow 1$.
	2.	$\Delta M = 0 \pm 1$. Except $0 \leftrightarrow 0$ when $\Delta J = 0$.	$\Delta M = 0 \pm 1$. Except $0 \leftrightarrow 0$ when $\Delta J = 0$.	$\Delta M = 0 \pm 1, \pm 2$.
	3.	Parity change.	No parity change.	No parity change.
With negligible configuration interaction	4.	One electron jumping, with $\Delta \ell = \pm 1$, Δn arbitrary.	No change in electron configuration; i.e., for all electrons, $\Delta \ell = 0$ and $\Delta n = 0$.	No change in electron configuration, or one electron jumping with $\Delta \ell = 0, \pm 2$ and Δn arbitrary.
Only for LS coupling	5.	$\Delta S = 0$.	$\Delta S = 0$.	$\Delta S = 0$.
	6.	$\Delta L = 0 \pm 1$. Except $0 \leftrightarrow 0$.	$\Delta L = 0$, $\Delta J = \pm 1$.	$\Delta L = 0 \pm 1 \pm 2$. Except $0 \leftrightarrow 0$, $0 \leftrightarrow 1$.

Table 4.1: Selection rules for discrete transitions (adapted from https://www.nist.gov/). Although in exact LS coupling M1 transitions cannot take place, a little departure from exact coupling gives rise to M1 transitions with $\Delta L = 0$ and $\Delta J = \pm 1$, as shown in this table.

where E_a and E_b are the energies of the initial and final states, respectively. Of course, conservation of energy would require $\omega_{ba} = \omega_{\mathbf{k}}$. Using time-dependent perturbation theory we can express, to first order, the probability per unit time for a transition from an arbitrary state $|A\rangle$ to another state $|B\rangle$ as

$$W^{BA} = \frac{2\pi}{\hbar} |\langle B|\delta\mathcal{H}|A\rangle|^2 \delta(E_B - E_A), \tag{4.117}$$

where $\delta\mathcal{H}$ is the perturbing potential in the Schrödinger representation. This is an example of the application of the so-called *Fermi golden rule* (see equations B.77 and B.82). When studying absorption and emission in atoms it usually happens that $kr \ll 1$, where $2\pi/k$ is the photon wavelength and r is an atomic dimension. Indeed, taking a wavelength of about 1000 Å and $r \sim 1$ Å we see that $kr \sim 2 \times 10^{-3}$. In such a case the vector potential can be considered constant over the region of the atom and it suffices to compute electric-dipole matrix elements. Using equation (4.99) we have the following expression for the probability of absorption,

$$W^{\text{abs}} = \frac{4\pi^2 e^2 \omega_{\mathbf{k}} n_{\mathbf{k}\lambda}}{V} \left| \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \langle b | \mathbf{R} | a \rangle \right|^2 \delta(E_B - E_A), \tag{4.118}$$

where the delta function

$$\delta(E_B - E_A) = \delta(E_b - E_a - \hbar\omega_{\mathbf{k}}) = \frac{1}{\hbar}\delta(\omega_{ba} - \omega_{\mathbf{k}}), \tag{4.119}$$

ensures the conservation of energy for the system as a whole (atom plus radiation field). Of course, a single frequency $\omega_{\mathbf{k}}$ is not a realistic radiation field, we need to take into account at

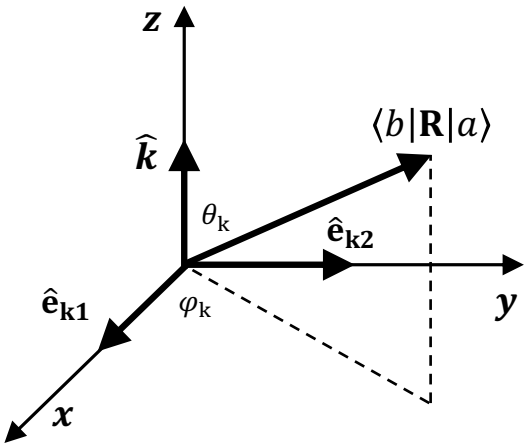


Figure 4.1: Reference frame for equations (4.125) and (4.125).

least a distribution of modes within a narrow interval of frequencies.[16] According to (4.5), the number of propagation modes in a cubic box within an interval $\mathrm{d}k$ and in the direction defined by the element of solid angle $\mathrm{d}\Omega$ is

$$\mathrm{d}N = \frac{V}{(2\pi)^3} k^2 \,\mathrm{d}k \,\mathrm{d}\Omega = \frac{V}{(2\pi)^3} \frac{\omega_{\mathbf{k}}^2}{c^3} \,\mathrm{d}\omega_{\mathbf{k}} \,\mathrm{d}\Omega, \tag{4.120}$$

with $k = \omega_{\mathbf{k}}/c$. Therefore the number of modes of a radiation field in a cubic box of volume V per unit energy interval is

$$\frac{\mathrm{d}N}{\mathrm{d}E} = \frac{1}{\hbar} \frac{\mathrm{d}N}{\mathrm{d}\omega_{\mathbf{k}}} = \frac{V}{(2\pi)^3 \hbar c^3} \,\omega_{\mathbf{k}}^2 \,\mathrm{d}\Omega. \tag{4.121}$$

Substituting the infinitely sharp density distribution $\delta(E_B - E_A) = \frac{1}{\hbar}\delta(\omega_{ba} - \omega_{\mathbf{k}})$ by the previous expression for $\mathrm{d}N/\mathrm{d}E$, we get

$$\mathrm{d}W^{\mathrm{abs}} = \frac{4\pi^2 e^2 \omega_{\mathbf{k}} n_{\mathbf{k}\lambda}}{V} \left| \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \langle b \,|\, \mathbf{R} \,|\, a \rangle \right|^2 \frac{V}{(2\pi)^3 \hbar c^3} \,\omega_{\mathbf{k}}^2 \,\mathrm{d}\Omega = \frac{e^2 \omega_{\mathbf{k}}^3 n_{\mathbf{k}\lambda}}{2\pi\hbar c^3} \left| \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \langle b \,|\, \mathbf{R} \,|\, a \rangle \right|^2 \mathrm{d}\Omega. \tag{4.122}$$

Here, quantities such as $\omega_{\mathbf{k}}$ and $n_{\mathbf{k}\lambda}$ must be understood as averages within the interval. For the emission process the calculation is just the same, except that we have to use the matrix element (4.100). The probabilities per unit time for absorption and emission of a photon with propagation vector $\mathbf{k}$, polarization λ, and frequency $\omega_{\mathbf{k}}$ contained within an element of solid angle $\mathrm{d}\Omega$ are therefore ($\alpha = e^2/\hbar c$)

$$\mathrm{d}W^{\mathrm{abs}} = \frac{\alpha \omega_{\mathbf{k}}^3 n_{\mathbf{k}\lambda}}{2\pi c^2} \left| \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \langle b \,|\, \mathbf{R} \,|\, a \rangle \right|^2 \mathrm{d}\Omega, \tag{4.123}$$

$$\mathrm{d}W^{\mathrm{em}} = \frac{\alpha \omega_{\mathbf{k}}^3 (n_{\mathbf{k}\lambda} + 1)}{2\pi c^2} \left| \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \langle b \,|\, \mathbf{R} \,|\, a \rangle \right|^2 \mathrm{d}\Omega. \tag{4.124}$$

Let's now sum over the two independent polarizations and integrate over the solid angle. We choose a reference frame where $\mathbf{k}$ lies along the z-coordinate axis, $\hat{\mathbf{e}}_{\mathbf{k}1}$ along the x-coordinate axis and $\hat{\mathbf{e}}_{\mathbf{k}2}$ along the y-axis. Calling θ_k and φ_k the angular coordinates of the vector $\langle b \,|\, \mathbf{R} \,|\, a \rangle$ we have then

$$\hat{\mathbf{e}}_{\mathbf{k}1} \cdot \langle b \,|\, \mathbf{R} \,|\, a \rangle = |\, \langle b \,|\, \mathbf{R} \,|\, a \rangle \,|\, \sin\theta_k \cos\varphi_k, \tag{4.125}$$

$$\hat{\mathbf{e}}_{\mathbf{k}2} \cdot \langle b \,|\, \mathbf{R} \,|\, a \rangle = |\, \langle b \,|\, \mathbf{R} \,|\, a \rangle \,|\, \sin\theta_k \sin\varphi_k, \tag{4.126}$$

[16] Furthermore, as we will see later, the atomic states have a natural width which allows a small range of frequencies to be absorbed or emitted.

$$\sum_{\lambda=1}^{2}\left|\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\langle b\,|\,\mathbf{R}\,|\,a\rangle\right|^2=|\,\langle b\,|\,\mathbf{R}\,|\,a\rangle\,|^2\,\sin^2\theta_k,\qquad\int\sin^2\theta_k\,\mathrm{d}\Omega=8\pi/3. \tag{4.127}$$

Assuming equal population numbers $n_{\mathbf{k}1}=n_{\mathbf{k}2}\equiv n_{\mathbf{k}}$ (unpolarized radiation), we have

$$\begin{aligned}\int\sum_{\lambda}\frac{\alpha\omega_{\mathbf{k}}^3 n_{\mathbf{k}\lambda}}{2\pi c^2}\left|\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\langle b\,|\,\mathbf{R}\,|\,a\rangle\right|^2\mathrm{d}\Omega&=\frac{\alpha\omega_{\mathbf{k}}^3 n_{\mathbf{k}}}{2\pi c^2}\int\sum_{\lambda}\left|\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\langle b\,|\,\mathbf{R}\,|\,a\rangle\right|^2\mathrm{d}\Omega\\&=\frac{\alpha\omega_{\mathbf{k}}^3 n_{\mathbf{k}}}{2\pi c^2}\,|\,\langle b\,|\,\mathbf{R}\,|\,a\rangle\,|^2\int\sin^2\theta_k\,\mathrm{d}\Omega=\frac{4\alpha\omega_{\mathbf{k}}^3 n_{\mathbf{k}}}{3\pi c^2}\,|\,\langle b\,|\,\mathbf{R}\,|\,a\rangle\,|^2.\end{aligned} \tag{4.128}$$

Dropping now the subscripts referring to the propagation vector and the polarization vector, which are no longer needed, the total probabilities per unit time for absorption and emission of a photon regardless of the direction of propagation or polarization are, respectively

$$W^{\mathrm{abs}}=\frac{4\alpha\omega^3 n}{3c^2}\,|\,\langle b\,|\,\mathbf{R}\,|\,a\rangle\,|^2,\qquad W^{\mathrm{em}}=\frac{4\alpha\omega^3(n+1)}{3c^2}\,|\,\langle b\,|\,\mathbf{R}\,|\,a\rangle\,|^2. \tag{4.129}$$

The only difference between these two formulas is that the absorption probability is proportional to n while the emission probability is proportional to $n+1$. This suggests to write W^{em} as a sum of two terms, $W^{\mathrm{em}}=W^{\mathrm{i}}+W^{\mathrm{s}}$, where

$$W^{\mathrm{i}}=\frac{4\alpha\omega^3 n}{3c^2}\,|\,\langle b\,|\,\mathbf{R}\,|\,a\rangle\,|^2,\qquad\text{and}\qquad W^{\mathrm{s}}=\frac{4\alpha\omega^3}{3c^2}\,|\,\langle b\,|\,\mathbf{R}\,|\,a\rangle\,|^2. \tag{4.130}$$

The W^{i} probability is identical to W^{abs} and is known as ***induced*** (or stimulated) emission probability per unit time, since it is proportional to n, the number of photons of frequency ω present in the radiation field. W^{s} is independent of n and is known as a ***spontaneous*** emission probability per unit time. Note that spontaneous emission can occur even if $n=0$, while the induced emission cannot.

$$W^{\mathrm{abs}}=W^{\mathrm{i}},\qquad W^{\mathrm{em}}=W^{\mathrm{abs}}+W^{\mathrm{s}}=W^{\mathrm{i}}+W^{\mathrm{s}}. \tag{4.131}$$

Atomic states may be degenerate and we need to modify the absorption and emission probabilities accordingly. Let W_{ab} be the transition probability per unit time from the state $|\alpha_a J_a M_a\rangle$, with degeneracy g_a, to the state $|\alpha_b J_b M_b\rangle$ with degeneracy g_b, where α_a and α_b represent the quantum numbers required to complete the specification of the states $|a\rangle$ and $|b\rangle$. To write down the new expressions we need to sum the probabilities for all possible transitions $|\alpha_a J_a M_a\rangle\to|\alpha_b J_b M_b\rangle$ and, since the probability for an electron being in any one of the initial states $|\alpha_a J_a M_a\rangle$ is $1/g_a$, the sum must be multiplied by this factor. Therefore the absorption probability to go from a state $|a\rangle$, with degeneracy g_a, to a state $|b\rangle$ with degeneracy g_b is

$$W_{ab}^{\mathrm{abs}}=\frac{1}{g_a}\sum_{M_a M_b}\frac{4\alpha\omega^3 n}{3c^2}\,|\,\langle\alpha_b J_b M_b\,|\,\mathbf{R}\,|\,\alpha_a J_a M_a\rangle\,|^2,\qquad E_a<E_b. \tag{4.132}$$

The emission probability to go from any of the sub-states in $|b\rangle$ to any of the sub-states in $|a\rangle$ is

$$W_{ba}^{\mathrm{em}}=\frac{1}{g_b}\sum_{M_a M_b}\frac{4\alpha\omega^3(n+1)}{3c^2}\,|\,\langle\alpha_a J_a M_a\,|\,\mathbf{R}\,|\,\alpha_b J_b M_b\rangle\,|^2,\qquad E_a<E_b. \tag{4.133}$$

The induced and spontaneous emission probabilities are, $(E_a<E_b)$,

$$W_{ba}^{\mathrm{i}}=\frac{1}{g_b}\sum_{M_a M_b}\frac{4\alpha\omega^3 n}{3c^2}\,|\,\langle\alpha_a J_a M_a\,|\,\mathbf{R}\,|\,\alpha_b J_b M_b\rangle\,|^2\;=\;\frac{g_a}{g_b}W_{ab}^{\mathrm{abs}}, \tag{4.134}$$

$$W_{ba}^{\mathrm{s}}=\frac{1}{g_b}\sum_{M_a M_b}\frac{4\alpha\omega^3}{3c^2}\,|\,\langle\alpha_a J_a M_a\,|\,\mathbf{R}\,|\,\alpha_b J_b M_b\rangle\,|^2. \tag{4.135}$$

Note that the probability of induced emission from the state $|b\rangle$ to the state $|a\rangle$, corresponding to the transition $|b\rangle \to |a\rangle + |\gamma\rangle$, is related to the absorption probability (transition $|a\rangle + |\gamma\rangle \to |b\rangle$) through

$$\frac{1}{g_a} W^{\mathrm{i}}_{ba} = \frac{1}{g_b} W^{\mathrm{abs}}_{ab}. \tag{4.136}$$

It is convenient to define the so-called *line strength* as

$$S_{ba} = S_{ab} = e^2 \sum_{M_a M_b} |\langle \alpha_b J_b M_b | \mathbf{R} | \alpha_a J_a M_a \rangle|^2. \tag{4.137}$$

Since there are no degeneracy factors (g_a, g_b) in this expression, the line strength is a symmetric quantity under the interchange $a \leftrightarrow b$. Using the line strengths, the transition probabilities can be written as

$$g_a W^{\mathrm{abs}}_{ab} = \frac{4\alpha\omega^3 n}{3e^2c^2} S_{ab} = g_b W^{\mathrm{i}}_{ba}, \tag{4.138}$$

$$g_b W^{\mathrm{s}}_{ba} = \frac{4\alpha\omega^3}{3e^2c^2} S_{ba}. \tag{4.139}$$

To do numerical calculations it may be useful to write these expressions in the SI system. This can be done by noting that $e^2 = 4\pi\varepsilon_0 \alpha \hbar c$. Thus

$$g_a W^{\mathrm{abs}}_{ab} = \frac{\omega^3 n}{3\pi\varepsilon_0 \hbar c^3} S_{ab}, \qquad g_b W^{\mathrm{s}}_{ba} = \frac{\omega^3}{3\pi\varepsilon_0 \hbar c^3} S_{ba}. \tag{4.140}$$

The operator $\mathbf{R}$ is an irreducible tensor (vector) operator and we can apply the Wigner-Eckart theorem to compute its matrix elements (see appendix B.6). In particular, if we write the individual components of the line strength as

$$S_{ab} = e^2 \sum_{M_a M_b} \Big[|\langle \alpha_b J_b M_b | R_x | \alpha_a J_a M_a \rangle|^2 + |\langle \alpha_b J_b M_b | R_y | \alpha_a J_a M_a \rangle|^2 + |\langle \alpha_b J_b M_b | R_z | \alpha_a J_a M_a \rangle|^2 \Big], \tag{4.141}$$

and apply (see equation (B.234) in page 318)

$$\sum_m \sum_{m'} \left| \left\langle \alpha j m \left| T^{(k)}_q \right| \alpha' j' m' \right\rangle \right|^2 = \frac{1}{2k+1} \left| \left\langle \alpha j \left\| \mathbf{T}^{(k)} \right\| \alpha' j' \right\rangle \right|^2, \tag{4.142}$$

for $k = 1$, we see that the three terms in that sum are each equal to

$$\frac{1}{3} |\langle \alpha_b J_b \| \mathbf{R} \| \alpha_a J_a \rangle|^2, \tag{4.143}$$

so that

$$S_{ab} = S_{ba} = e^2 |\langle \alpha_b J_b \| \mathbf{R} \| \alpha_a J_a \rangle|^2. \tag{4.144}$$

In the LS coupling this matrix element would be written as

$$\langle \alpha_b J_b \| \mathbf{R} \| \alpha_a J_a \rangle = \langle \alpha_b L_b S_b J_b \| \mathbf{R} \| \alpha_a L_a S_a J_a \rangle . \tag{4.145}$$

It can be seen (we skip this) that, since $\mathbf{R}$ does not contain spin operators, this matrix element can be transformed to

$$\langle \alpha_b L_b S_b J_b \| \mathbf{R} \| \alpha_a L_a S_a J_a \rangle = (-1)^{L_b + S_b + J_a + 1} \delta(S_a, S_b) \sqrt{(2J_b+1)(2J_a+1)} \times \begin{Bmatrix} J_b & 1 & J_a \\ L_a & S_b & L_b \end{Bmatrix} \langle \alpha_b L_b \| \mathbf{R} \| \alpha_a L_a \rangle . \tag{4.146}$$

The Kronecker δ and the $6j$ symbol enforce the following selection rules,[17] which are of course the same rules that were already deduced in the previous section for electric-dipole transitions

$$\Delta S = 0, \qquad \Delta L = 0, \pm 1, \qquad L_a + L_b \geq 1, \qquad \Delta J = 0, \pm 1, \qquad J_a + J_b \geq 1. \quad (4.147)$$

Therefore the line strength in LS coupling becomes

$$\begin{aligned} S_{ba} = S_{ab} &= e^2 |\, \langle \alpha_b J_b \,\|\, \mathbf{R} \,\|\, \alpha_a J_a \rangle \,|^2 \\ &= e^2 (2J_b + 1)(2J_a + 1) \times \left\{ \begin{array}{ccc} J_b & 1 & J_a \\ L_a & S & L_b \end{array} \right\}^2 |\, \langle \alpha_b L_b \,\|\, \mathbf{R} \,\|\, \alpha_a L_a \rangle \,|^2, \end{aligned} \quad (4.148)$$

with $S = S_a = S_b$.

4.5.2 Einstein coefficients and Planck's Law.

It is interesting to note that, already in 1916, Einstein proposed a treatment of emission and absorption of radiation by atoms before the development of quantum mechanics. He considered an enclosure containing a single kind of atoms (with only two energy levels) and radiation, in equilibrium at absolute temperature T. When one of those atoms interacts with radiation of energy density $\rho(\omega)$ per unit frequency interval it may suffer a transition from the lower to the upper level, $|a\rangle \to |b\rangle$, at a rate proportional to $\rho(\omega_{ba})$, with a proportionality constant that we may denote as B_{ab}. The time evolution of the number of atoms in the two states, N_a and N_b, due to this process is described by

$$-W_{ab}^{\rm abs} N_a = \frac{{\rm d}N_a}{{\rm d}t} = -\frac{{\rm d}N_b}{{\rm d}t} = -B_{ab} N_a \rho(\omega_{ba}). \quad (4.149)$$

An atom only has a significant probability of interacting with that part of the distribution $\rho(\omega)$ with a frequency close to $\omega_{ba} = (E_b - E_a)/\hbar$, which is the atom resonant frequency. Since atoms and radiation are in thermal equilibrium, symmetry arguments lead us to expect that the radiation field will also cause transitions from the upper to lower levels at a rate dependent on the energy density, but with a constant of proportionality B_{ba}. This is the process of induced (or stimulated) emission in which the radiation at angular frequency ω causes the atom to emit radiation of the same frequency,[18] and leads to the following evolution equation

$$-W_{ba}^{\rm i} N_b = \frac{{\rm d}N_b}{{\rm d}t} = -\frac{{\rm d}N_a}{{\rm d}t} = -B_{ba} N_b \rho(\omega_{ba}). \quad (4.150)$$

The relation between the rate of induced emission and Einstein's B coefficient is

$$W_{ba}^{\rm i} = B_{ba}\, \rho(\omega_{ba}). \quad (4.151)$$

The overall symmetry between up and down processes is broken by the spontaneous emission, which causes an atom to fall down to the lower level even when no external radiation is present. Einstein introduced the coefficient A_{ba} to represent the rate of this process. We now know that this coefficient corresponds to the spontaneous emission probability: $A_{ba} = W_{ba}^{\rm s}$, where $W_{ba}^{\rm s}$ was defined in the electric dipole approximation in equation (4.130). The evolution equation for this process is

$$-W_{ba}^{\rm s} N_b = \frac{{\rm d}N_b}{{\rm d}t} = -\frac{{\rm d}N_a}{{\rm d}t} = -A_{ba} N_b. \quad (4.152)$$

Of course, all three processes just described must be taken into account simultaneously. This leads to

$$\frac{dN_b}{dt} = N_a B_{ab} \rho(\omega_{ba}) - N_b B_{ba} \rho(\omega_{ba}) - N_b A_{ba}, \qquad \frac{dN_a}{dt} = -\frac{dN_b}{dt}. \quad (4.153)$$

[17] A very short introduction to $6j$ symbols is given in section B.5.3.

[18] This increase in the amount of light at the incident frequency is fundamental to the operation of lasers.

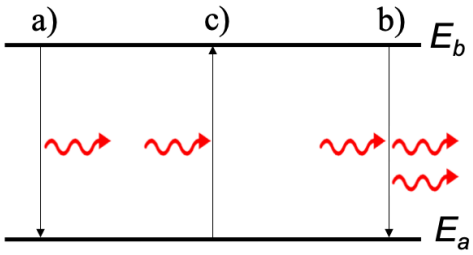


Figure 4.2: Two-level system employed by Einstein to deduce Planck's law. Note the spontaneous (a) and induced (b) emissions as well as the absorption (c) processes.

The first equation gives the rate of change of N_b in terms of the absorption, stimulated emission, and spontaneous emission, respectively. The second equation is a consequence of having only two levels, so that atoms leaving level b must go into level a. When $\rho(\omega) = 0$ the equations have a decaying exponential solution,

$$N_b(t) = N_b(0)e^{-A_{ba}t}, \tag{4.154}$$

implying a mean lifetime for the atom given by $1/\tau = A_{ba}$. At thermal equilibrium the rates of change of N_b and N_a are both zero, so from (4.153) we get

$$N_b A_{ba} = N_a B_{ab}\rho(\omega_{ba}) - N_b B_{ba}\rho(\omega_{ba}), \qquad \rho(\omega_{ba}) = \frac{N_b A_{ba}}{N_a B_{ab} - N_b B_{ba}}$$

$$\rho(\omega_{ba}) = \frac{A_{ba}}{B_{ba}} \frac{1}{\dfrac{N_a}{N_b}\dfrac{B_{ab}}{B_{ba}} - 1}. \tag{4.155}$$

Furthermore, statistical mechanics tells us that the number of particles (atoms in this case) with energy E, in thermal equilibrium at temperature T, is proportional to the Boltzmann factor,[19] $\exp(-E/kT)$, where k is the Boltzmann constant. Taking into account possible degeneracies in the states, g_a and g_b, we may write

$$N_a \sim g_a \exp\left(-\frac{E_a}{kT}\right), \qquad N_b \sim g_b \exp\left(-\frac{E_b}{kT}\right),$$

$$\frac{N_a}{N_b} = \frac{g_a}{g_b} \exp\left(\frac{E_b - E_a}{kT}\right) = \frac{g_a}{g_b} \exp\left(\frac{\hbar\omega_{ba}}{kT}\right). \tag{4.156}$$

From equations (4.155) and (4.156) we get

$$\rho(\omega_{ab}) = \frac{A_{ba}}{B_{ba}} \frac{1}{\dfrac{g_a B_{ab}}{g_b B_{ba}} \exp(\hbar\omega_{ab}/kT) - 1}. \tag{4.157}$$

When Einstein derived this equation it was already known that the energy density of radiation in thermal equilibrium with matter at a temperature T was given by Planck's law

$$\rho(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{\exp(\hbar\omega/kT) - 1}. \tag{4.158}$$

Comparing both equations we conclude that

$$A_{ba} = \frac{\hbar\omega^3}{\pi^2 c^3} B_{ba}, \qquad \frac{1}{g_a} B_{ba} = \frac{1}{g_b} B_{ab}. \tag{4.159}$$

[19] See appendix B.10.

The Einstein coefficients are properties of the atom and these relationships hold for any type of radiation. Their calculation from first principles could only be done after quantum mechanics was developed. Using equation (4.140) we can relate Einstein's A coefficient to the line strengths of the decay in the electric-dipole approximation,

$$A_{ba} = W^{\rm s}_{ba} = \frac{\omega^3}{3\pi\varepsilon_0\hbar c^3 g_b} S_{ba}. \tag{4.160}$$

The B coefficient is related to the rate of induced emission by $W^{\rm i}_{ba} = B_{ba}\,\rho(\omega_{ba})$. It is useful to write it also in terms of the A coefficient and the line strength,

$$B_{ba} = \frac{\pi^2 c^3}{\hbar\omega^3} A_{ba} = \frac{\pi}{3\varepsilon_0\hbar^2 g_b} S_{ba}, \tag{4.161}$$

$$B_{ab} = \frac{\pi}{3\varepsilon_0\hbar^2 g_a} S_{ba}. \tag{4.162}$$

The A_{ba} coefficient can be determined from the lifetime of the state, since $N_b(t) = N_b(0)e^{-A_{ba}t}$ and $1/\tau = A_{ba}$. Once this is done, the remaining coefficients can be computed using the previous relations.

When designing a laser it is very convenient to work with a configuration where the probability of stimulated emission (coefficient B) is larger than the probability of spontaneous emission (coefficient A). Since the ratio A_{ba}/B_{ba} is proportional to ω^3, it is clear that the higher the energy (frequency) of the EM radiation the more difficult it is to build the laser. The ratio of spontaneous over stimulated (induced) emission rates is

$$\frac{W^s}{W^i} = \frac{A}{B}\frac{1}{\rho(\omega)} = \exp\left(\frac{\hbar\omega_{ba}}{kT}\right) - 1. \tag{4.163}$$

It is worth stressing the fact that the quantum-mechanical treatment of the radiation field gives rise automatically to the spontaneous part of the emission probability, while in the semiclassical theory it has to be introduced *ad hoc* to reconcile Planck's radiation law with thermodynamic requirements. Indeed, Einstein was *forced to invent* spontaneous emission in order to reproduce Planck's formula.

4.5.3 Lifetime of the first excited state of Hydrogen.

As an example, let's compute the A_{ba} coefficient for the $2p \to 1s$ transition in Hydrogen. The formulas to employ are

$$A_{ba} = \frac{\omega^3}{3\pi\varepsilon_0\hbar c^3 g_b} S_{ab}, \quad \text{and} \quad S_{ab} = e^2 \sum_{M_a M_b} |\,\langle \alpha_b J_b M_b \,|\, \mathbf{R} \,|\, \alpha_a J_a M_a\rangle\,|^2.$$

Here we only have one electron, so $\mathbf{R}$ becomes $\mathbf{r}$ and

$$\begin{aligned} S_{ab} &= e^2\Big(|\,\langle 2p_{-1}\,|\,\mathbf{r}\,|\,1s\rangle\,|^2 + |\,\langle 2p_0\,|\,\mathbf{r}\,|\,1s\rangle\,|^2 + |\,\langle 2p_{+1}\,|\,\mathbf{r}\,|\,1s\rangle\,|^2\Big) \\ &= e^2\Big(|\,\langle 2p_{-1}\,|\,x\,|\,1s\rangle\,|^2 + |\,\langle 2p_0\,|\,x\,|\,1s\rangle\,|^2 + |\,\langle 2p_{+1}\,|\,x\,|\,1s\rangle\,|^2 + \\ &\quad |\,\langle 2p_{-1}\,|\,y\,|\,1s\rangle\,|^2 + |\,\langle 2p_0\,|\,y\,|\,1s\rangle\,|^2 + |\,\langle 2p_{+1}\,|\,y\,|\,1s\rangle\,|^2 + \\ &\quad |\,\langle 2p_{-1}\,|\,z\,|\,1s\rangle\,|^2 + |\,\langle 2p_0\,|\,z\,|\,1s\rangle\,|^2 + |\,\langle 2p_{+1}\,|\,z\,|\,1s\rangle\,|^2\Big). \end{aligned} \tag{4.164}$$

To do the actual calculations it is convenient to use spherical components for $\mathbf{r}$,

$$\begin{aligned} x &= -\frac{1}{\sqrt{2}}(r_+ - r_-), \qquad z = r_0, \qquad y = +\frac{i}{\sqrt{2}}(r_+ + r_-), \\ r_- &= \frac{1}{\sqrt{2}}(x - iy), \qquad r_0 = z, \qquad r_+ = -\frac{1}{\sqrt{2}}(x + iy), \end{aligned} \tag{4.165}$$

and write them in terms of spherical harmonics,

$$r_- = \sqrt{\frac{4\pi}{3}}\, r\, Y_1^{-1}, \qquad r_0 = z = \sqrt{\frac{4\pi}{3}}\, r\, Y_1^{0}, \qquad r_+ = \sqrt{\frac{4\pi}{3}}\, r\, Y_1^{+1}. \tag{4.166}$$

Therefore

$$\begin{aligned} S_{ab} = e^2\Big(&|\,\langle 2p_{-1}\,|\,r_-\,|\,1s\rangle\,|^2 + |\,\langle 2p_0\,|\,r_-\,|\,1s\rangle\,|^2 + |\,\langle 2p_{+1}\,|\,r_-\,|\,1s\rangle\,|^2 \\ &+ |\,\langle 2p_{-1}\,|\,r_0\,|\,1s\rangle\,|^2 + |\,\langle 2p_0\,|\,r_0\,|\,1s\rangle\,|^2 + |\,\langle 2p_{+1}\,|\,r_0\,|\,1s\rangle\,|^2 \\ &+ |\,\langle 2p_{-1}\,|\,r_+\,|\,1s\rangle\,|^2 + |\,\langle 2p_0\,|\,r_+\,|\,1s\rangle\,|^2 + |\,\langle 2p_{+1}\,|\,r_+\,|\,1s\rangle\,|^2\Big) \\ = e^2\frac{4\pi}{3}|\,\langle 2p\,|\,r\,|\,1s\rangle\,|^2\Big(&|\,\langle Y_1^{-1}\,|\,Y_1^{-1}\,|\,Y_0^0\rangle\,|^2 + |\,\langle Y_1^{0}\,|\,Y_1^{-1}\,|\,Y_0^0\rangle\,|^2 + |\,\langle Y_1^{1}\,|\,Y_1^{-1}\,|\,Y_0^0\rangle\,|^2 \\ &+ |\,\langle Y_1^{-1}\,|\,Y_1^{0}\,|\,Y_0^0\rangle\,|^2 + |\,\langle Y_1^{0}\,|\,Y_1^{0}\,|\,Y_0^0\rangle\,|^2 + |\,\langle Y_1^{1}\,|\,Y_1^{0}\,|\,Y_0^0\rangle\,|^2 \\ &+ |\,\langle Y_1^{-1}\,|\,Y_1^{1}\,|\,Y_0^0\rangle\,|^2 + |\,\langle Y_1^{0}\,|\,Y_1^{1}\,|\,Y_0^0\rangle\,|^2 + |\,\langle Y_1^{1}\,|\,Y_1^{1}\,|\,Y_0^0\rangle\,|^2\Big). \end{aligned}$$

This allows us to use directly *Gaunt's formula* (B.199) to compute the integral of the products of three spherical harmonics. They are also tabulated in several places (Weissbluth, table 11.1, page 246). The results we need are

$$\frac{4\pi}{3}\left|\langle Y_1^{-1}\,|\,Y_1^{-1}\,|\,Y_0^0\rangle\right|^2 = \frac{4\pi}{3}\left|\langle Y_1^{0}\,|\,Y_1^{0}\,|\,Y_0^0\rangle\right|^2 = \frac{4\pi}{3}\left|\langle Y_1^{1}\,|\,Y_1^{1}\,|\,Y_0^0\rangle\right|^2 = \frac{1}{3}, \tag{4.167}$$

and all other matrix elements are zero. Therefore

$$\begin{aligned} S_{ab} = e^2\frac{4\pi}{3}|\,\langle 2p\,|\,r\,|\,1s\rangle\,|^2\Big(&|\,\langle Y_1^{-1}\,|\,Y_1^{-1}\,|\,Y_0^0\rangle\,|^2 + |\,\langle Y_1^{0}\,|\,Y_1^{0}\,|\,Y_0^0\rangle\,|^2 + |\,\langle Y_1^{1}\,|\,Y_1^{1}\,|\,Y_0^0\rangle\,|^2\Big) \\ &= e^2|\,\langle 2p\,|\,r\,|\,1s\rangle\,|^2. \end{aligned}$$

For the radial integral we can use the Hydrogen radial functions or look up the result in tables, which is

$$|\,\langle 2p\,|\,r\,|\,1s\rangle\,|^2 = \left(\int u_{2\mathrm{p}}(r) u_{1\mathrm{s}}(r)\, r\,\mathrm{d}r\right)^2 = \frac{2^{15}}{3^9}\, a_0^2 \approx 1.665\, a_0^2, \tag{4.168}$$

where a_0 is the Hydrogen Bohr radius. Therefore,

$$S_{ab} = \frac{2^{15}}{3^9}\, e^2 a_0^2 \approx 1.665\, e^2 a_0^2. \tag{4.169}$$

In rectangular coordinates the calculation is a bit longer and would proceed as follows

$$\begin{aligned} |\,\langle 2p_{-1}\,|\,x\,|\,1s\rangle\,|^2 &= \left|\left\langle 2p_{-1}\,\middle|\,-\frac{1}{\sqrt{2}}(r_+ - r_-)\,\middle|\,1s\right\rangle\right|^2 = \frac{1}{2}\Big|\,\langle 2p_{-1}\,|\,r_+\,|\,1s\rangle - \langle 2p_{-1}\,|\,r_-\,|\,1s\rangle\,\Big|^2 \\ &= \frac{1}{2}\Big|\,\langle 2p_{-1}\,|\,r_+\,|\,1s\rangle\,\Big|^2 + \frac{1}{2}\Big|\,\langle 2p_{-1}\,|\,r_-\,|\,1s\rangle\,\Big|^2 - \Big|\,\langle 2p_{-1}\,|\,r_+ r_-\,|\,1s\rangle\,\Big| \\ = \frac{4\pi|\,\langle 2p\,|\,r\,|\,1s\rangle\,|^2}{3}&\left[\frac{1}{2}\Big|\,\langle Y_1^{-1}\,|\,Y_1^{+1}\,|\,Y_0^0\rangle\,\Big|^2 + \frac{1}{2}\Big|\,\langle Y_1^{-1}\,|\,Y_1^{-1}\,|\,Y_0^0\rangle\,\Big|^2 - \Big|\,\langle Y_1^{-1}\,|\,Y_1^{+1}Y_1^{-1}\,|\,Y_0^0\rangle\,\Big|\right] \\ &= \frac{4\pi}{3}\frac{1}{2}|\,\langle 2p\,|\,r\,|\,1s\rangle\,|^2\,\Big|\,\langle Y_1^{-1}\,|\,Y_1^{-1}\,|\,Y_0^0\rangle\,\Big|^2 = \frac{1}{3}\frac{1}{2}\frac{2^{15}}{3^9}\, a_0^2 = \frac{2^{14}}{3^{10}}\, a_0^2. \end{aligned}$$

The matrix element containing four spherical harmonics is zero because it is proportional to $\int_0^{2\pi} e^{i\varphi}\mathrm{d}\varphi = 0$, and the others follow from Gaunt's formula. The remaining non-zero matrix

elements are calculated in the same way. The list is

$$\begin{aligned} |\langle 2p_{-1}|x|1s\rangle|^2 = |\langle 2p_{+1}|x|1s\rangle|^2 = \frac{2^{14}}{3^{10}}a_0^2, \\ |\langle 2p_{-1}|y|1s\rangle|^2 = |\langle 2p_{+1}|y|1s\rangle|^2 = \frac{2^{14}}{3^{10}}a_0^2, \\ |\langle 2p_0|z|1s\rangle|^2 = \frac{2^{13}}{3^{10}}a_0^2. \end{aligned} \tag{4.170}$$

To get the line strength, the matrix elements in (4.170) must be added. The result is

$$S_{2p,1s} = e^2\sum|\langle 2p|\mathbf{r}|1s\rangle|^2 = 3\times 0.555\, e^2a_0^2 = 1.665\, e^2a_0^2. \tag{4.171}$$

To compute the Einstein coefficient note that the degeneracy factor for the 2p state is $g_{2\mathrm{p}} = 3$ and[20]

$$\omega_{2p,1s} = \frac{E_{2p}-E_{1s}}{\hbar} = \left(1-\frac{1}{2^2}\right)\frac{m_ec^2\alpha^2}{2\hbar} = \frac{3}{8}\frac{m_ec^2\alpha^2}{\hbar}. \tag{4.172}$$

Therefore

$$\begin{aligned} A_{2p,1s} = \frac{\omega^3}{3\pi\varepsilon_0\hbar c^3 g_b}S_{2p,1s} = \frac{\left(\frac{3}{8}\frac{m_ec^2\alpha^2}{\hbar}\right)^3}{3\pi\varepsilon_0\hbar c^3 g_b}\times 1.665\, e^2a_0^2 \\ = \frac{3\times 1.665\, m_ec^2\alpha^5}{128\hbar} = 6.28\times 10^8\,\mathrm{s}^{-1}. \end{aligned} \tag{4.173}$$

This can be translated into a spontaneous radiative lifetime for the 2p state of $\tau = 1/A = 0.16\times 10^{-8}$ sec. It should be noted that these results hold only for unpolarized light.

4.5.4 Emission intensities.

A relevant magnitude when analyzing spectra is that of emission intensities. Transitions between discrete states in atomic spectra in the laboratory are observed as a series of lines. If they are in the visible region of the spectrum, we will observe that some lines are shinier (more intense) than others. For a given sample of atoms, the line intensity is proportional to the emitted power (energy per unit of time). Thus, the total intensity of a spectral line of frequency ω_{ba} is given by the number of decays per unit time and unit of solid angle multiplied by the energy released in the decay,

$$I_{\mathrm{line}} = \frac{1}{4\pi}\hbar\omega_{ba}\frac{\mathrm{d}N_b}{\mathrm{d}t} = \frac{1}{4\pi}\hbar\omega_{ba}A_{ba}N_b, \tag{4.174}$$

where $A_{ba} = W^s_{ba}$ is the Einstein coefficient and N_b is the number of excited atoms in the upper (initial) level. Equation (4.140) provides a relation between the line strength and the A_{ba} Einstein coefficient, so we can write (in the SI system)

$$I_{\mathrm{line}} = \frac{1}{g_b}\frac{\hbar\omega_{ba}^4 S_{ba}N_b}{12\pi^2\varepsilon_0\hbar c^3}, \tag{4.175}$$

where g_b is the degeneracy factor for the initial b state.

Let us compute now the total power emitted in a given transition by a sample of atoms in thermal equilibrium with radiation of energy density ρ. If the energy released in the transition is $E_b - E_a = \hbar\omega_{ba}$, the total power is

$$\begin{aligned} P = \hbar\omega_{ba}(B_{ba}\rho + A_{ba})N_b = \hbar\omega_{ba}\left(\frac{\pi^2c^3}{\hbar\omega_{ba}^3}A_{ba}\rho + A_{ba}\right)N_b = \hbar\omega_{ba}A_{ba}N_b\left(\frac{\pi^2c^3}{\hbar\omega_{ba}^3}\rho + 1\right) \\ = \hbar\omega_{ba}A_{ba}N_b\left(\frac{\pi^2c^3}{\hbar\omega_{ba}^3}\frac{\hbar\omega_{ba}^3}{\pi^2c^3}\frac{1}{\exp(\hbar\omega_{ba}/kT)-1} + 1\right) = \frac{\hbar\omega_{ba}A_{ba}N_b}{1-\exp(-\hbar\omega_{ba}/kT)}. \end{aligned} \tag{4.176}$$

[20] We are not taking into account the spin degree of freedom here because $\Delta S = 0$ for transitions in LS coupling, which is the case in hydrogen. If we take into account the spin and set $g_{2\mathrm{p}} = 6$, then a factor 2 must be included also in equation (4.171) and the final result does not change.

4.5.5 Line broadening.

Up to this point, we have been assuming that the atomic states $|a\rangle$ and $|b\rangle$ are solutions of a time-independent Schrödinger equation, hence stationary states. However, this cannot represent the real system because we have found that there exists always a spontaneous emission process for all states except the lowest one, so they cannot be stationary. To make a self-consistent treatment of radiative processes we must take this into account. This can be done using time-dependent perturbation theory, which you have seen in your *Quantum Physics II* course. You can review it also by reading appendix (B.3), but here is a brief reminder.

Consider the Hamiltonian $\mathcal{H} = \mathcal{H}_0 + V$, where $\mathcal{H}_0$ is the time-independent unperturbed Hamiltonian, which we know how to solve exactly, and V is a small time-dependent perturbation. We look for solutions to the time-dependent Schrödinger equation

$$i\hbar\frac{\partial\psi(\mathbf{r},t)}{\partial t} = \mathcal{H}\psi(\mathbf{r},t) = (\mathcal{H}_0 + V)\psi(\mathbf{r},t) \tag{4.177}$$

in terms of an orthonormal set of $\mathcal{H}_0$ eigenfunctions $\phi_k(\mathbf{r},t) = \phi_k(\mathbf{r})e^{-iE_kt/\hbar}$, which satisfy $\mathcal{H}_0\phi_k = E_k\phi_k$. We try the following wavefunction

$$\psi(\mathbf{r},t) = \sum_k c_k(t)\phi_k(\mathbf{r})e^{-iE_kt/\hbar}. \tag{4.178}$$

Upon inserting this expression in both sides of the Schrödinger equation we have

$$i\hbar\frac{\partial\psi(\mathbf{r},t)}{\partial t} = \underbrace{\sum_k E_kc_k(t)\phi_k(\mathbf{r})e^{-iE_kt/\hbar}}_{\mathcal{H}_0\psi(\mathbf{r},t)} + i\hbar\sum_k \frac{\partial c_k}{\partial t}\phi_k(\mathbf{r})e^{-iE_kt/\hbar}, \tag{4.179}$$

$$(\mathcal{H}_0 + V)\psi(\mathbf{r},t) = \mathcal{H}_0\psi(\mathbf{r},t) + \sum_k Vc_k(t)\phi_k(\mathbf{r})e^{-iE_kt/\hbar}. \tag{4.180}$$

Equating the right-hand sides of the two previous equations we find

$$i\hbar\sum_k \frac{\partial c_k}{\partial t}\phi_k(\mathbf{r})e^{-iE_kt/\hbar} = \sum_k Vc_k(t)\phi_k(\mathbf{r})e^{-iE_kt/\hbar}. \tag{4.181}$$

Multiply this expression on the left by $\phi_\ell^*(\mathbf{r})$, integrate over $\mathbf{r}$, and apply the normalization condition $\langle\phi_\ell|\phi_k\rangle = \delta_{\ell k}$,

$$i\hbar\sum_k \frac{\partial c_k}{\partial t}\langle\phi_\ell|\phi_k\rangle\, e^{-iE_kt/\hbar} = \sum_k \langle\phi_\ell|V|\phi_k\rangle\, c_k(t)e^{-iE_kt/\hbar},$$

$$i\hbar\frac{\partial c_\ell}{\partial t} = \sum_k \langle\phi_\ell|V|\phi_k\rangle\, c_k(t)e^{i(E_\ell - E_k)t/\hbar}. \tag{4.182}$$

We have as many differential equations of this sort as the number of $\mathcal{H}_0$ eigenfunctions. By solving them we get the coefficients of expansion (4.178). The quantity $|c_\ell(t)|^2$ is the probability of finding the system in the state ϕ_l at the time t.

For our purpose here we will reduce the sum to a single term so that only two states, ϕ_k (initial) and ϕ_ℓ (final), are involved in the transition,

$$i\hbar\frac{\partial c_\ell}{\partial t} = \langle\phi_\ell|V|\phi_k\rangle\, c_k(t)e^{i(E_\ell - E_k)t/\hbar}. \tag{4.183}$$

A further assumption is that the initial state decays at a rate given by

$$c_k(t) = e^{-\Gamma t/2\hbar} = e^{-t/2\tau}, \qquad \tau = \frac{\hbar}{\Gamma} = \frac{1}{A}, \tag{4.184}$$

where τ is the mean life (or lifetime) of the state.[21] The integration of the differential equation for $c_\ell(t)$ under these assumptions gives[22]

$$c_\ell(t) = -\frac{i}{\hbar}\,\langle\phi_\ell|V|\phi_k\rangle \int_0^t e^{i(E_\ell - E_k)t_1/\hbar}\, e^{-\Gamma t_1/2\hbar}\, dt_1. \tag{4.185}$$

Now, calling $\alpha = -\Gamma/2\hbar < 0$ and $\beta = (E_l - E_k)/\hbar$ we have

$$\int_0^t e^{(\alpha+i\beta)t_1}\, dt_1 = \frac{e^{(\alpha+i\beta)t} - 1}{\alpha + i\beta}, \qquad \lim_{t\to\infty}\left|\frac{e^{(\alpha+i\beta)t} - 1}{\alpha + i\beta}\right|^2 = \frac{1}{\alpha^2 + \beta^2}, \tag{4.186}$$

because $e^{\alpha t} \to 0$ for $\alpha < 0$. Therefore,

$$|c_\ell(\infty)|^2 = \frac{|\,\langle\phi_\ell|V|\phi_k\rangle\,|^2}{(E_\ell - E_k)^2 + \frac{1}{4}\Gamma^2}. \tag{4.187}$$

Using the Fermi golden rule (4.117), $W_{\ell k} = \frac{2\pi}{\hbar}\,|\,\langle\phi_\ell|V|\phi_k\rangle\,|^2\,\delta(E_\ell - E_k)$, with the δ-function replaced by the density of states dN/dE, we can rewrite the previous expression as

$$|c_\ell(\infty)|^2\,\frac{dN}{dE} = \frac{\hbar}{2\pi}\,\frac{W_{\ell k}}{(E_\ell - E_k)^2 + \frac{1}{4}\Gamma^2}. \tag{4.188}$$

This formula has in principle a rather general applicability, but we are interested here in atomic spontaneous emission. Let $|B\rangle = |b;0\rangle$ be the initial state which consists of an atomic state $|b\rangle$ and zero photons in the radiation field, and $|A\rangle = |a;1\rangle$ the final state with one photon in the radiation field. The total energy of the system is related to the atomic and photon energies through

$$E_B - E_A = E_b - E_a - \hbar\omega. \tag{4.189}$$

We need to replace the transition probability per unit time $W_{\ell k}$ with the spontaneous emission probability $W^{\rm s}$, which is equal to the Einstein A coefficient. Performing the substitution $W_{\ell k} = W^{\rm s} = A = \Gamma/\hbar$ we arrive at

$$|c_\ell(\infty)|^2\,\frac{dN}{dE} = \frac{1}{2\pi}\,\frac{\Gamma}{(E_b - E_a - \hbar\omega)^2 + \frac{1}{4}\Gamma^2}. \tag{4.190}$$

This can also be written as

$$|c_\ell(\infty)|^2\,\frac{dN}{d\omega} = \frac{\hbar}{2\pi}\,\frac{\Gamma}{(E_b - E_a - \hbar\omega)^2 + \frac{1}{4}\Gamma^2}. \tag{4.191}$$

Obviously $|c_\ell(\infty)|^2$ is the probability for a transition $|B\rangle \to |A\rangle$ to occur within a time much longer that $\tau = \hbar/\Gamma$, therefore the product $|c_\ell(\infty)|^2\,(dN/d\omega)$ can be interpreted to be the number of modes per unit angular frequency interval that are excited by spontaneous emission, with any value of polarization and direction of propagation of the emitted photon. The following normalized function

$$L(\omega) = \frac{\hbar}{2\pi}\,\frac{\Gamma}{(E_b - E_a - \hbar\omega)^2 + \frac{1}{4}\Gamma^2}, \qquad \int L(\omega)\, d\omega = 1, \tag{4.192}$$

is known as a *Lorentzian line shape*. It appears in many different situations involving resonant processes. Note that the $L(\omega)$ function has a maximum value for $E_b - E_a = \hbar\omega$.

[21] Note that the probability of finding the particle in state k is proportional to $|c_k(t)|^2 = e^{-\Gamma t/\hbar} = e^{-t/\tau}$, hence the mean life is τ.

[22] In this step we are assuming that the perturbation V is approximately time-independent in the time interval $0 \to t$, so that the $\langle\phi_\ell|V|\phi_k\rangle$ matrix element can be pulled out of the integral.

This analysis shows that the initial excited atomic state is broadened to a width $\Gamma = \hbar/\tau$ because of the spontaneous emission process, resulting in a Lorentzian shape in the emission spectrum whose full width at half-maximum is Γ. It can be shown that a similar analysis for the absorption process would lead to the result that the absorption line is also broadened due to the finite lifetime of the excited state. This is to be expected since there is an intrinsic symmetry between absorption and emission. This broadening effect is usually called *natural*, since it is always present even if other sources of broadening (Doppler and pressure broadening) could be completely eliminated, which cannot.

Random motions of the atoms in gas give rise to a Doppler effect which broadens any given line. The profile of the Doppler broadening is Gaussian as a result of the Maxwellian distribution of atom velocities. Another source is the so-called *collision* or *pressure* broadening, since the atomic collision rate in gasses increases with the pressure. Besides radiating a free photon, an excited atom can lose energy by a more complicated process involving a collision with a neighbor atom or with the walls of the containing vessel. These collisions interrupt the process of light emission and effectively reduce the lifetime of the excited state, resulting in an additional broadening which has also a Lorentzian shape, but with the parameter $\Gamma = \hbar/\tau$ replaced by $\Gamma_c = \hbar/\tau_c$, where τ_c is the mean time between collisions.

4.6 Atomic spectra.

A few general features of atomic spectra are:

- Atomic spectra lines arise in most cases from single-electron excitations. In the visible (or near visible) region of the spectra, one of the outermost electrons is responsible for the transition. In the X-ray region an inner shell electron is involved.
- Most of the atomic spectra lines are of the electric dipole type, with transition probabilities in the range of 10^7 to 10^9 sec^{-1}. Magnetic dipole and electric quadrupole transitions are less probable by a factor of about 10^5.
- Very few spectra can be described by pure LS coupling or pure jj coupling. When going from light to heavy atoms there is a continuous transition from predominantly LS coupling to predominantly jj coupling.
- As the departure from LS coupling increases, intercombination lines (due to transitions with $\Delta S \neq 0$) become more probable. Such lines are weak in the spectra of light atoms and become fairly strong in heavy atoms.

The study of spectra is usually done by classifying the atoms into groups based on the number and type of electrons that are excitable. The best-understood spectrum is that of Hydrogen, whose main features are contained in the expression

$$E_n = -\frac{R_{\rm H}}{n^2}, \qquad \frac{hc}{\lambda} = R_{\rm H}\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right), \tag{4.193}$$

where

$$R_{\rm H} = \frac{R_\infty}{1+\dfrac{m}{M}} = \frac{\mu e^4}{2\hbar^2} = 109678.758 \text{ cm}^{-1} = 13.595 \text{ eV},$$

$$\mu = \frac{mM}{m+M}, \qquad R_\infty = \frac{me^4}{2\hbar^2} = 1097373.3157 \text{ cm}^{-1} = 13.6057 \text{ eV}.$$

M is the proton mass, m is the electron mass, $R_{\rm H}$ is the ionization potential of Hydrogen, and R_∞ is the Rydberg constant. The major series of lines in the spectrum of Hydrogen are those associated with $n_1 = 1, 2, 3, \ldots$, named as shown in table (4.2)

Series	n_1	n_2	Longest wavelength in Å
Lyman	1	$2, 3, \ldots$	1215.68
Balmer	2	$3, 4, \ldots$	6562.79
Ritz–Paschen	3	$4, 5, \ldots$	18751
Brackett	4	$5, 6, \ldots$	40510
Pfund	5	$6, 7, \ldots$	74560

Table 4.2: Main series in the Hydrogen spectra.

- Beyond Hydrogen, the alkali elements (Li, Na, K, Rb, Cs, Fr) have a single s electron outside of a filled s or p shell. The ground state for all of them is $^2\mathrm{S}_{1/2}$. Because of the spherical symmetry of the filled shells the outer electron moves in a central field, therefore the spectra of these elements have very close similarities to that of Hydrogen. The single electron in alkali elements is more loosely bound than the inner electrons and than the electron in Hydrogen, hence the ionization potentials of these elements are lower. However, the doublet splitting due to spin-orbit interaction is much larger than in Hydrogen.

  One would be tempted to use equation (4.193) in the valence electrons of these atoms. A better model for alkali atoms can be found by writing the energy of each (n, ℓ) term as

  $$E_{n\ell} = -\frac{R_H}{[n - \delta(\ell)]^2}, \tag{4.194}$$

  where $\delta(\ell)$ is a phenomenological parameter named *quantum defect*. According to this, the energy emitted in a transition $(n_2, \ell_2) \to (n_1, \ell_1)$ is

  $$\frac{hc}{\lambda} = R_\mathrm{H}\left(\frac{1}{[n_1 - \delta(\ell_1)]^2} - \frac{1}{[n_2 - \delta(\ell_2)]^2}\right). \tag{4.195}$$

  The quantum defect accounts for the penetration of the inner shells by the valence electrons. Although the quantum defect should in principle depend on both n and ℓ, it is empirically found to depend mainly on ℓ only. The discussion here is also valid for other atoms with a single valence electron as, for example, the singly charged ions of the divalent alkaline-earth metals: Be^+, Mg^+, Ca^+, etc.

Z	Element	Configuration	Term	Ion. Pot. (eV)
3	Li	[He]2s	$^2\mathrm{S}_{1/2}$	5.39
11	Na	[Ne]3s	$^2\mathrm{S}_{1/2}$	5.14
19	K	[Ar]4s	$^2\mathrm{S}_{1/2}$	4.34
37	Rb	[Kr]5s	$^2\mathrm{S}_{1/2}$	4.18
55	Cs	[Xe]6s	$^2\mathrm{S}_{1/2}$	3.89
87	Fr	[Rn]7s	$^2\mathrm{S}_{1/2}$	4.00

- Cu, Ag and Au have an s valence electron outside of a filled d shell. However, since the binding energies of the s and d electrons are comparable, their spectra are interpretable on the basis of either s or d excitations.

- Elements with a single p electron outside of closed shells are: B, Al, Ga, In and Tl. In boron, for example, the ground state configuration is $1s^2 2s^2 2p$. The spectrum of these atoms resembles that of an alkali atom but with the lowest ^2S term missing. In the same category are F, Cl, Br, I and At, whose ground state configurations are np^5. This configuration also gives rise to the single ^2P term but the multiplet structure is inverted because the shell is more than half-filled.

So far we have discussed atoms with a single valence electron (electrons outside closed shells). We briefly describe now the spectrum structure of atoms with two valence electrons or two valence vacancies.

- Helium has two valence electrons. The energy levels of the Helium atom were discussed in 3.1. The structure shows two families of levels according to the total spin of the atom: singlets and triplets. For the same value of L, the triplet $S = 1$ states are at lower energy than the $S = 0$ singlets. Due to the $\Delta S = 0$ selection rule there are not optical transitions between singlets and triplets.

- The alkaline earth elements: Be, Mg, Ca, Sr, Ba, Ra, Zn, Cd and Hg have also two electrons (ns^2, $n > 2$) outside closed shells. Their spectrum is generically similar to that of helium. However, the spin-orbit interaction is more relevant than in He and the selection rule ΔS is not as strict as for this element. The ionization energy of these elements is, evidently, smaller than for helium.

- For C, Si, Ge, Sn and Pb the two valence electrons are in a np^2 configuration. The ground term is $^3P_0$ (Hund's rules). In the first excited states of these elements, which is $np(n+1)s$, a departure from the LS coupling into the jj coupling is observed as n increases (for heavy elements). This discussion also applies to O, S, Se, Te and Po, since they have a np^4 structure that produces the same terms as np^2, but with inverted multiplet structure.

Elements with three valence electrons in a np^3 configuration are: N, P, As, Sb and Bi. The lowest term is ^4S while the other np^3 terms, ^2P and ^2D, are meta-stable because of the ΔS selection rule (heavy elements deviate from this behavior because LS coupling does not apply well to them). The noble gases Ne, Ar, Kr, Xn and Rn, show a stable np^6 configuration with a $^1S_0$ ground state.

There are elements in which the occupation of s orbitals competes with the occupation of d orbitals. These elements have s electrons outside partially filled d shells and can be classified according to their unfilled d shells. The $3d$ (iron) group: Sc, Ti, V, Cr, Mn, Fe, Co, Ni; the $4d$ group: Y, Zr, Nb, Mo, Tc, Ru, Rb, Pd; and the $5d$ group: Lu, Hf, Ta, W, Re, Os, Ir, Pt. Due to the large number of terms produced by the d orbital electrons, the spectra of these atoms are very rich in the visible and ultraviolet regions. Their electronic configuration with unfilled d shells is also at the origin of the important magnetic properties of these materials.

Finally, we have the materials characterized by unfilled f shells. These are the lanthanides ($4f$): Ce, Pr, Nd, Pn, Sm, Eu, Gd, Tb, Dy, Ho, Er, Tm, Yb; and the actinides ($5f$): Ac, Th, Pa, U, Np, Pu, Am, Cm, Bk, Cf. Again, rich spectra and peculiar magnetic properties arise from the partially filled shells.

4.6.1 X-ray spectra

X-rays are high-energy electromagnetic radiation with wavelengths ranging from 0.01 to 10 nanometers, corresponding to frequencies in the range 3×10^{16} Hz to 3×10^{19} Hz and energies in the range 100 eV to 100 keV. X-ray wavelengths are shorter than those of UV rays and typically longer than those of gamma rays.

Spectra originating from transitions by inner shell (core) electrons are known as characteristic X-ray spectra. They are produced when an electron from an inner shell is removed from the atom (ionization), creating a hole. This can be done by electron bombardment, *bremsstrahlung* X-ray irradiation (see below), or other means. The atom reverts to the ground state through transitions where the hole is filled by an electron of an outer shell accompanied by the release of an X-ray photon (See Fig.4.3).

The vacancy in a closed shell produces the same term as a single electron but with an inverted multiplet structure and the selection rules are the same as for a single electron: $\Delta n \neq 0$, $\Delta \ell = \pm 1$, $\Delta j = 0, \pm 1$. For historical reasons the electronic shells are named K, L, M, N,

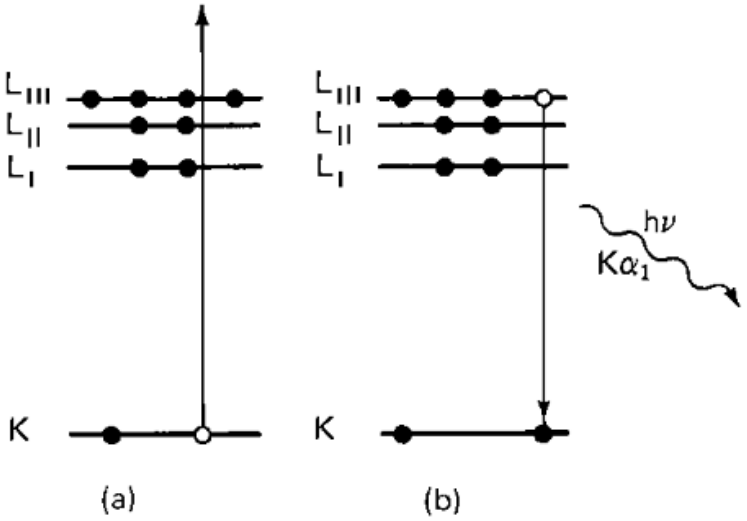


Figure 4.3: X-ray absorption and fluorescence. (a) Excitation; (b) deexcitation by fluorescence. Reproduced from Weissbluth 1978.

O... which was later discovered that correspond to principal quantum numbers 1, 2, 3, 4, 5... Therefore the K shell corresponds to $n = 1$, $\ell = 0$, $j = 1/2$, whereas the L shell is in principle three-fold degenerate, although screening effects and the spin-orbit interaction remove this degeneracy. In X-rays terminology the quantum number correspondence is

$$n = 2, \quad \ell = \begin{cases} 0, & j = \frac{1}{2} \quad (\mathrm{L_I}) \\ 1, & j = \frac{1}{2}, \frac{3}{2} \quad (\mathrm{L_{II}}, \mathrm{L_{III}}) \end{cases} .$$

The M level consists of 5 components, the N level of 7, etc.

If the initial vacancy occurs in the K shell, an electron from the L shell can fill the hole and emit an X-ray photon. Two possible transitions are allowed by the selection rules. Using the notation (n, ℓ, j), they can be written as $(2, 1, 3/2) \to (1, 0, 1/2)$ and $(2, 1, 1/2) \to (1, 0, 1/2)$. The emission lines corresponding to these transitions are labeled, according to the International Union of Pure and Applied Chemistry (IUPAC) nomenclature, as K-L_3 the more energetic one, and K-L_2 the less energetic. If the vacancy in the K shell is filled by a M-shell electron the lines are tagged as K-M_3 and K-M_2. After the K-L photon is emitted the vacancy moves to the L shell. Holes in the L shell are filled by electrons from the M shell, giving rise to a series of L lines. This process continues until the atom reaches the ground state again, so the emission of the K series is always accompanied by the emission of the L, M, N ... series, as appropriate. Historically the Sieghbahn notation (named after Manne Sieghbahn) was employed in X-ray spectra. Although still widely used in spectroscopy, this notation is unsystematic and often confusing. In the Sieghbahn notation K-L x-rays are known as K_α, K-M x-rays as K_β, etc. The correspondence between the two notations for some main spectral lines is specified in Fig 4.4.

The energy of an electron in an inner shell with principal quantum number n can be written as:

$$E_n = -\frac{(Z_n^{\mathrm{eff}})^2}{n^2} R_{\mathrm{H}}, \tag{4.196}$$

where Z_n^{eff} is the effective nuclear charge. The difference between Z and Z_n^{eff} is caused by the screening effect of the other electrons. This can be re-written as:

$$Z_n^{\mathrm{eff}} = Z - \sigma_n, \tag{4.197}$$

where σ_n is a phenomenological **screening parameter** accounting for the screening of the nucleus by the other electrons. For the K shell of sodium ($Z = 11$), for example, the measured

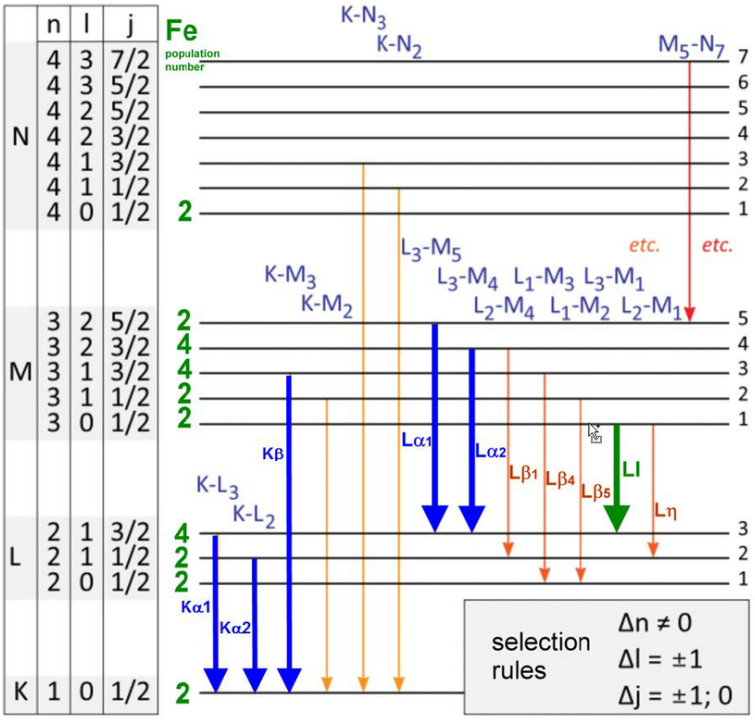


Figure 4.4: Energy level diagram showing electron transitions producing Fe K and L x-rays.

value is $E_1 = -1072$ eV as compared to

$$-\frac{(11)^2}{1^2}R_{\rm H} = -1650 \text{ eV}.$$

This implies $Z_n^{\rm eff} = 8.88$ and $\sigma_1 = 2.12$. This describes the combined screening due to the other electron in the K-shell and that of the remaining nine electrons in the higher shells. It can be seen that the value $\sigma_1 = 2.12$ may be interpreted as a contribution of ~ 1 from the other electron in the K shell plus another contribution of ~ 1 from the remaining nine electrons. Most of the probability density for these outer electrons lies outside the K shell, but there is still some fraction of it closer to the nucleus than the expected value of the K shell radius.

The energy of an X-ray transition $n' \to n$ is given by:

$$\hbar\omega = |E_{n'} - E_n| = \left| -\frac{(Z_{n'}^{\rm eff})^2}{(n')^2} + \frac{(Z_n^{\rm eff})^2}{(n)^2} \right| R_{\rm H}. \tag{4.198}$$

In practice, a crude approximation as $Z_n^{\rm eff} = Z_{n'}^{\rm eff} = (Z - \sigma_n)$ can be done to get, for the K shell,

$$\hbar\omega \approx (Z - \sigma_{\rm K})^2 R_{\rm H} \left(\frac{1}{1^2} - \frac{1}{(n')^2} \right), \tag{4.199}$$

where $n' > 1$ and, for heavy atoms, $\sigma_{\rm K} \sim 3$. A similar expression applies to the L shell:

$$\hbar\omega \approx (Z - \sigma_{\rm L})^2 R_{\rm H} \left(\frac{1}{2^2} - \frac{1}{(n')^2} \right), \tag{4.200}$$

with $n' > 2$ and $\sigma_{\rm L} \sim 10$. These expressions are examples of the so-called called *Moseley's law*, which has just an empirical justification. The fact that the $Z_n^{\rm eff} = Z_{n'}^{\rm eff}$ approximation can be applied means that the transition energy is dominated by the energy of the lower shell. Moseley's law and equation (4.196) do not explain the sub-shell structure shown in figure (4.4).

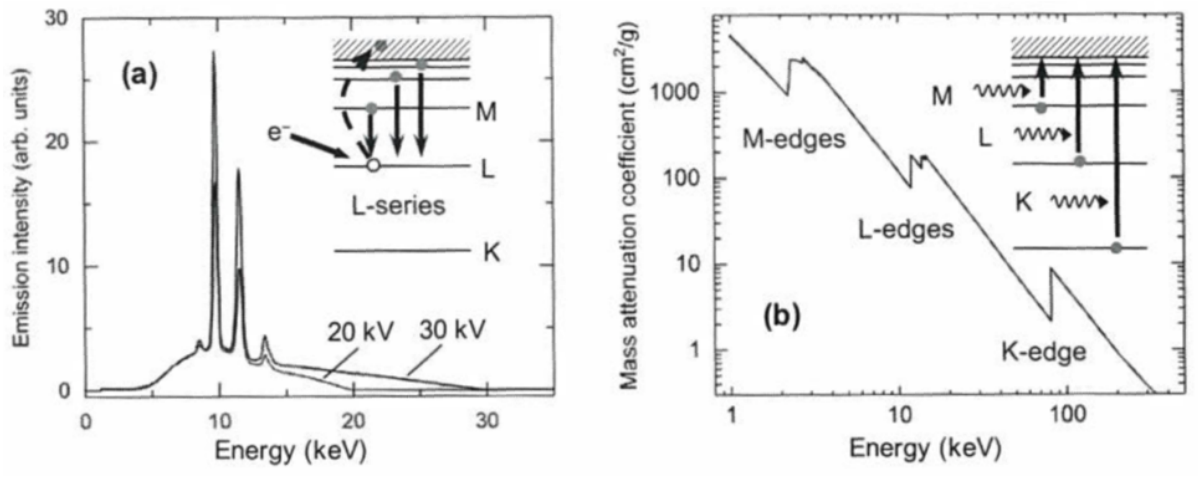


Figure 4.5: X-ray emission and absorption spectra of gold. (a) Emission spectrum at two different electron voltages showing sharp lines caused by L-series transitions. (b) Mass attenuation coefficient, which is proportional to the absorption coefficient. Reproduced from Fox (2019)

The left plot of figure (4.5) shows the X-ray emission spectra of gold for two different energies of the electron beam which collides with the gold target. The sharp peaks are generated by radiative transitions following the ejection of L-shell electrons. The continuum part of the spectrum is due to bremsstrahlung.[23] The right plot shows the gold mass attenuation coefficient (proportional to the absorption coefficient) for X-rays of energies up to 300 keV. A series of sharp edges can be seen. They occur when the incoming X-ray photon has enough energy to promote an electron from an inner shell to empty states above the highest occupied shell (either bound or ionized states). The absorption probability decreases beyond the edges due to the reduced overlap between the initial localized electron wave function and the delocalized states of the continuum, as described by the photoelectric cross-section. The energy of the absorption edge can be derived from equation (4.198) setting $E_{n'} = 0$, since the binding energy of valence electrons is negligible compared to X-ray energies. We have

$$E^{\text{edge}} = \frac{(Z_n^{\text{eff}})^2}{n^2} R_{\text{H}} \equiv \frac{(Z - \sigma_n)^2}{n^2} R_{\text{H}}. \tag{4.201}$$

The absorption edge provides a measurement of the inner shell energy.

The emission of X-rays is not the only way for an atom with a vacancy in an inner shell to loose energy. The excitation energy of the atom can be decreased by ejecting an electron from the outer shells without the emission of any photon. Such radiationless deexcitation of atoms by emission of electrons rather than X-rays is known as the *Auger* effect, and the emitted electrons are called Auger electrons. Of course, an Auger process leaves the atom in a doubly ionized state which can further relax by photon emission, so an Auger electron is usually accompanied by X-rays.

[23] Bremsstrahlung radiation is produced when energetic electrons scatter off atoms without ejecting a core electron from the target. The acceleration of the electron associated with its change of direction causes it to radiate energy in the form of X-rays.

4.7 Problems.

1. Write the wave function corresponding to the three-photon state $|1_{\mathbf{k}_1,+}, 1_{\mathbf{k}_2,+}, 1_{\mathbf{k}_3,-}\rangle$ in the Fock space.

2. A general two-photon state can be written as

$$|\text{state}\rangle = \sum_{\mathbf{k}_1,\lambda_1,\mathbf{k}_2,\lambda_2} f(\mathbf{k}_1, \hat{\mathbf{e}}_{\lambda_1}; \mathbf{k}_2, \hat{\mathbf{e}}_{\lambda_2}) a^\dagger_{\mathbf{k}_1,\lambda_1} a^\dagger_{\mathbf{k}_2,\lambda_2} |0\rangle,$$

where, apart from normalization, $f(\mathbf{k}_1, \hat{\mathbf{e}}_{\lambda_1}; \mathbf{k}_2, \hat{\mathbf{e}}_{\lambda_2})$ is any function of $\mathbf{k}_1, \hat{\mathbf{e}}_{\lambda_1}$ and $\mathbf{k}_2, \hat{\mathbf{e}}_{\lambda_2}$ (f can be called the wave function of the state).
Show that this state may be written as

$$|\text{state}\rangle = \sum_{\mathbf{k}_1,\lambda_1,\mathbf{k}_2,\lambda_2} g(\mathbf{k}_1, \hat{\mathbf{e}}_{\lambda_1}; \mathbf{k}_2, \hat{\mathbf{e}}_{\lambda_2}) a^\dagger_{\mathbf{k}_1,\lambda_1} a^\dagger_{\mathbf{k}_2,\lambda_2} |0\rangle,$$

with $g(\mathbf{k}_1, \hat{\mathbf{e}}_{\lambda_1}; \mathbf{k}_2, \hat{\mathbf{e}}_{\lambda_2}) = \{f(\mathbf{k}_1, \hat{\mathbf{e}}_{\lambda_1}; \mathbf{k}_2, \hat{\mathbf{e}}_{\lambda_2}) + f(\mathbf{k}_2, \hat{\mathbf{e}}_{\lambda_2}; \mathbf{k}_1, \hat{\mathbf{e}}_{\lambda_1})\}/2$, symmetric under the interchange of labelling. Discuss the implications of this result.

3. The free-radiation field state inside a cubic enclosure can be written as

$$|c\rangle = \exp\left(-\frac{1}{2}|c|^2\right) \sum_{n=0}^{\infty} \frac{c^n}{\sqrt{n!}} |n\rangle ,$$

where $c = |c|e^{i\delta}$ is any complex number and $|n\rangle$ is a state of n photons with some specific wavevector $\mathbf{k}$ and transverse polarization λ, omitted in the expression ($|n\rangle \equiv |n_{\mathbf{k}\lambda}\rangle$). Such a state $|c\rangle$ is called *a coherent state* and is extremely useful in quantum optics and other bosonic quantum field theories. Derive the following properties of $|c\rangle$.

   (a) $|c\rangle$ is normalized: $\langle c|c\rangle = 1$.

   (b) $|c\rangle$ is an eigenstate of $a_{\mathbf{k}\lambda}$ with the complex eigenvalue c: $a_{\mathbf{k}\lambda}|c\rangle = c\,|c\rangle$.

   (c) The expected number of photons in the enclosure $\overline{N} \equiv \langle c|N|c\rangle$ is equal to $|c|^2$.

   (d) The RMS fluctuation of the photon number is given by:

$$(\Delta N)^2 \equiv \langle c|N^2|c\rangle - \overline{N}^2 = |c|^2.$$

   (e) The expected value of the electric field $\mathbf{E}$ is given by:

$$\overline{\mathbf{E}} \equiv \langle c|\mathbf{E}|c\rangle = -2\sqrt{\frac{2\pi\hbar\omega_{\mathbf{k}}}{V}}|c|\sin(\mathbf{k}\cdot\mathbf{r} - \omega_{\mathbf{k}}t + \delta)\ \hat{\mathbf{e}}_{\mathbf{k}\lambda},$$

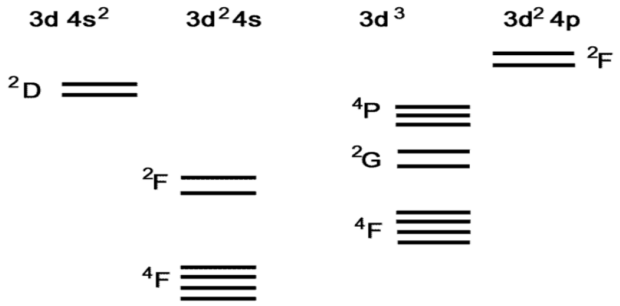
   where V is the volume of the cubic enclosure.

   (f) The RMS fluctuation ΔE in the electric field is given by:

$$(\Delta E)^2 \equiv \langle c|E^2|c\rangle - \langle c|E|c\rangle^2 = \frac{2\pi\hbar\omega_{\mathbf{k}}}{V}.$$

4. Using the properties of the 6j Wigner coefficients demonstrate the E1 selection rules: $\Delta L = 0, \pm 1$, $L_a + L_b \geq 1$, $\Delta J = 0, \pm 1$ and $J_a + J_b \geq 1$

5. The diagram shows the energy levels of Ti^+. Indicate all possible electric-dipole transitions (E1) in LS coupling assuming all multiplets are normal (higher J terms are more energetic).



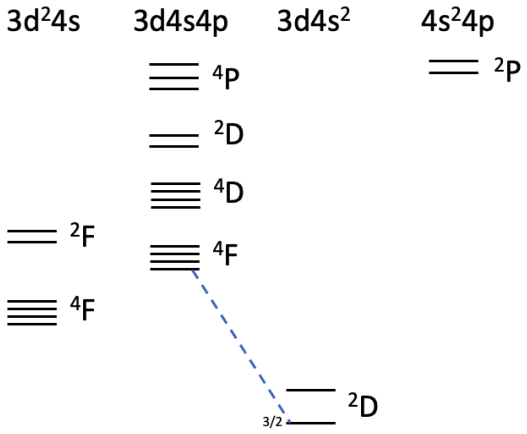
6. The diagram shows the ground state and first excited states of Sc (Scandium, $Z = 21$).

   a) Indicate the parity of every state.

   b) Make a table with the allowed electric dipole (E1) transitions. Assume normal ordering of the multiplets.

   c) Indicate, as a function of two 6-j symbols, the line-strength ratio between the $^2P_{3/2} \rightarrow {}^2D_{5/2}$ and the $^2P_{1/2} \rightarrow {}^2D_{3/2}$ lines.

   d) The transition indicated with a dashed line is experimentally observed. Discuss the presence of that transition in the spectrum of Scandium.



7. For each of the following transitions among LS-coupled levels, state whether it is allowed or forbidden for electric-dipole transitions. For forbidden transitions, which

selection rules(s) is/are violated?

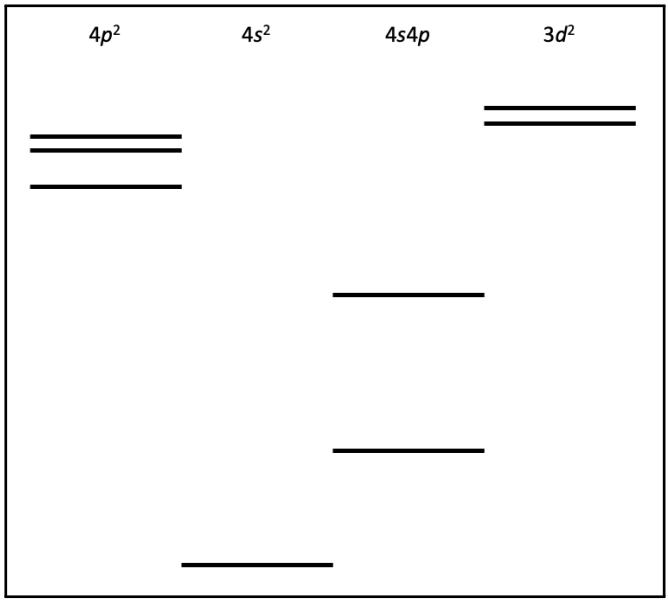
$$^3P_1 \rightarrow {}^1S_0$$
$$^2P_{3/2} \rightarrow {}^2S_{1/2}$$
$$^2D_{3/2} \rightarrow {}^2S_{1/2}$$
$$^4D_{3/2} \rightarrow {}^2P_{1/2}$$
$$^3F_4 \rightarrow {}^3G_4$$
$$^3G_4 \rightarrow {}^3H_5$$
$$^3H_5 \rightarrow {}^3G_3$$
$$^4I_{11/2} \rightarrow {}^4K_{11/2}$$

8. Determine the possible atomic terms for the p^2 and sp electronic configurations. Determine also the electric dipole transitions which can take place from the p^2 to sp terms assuming LS coupling.

9. The electronic configuration of the phosphorus ground state is $[\mathrm{Ne}](3\mathrm{s})^2(3\mathrm{p})^3$.

   - Find the values of L, S and J for the phosphorous ground state.
   - Find the possible values of L, S and J for the excited configuration $(3\mathrm{p})^2(4\mathrm{s})$.
   - One of the terms of the previous excited configuration is ^{4}P. Taking into account that the energy increases with the value of J, draw a Grotrian diagram showing the levels of this term and the ground-state level. Show also the corresponding parities.
   - Draw all possible radiative transitions between the previous levels. For each transition list the appropriate selection rules and the type of radiation emitted.

10. The next figure shows the ground state of Calcium ($[\mathrm{Ar}]4s^2$), also named as Ca I, and some of its first excited terms. Over every term there is its corresponding configuration.



a) Assign quantum numbers, L and S, to each term using the Hund rules and indicate, with spectroscopic notation, which one corresponds to each term.

b) In the $3d^2$ configuration there are terms that are in principle possible, but are not observed in reality. Discuss the relation of the energy of these terms with the ionization energy of Ca.

c) Make a diagram showing the levels of the terms due to the spin-orbit interaction, sorted in energy according to Hund's rules.

d) Draw in the previous diagram the allowed dipolar electric transitions (E1) between the considered levels.

e) Aside from the terms of the figure, Ca I presents a ^{3}P term corresponding to the [Ar]$4s5p$ configuration. This term splits into three levels with energies: 36547.688, 36554.749 and 36575.119 cm^{-1} above the ground state. Compute the approximate value of $\hbar^2\zeta$, where ζ is the proportionality factor in the spin-orbit matrix elements, $\langle \alpha LSJM_J \,|\, \mathcal{H}_{\rm so} \,|\, \alpha LSJM_J \rangle$, of multielectron atoms explained in Eq. (3.92).

11. For an F multiplet decaying into a D multiplet, find the relative intensity of $^2F_{7/2} \to {}^2D_{5/2}$ and $^2F_{5/2} \to {}^2D_{3/2}$ transitions.
Hint: you need to use a 6j table. Assume that the difference between the frequencies of the photons emitted in both transitions is negligible, so that the relative intensity can be approximated by the ratio of line strengths.

12. **Thomson scattering:** When discussing the interaction of the EM radiation field we neglected the interaction hamiltonian given by

$$\mathcal{H}_{\rm int} = \frac{e^2}{2mc^2} A^2$$

in equation (4.69). This term is very relevant when we consider the scattering of photons with energies $\hbar\omega_{\mathbf{k}}$ much larger than the binding energy of the electrons,

such that the electrons can be considered to be free (only their kinetic energy has to be accounted for), but much smaller than mc^2. This last condition implies that the energy $\hbar\omega_{\mathbf{k}'}$ of the scattered photon ($\mathbf{k}' \neq \mathbf{k}$) is not changed, $\hbar\omega_{\mathbf{k}'} = \hbar\omega_{\mathbf{k}}$, because for small recoil momenta the recoil energy may be neglected.

1) Show that

$$\langle b; 1_{\mathbf{k}'\lambda'} \,|\, \mathcal{H}_{\text{int}} \,|\, a; 1_{\mathbf{k}\lambda} \rangle = \frac{2\pi e^2 \hbar}{mV\omega_{\mathbf{k}}} \left[\hat{\mathbf{e}}_{\mathbf{k}'\lambda'} \cdot \hat{\mathbf{e}}_{\mathbf{k}\lambda}\right] \delta_{\mathbf{p}_b + \hbar\mathbf{k}', \mathbf{p}_a + \hbar\mathbf{k}},$$

where a and b indicate the electron state. Recall that, since we are considering free electrons in a cubic enclosure of volume V (we can disregard the electron spin in the exercise),

$$\langle \mathbf{r}|a\rangle = \frac{1}{\sqrt{V}} e^{i\mathbf{p}_a \mathbf{r}/\hbar}, \qquad \langle \mathbf{r}|b\rangle = \frac{1}{\sqrt{V}} e^{i\mathbf{p}_b \mathbf{r}/\hbar}.$$

2) We are not interested in the recoil of the free electron. Thus, we can sum over all the possible final momentum values of $\mathbf{p}_b$ to obtain

$$\left\langle 1_{\mathbf{k}',\lambda'} \,\middle|\, \mathcal{H}_{\text{int}} \,\middle|\, 1_{\mathbf{k},\lambda} \right\rangle = \sum_{\mathbf{p}_b} \frac{2\pi e^2 \hbar}{mV\omega_{\mathbf{k}}} \hat{\mathbf{e}}_{\mathbf{k}'\lambda'} \cdot \hat{\mathbf{e}}_{\mathbf{k}\lambda} \delta_{\mathbf{p}_b + \hbar\mathbf{k}', \mathbf{p}_a + \hbar\mathbf{k}} = \frac{2\pi e^2 \hbar}{mV\omega_{\mathbf{k}}} \hat{\mathbf{e}}_{\mathbf{k}'\lambda'} \cdot \hat{\mathbf{e}}_{\mathbf{k}\lambda}.$$

Now, use equations (4.117) and (4.120) and show that the transition probability to a group of states is

$$\frac{dW^{\mathbf{k}\lambda \to \mathbf{k}'\lambda'}}{d\Omega} = \frac{c}{V} r_0^2 \left[\hat{\mathbf{e}}_{\mathbf{k}'\lambda'} \cdot \hat{\mathbf{e}}_{\mathbf{k}\lambda}\right]^2,$$

where $r_0 = e^2/(mc^2) \approx 2.818$ f is called the *classical electron radius*.

3) If the transition probability is divided by c/V, which is the photon flux, we obtain the Thomson cross-section

$$\frac{d\sigma}{d\Omega} = r_0^2 \left[\hat{\mathbf{e}}_{\mathbf{k}'\lambda'} \cdot \hat{\mathbf{e}}_{\mathbf{k}\lambda}\right]^2.$$

We shall study the physical meaning of a cross-section in Chapter 5.

We want to obtain the unpolarized cross-section summing over the polarization of final states and averaging over the polarization of initial states. Demonstrate that

$$\frac{d\sigma^{\text{unp}}}{d\Omega} = \frac{r_0^2}{2} \sum_{\lambda=1}^{2} \sum_{\lambda'=1}^{2} \left[\hat{\mathbf{e}}_{\mathbf{k}'\lambda'} \cdot \hat{\mathbf{e}}_{\mathbf{k}\lambda}\right]^2 = \frac{r_0^2}{2}(1 + \cos^2\theta_k),$$

where the θ_k angle is defined in a reference system where $\hat{\mathbf{k}} = \mathbf{k}/k$ is the z axis unitary vector (see figure 4.1).

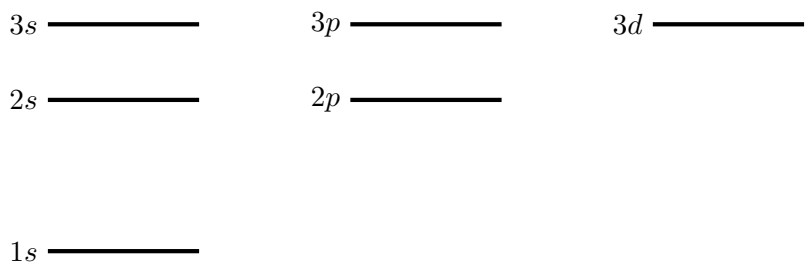
4) Now, integrate this result over the solid angle to show that the total cross-section is

$$\sigma_{\text{total}} = \frac{8\pi}{3} r_0^2 = 6.65 \times 10^{-25} \text{ cm}^2.$$

13. What is the the $3p \to 1s$ Einstein A coefficient in Hydrogen?
Hint: you may consider that:

$$\left(\int_0^\infty u_{n\mathrm{p}}(r) u_{1\mathrm{s}}(r)\, r\, \mathrm{d}r\right)^2 = \frac{2^8 n^7 (n-1)^{2n-5}}{(n+1)^{2n+5}} a_0^2.$$

14. The following Grotrian diagram shows (not to scale) the energies of the six first Hydrogen orbitals



Draw, using arrows in this diagram, all possible E1 transitions. Compute the lifetime of the $3p$ state. You will need the following formula

$$\left(\int_0^\infty u_{n\mathrm{p}}(r) u_{2\mathrm{s}}(r)\, r\, \mathrm{d}r\right)^2 = \frac{2^{17} n^7 (n^2-1)(n-2)^{2n-6}}{(n+2)^{2n+6}} a_0^2,$$

where $u_{n\ell}(r) = r R_{n\ell}(r)$ is the radial part of the Hydrogen-like wave functions.

15. A Hydrogen atom in its ground state is excited by a time-dependent perturbation

$$\mathcal{H}_0 = e\mathcal{E}_0 z e^{-t/\tau}, \qquad t \geq 0.$$

Find, at first order in time-dependent perturbation theory, the transition probability to the $|n\ell m\rangle$ state.
Hint: Consider: $z = r\sqrt{\frac{4\pi}{3}} Y_{10}(\theta, \phi)$ and the radial integral in the hint to problem (13).

16. We are investigating hydrogen plasma at a temperature of 4500 °C. Calculate the probability, per atom and second, for stimulated emission of photons between the $2p$ and $1s$ states if the lifetime of $2p$ state is 1.6 ns.

17. Replace the Coulombian interaction between a proton and an electron with a three-dimensional harmonic oscillator

$$V(r) = V_0 + \frac{1}{2} m\omega^2 r^2,$$

where r is the distance between the two charges and m is the electron mass.

a) Show that separation of variables in cartesian coordinates turns this system into three one-dimensional oscillators. Take the solution of a one-dimensional oscillator from a textbook and demonstrate that the allowed energies are

$$E_n = V_0 + \left(n + \frac{3}{2}\right)\hbar\omega,$$

where n_x, n_y, n_z are integer numbers and $n = n_x + n_y + n_z$. What is the wave function for a generic state $|n_x, n_y, n_z\rangle$ of the three-dimensional oscillator? Write explicitly the wave function for $|0,0,0\rangle$ (the ground state) and for $|1,0,0\rangle$ (one of the first excited states).

b) Discuss the degeneracy of E_n, in particular that of the ground state and first excited states of the oscillator.

c) Determine the parameters V_0 and ω if the energies of the ground and first excited states coincide with the ground and first excited states of the Hydrogen atom.

d) Determine the lifetime of the first excited state of the three-dimensional oscillator.

18. Radiative transitions from $n = 109$ to $n = 108$ of Hydrogen states are observed in radio–astronomy. Compute the frequency and wavelength of the radiation emitted in this transition. This same transition has been also observed in excited helium atoms. Without performing any calculation try to estimate the frequency of the radiation emitted. This kind of radiative transitions are very difficult to observe in laboratory experiments, can you tell why?

19. The ground-state electronic configuration of the alkali metal cesium, Cs, is [Xe]$6s$. The approximate values of the quantum defects are: $\delta(0) \approx 4.11$, $\delta(1) \approx 3.61$, $\delta(2) \approx 2.45$, $\delta(3) \approx 0.02$. The $6s$ electron is promoted to the $7p$ shell. What wavelengths can be observed in E1 emission?

20. Francium, Fr, is an alkali metal with electronic configuration [Rn]$7s^1$. Its first ionization potential is equal to 4.07 eV, and the wavelength of the transition $8p \to 7s$ is 428 nm.

a) Compute the values of the quantum defects $\delta(0)$ and $\delta(1)$ for Fr.

b) Make an estimation of the $7p \to 7s$ transition wavelength.

c) Compute the minimum value of $\delta(2)$ for the $8p \to 6d$ (emission) transition to be possible.

21. To study the composition of an alloy sample of iron ($Z = 26$) and manganese ($Z = 27$), we plan to use X rays with energy equal to the average value of the two K shell absorption edges of these metals. Compute this average energy.

22. A sample of tritium atoms is excited into $2p$ states. Tritium is unstable and decays as ${}^3_1\mathrm{H} \to {}^3_2\mathrm{He} + e^- + \bar{\nu}_e$. What is the relative intensity of the $3p \to 1s$ and $2p \to 1s$ lines of the resulting ${}^3_2\mathrm{He}^+$ ions from the tritium decays?

Hint: you need to take into account the probability amplitudes for an initial $2p$ tritium state to end up populating (after the tritium decay) the $3p$ and $2p$ ${}^3_2\mathrm{He}^+$ states.

23. A gas enclosure of 125 cm^3 volume contains Lithium vapour at thermal equilibrium at a temperature of 970 K and a pressure of 135 Pa. The radiated power for the $2p \to 2s$ transition, that corresponds to 670.8 nm wavelength, is 10.41 mW.[a]

    a) What is the ratio of the $2s$ and $2p$ populations, i.e. $\frac{N_{2p}}{N_{2s}}$.

    b) Calculate the lifetime of the $2p$ state of Lithium.

[a] In fact there are two levels in the ^{2}P Li term, $^2P_{3/2}$ and $^2P_{1/2}$. Their transition wavelength differs less than the quoted precision and can be considered degenerate in this problem.

Chapter 5

Scattering.

Particles bound by a potential, as electrons in an atom, have a discrete spectrum of negative energies. Free particles, on the other hand, may have any positive energy in a continuum spectrum. Scattering theory studies processes where a continuum initial state is transformed into a continuum final state through the action of some potential. By performing scattering measurements is how we primarily get information about distributions of mass, charge and potential energy for molecular and atomic systems. At a more fundamental level, it is also the primary way to study the properties of elementary particles and their interactions. This chapter deals mainly with the theory of scattering in a simple but important case, the elastic scattering of a non-relativistic particle in a local potential. In section 5.9, a brief introduction to scattering by complex potentials and absorption processes is given.

5.1 Introduction: classical scattering theory.

Consider a particle that approaches a scattering center (a heavy nucleus, for example) with an impact parameter b (see figure (5.1)). Provided the interaction is symmetric in the azimuth angle, the trajectory remains in a plane and the particle emerges at some scattering angle θ.

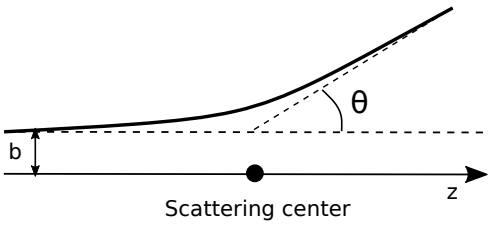


Figure 5.1: Impact parameter of a classic scattering trajectory.

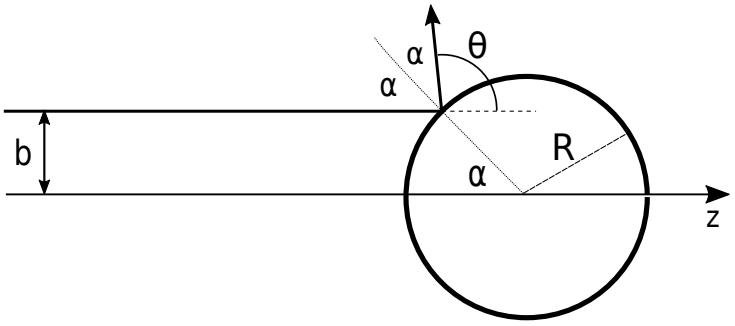


Figure 5.2: Elastic hard-sphere scattering. According to this figure, $b = R \sin \alpha$.

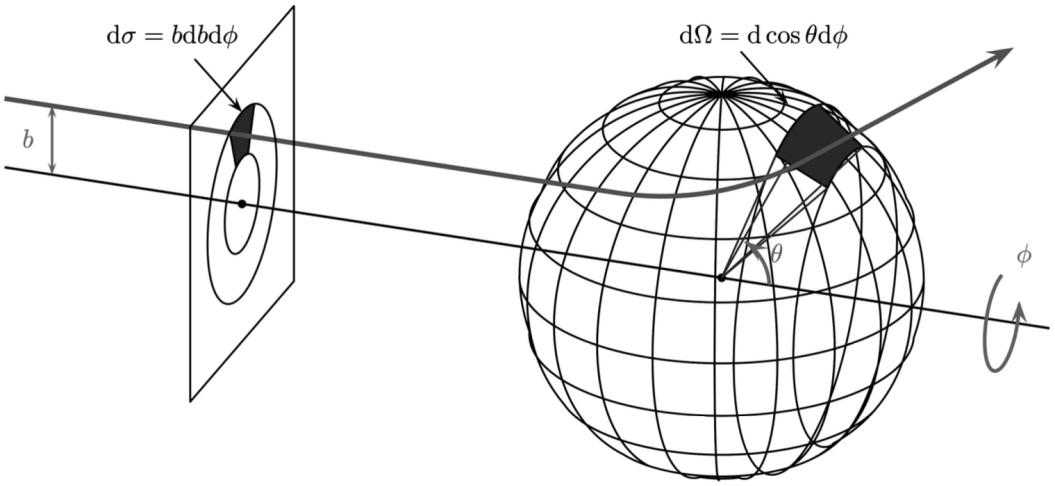


Figure 5.3: Particles incident in the area $\mathrm{d}\sigma$ scatter into the solid angle $\mathrm{d}\Omega$. (taken from Rubbia [12]).

Suppose the target is a hard sphere of radius R and the incident particle bounces off elastically. According to figure (5.2), the impact parameter can be written as $b = R\sin\alpha$, and the scattering angle as $\theta = \pi - 2\alpha$, so

$$b = R\sin\left(\frac{\pi}{2} - \frac{\theta}{2}\right) = R\cos\left(\frac{\theta}{2}\right) \tag{5.1}$$

$$\theta = \begin{cases} 2\arccos(b/R), & \text{if } b \le R \\ 0, & \text{if } b \ge R \end{cases}. \tag{5.2}$$

In a more general situation we should consider particles that come in through an infinitesimal cross-sectional area, $\mathrm{d}\sigma$, and scatter into the corresponding infinitesimal solid angle $\mathrm{d}\Omega$ (see figure (5.3)). The differential scattering cross-section, $D(\theta)$, is defined as the proportionality factor between $\mathrm{d}\sigma$ and $\mathrm{d}\Omega$.

$$\mathrm{d}\sigma = D(\theta)\mathrm{d}\Omega. \tag{5.3}$$

Although $D(\theta) = \mathrm{d}\sigma/\mathrm{d}\Omega$ is called *differential cross-section*, and is almost always written just as $\mathrm{d}\sigma/\mathrm{d}\Omega$, it is neither a cross-section nor a differential quantity. However, this terminology is universally used. In terms of the impact parameter, b, and the azimuthal angle, φ, we have

$$\mathrm{d}\sigma = b\mathrm{d}\varphi\,\mathrm{d}b \qquad \mathrm{d}\Omega = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi, \tag{5.4}$$

therefore

$$D(\theta) = \frac{b}{\sin\theta}\left|\frac{\mathrm{d}b}{\mathrm{d}\theta}\right|. \tag{5.5}$$

We take the absolute value because b is usually a decreasing function of θ, therefore the derivative is negative.[1]

In the case of hard-sphere scattering we have

$$\frac{\mathrm{d}b}{\mathrm{d}\theta} = -\frac{1}{2}R\sin(\theta/2), \qquad D(\theta) = \frac{R\cos(\theta/2)}{\sin\theta}\,\frac{1}{2}R\sin(\theta/2) = \frac{R^2}{4}. \tag{5.6}$$

The differential cross-section is actually independent of θ in this very particular case. The total cross-section is defined as

$$\sigma \equiv \int D(\theta)\,\mathrm{d}\Omega. \tag{5.7}$$

[1] Note also that $0 \le \theta \le \pi$.

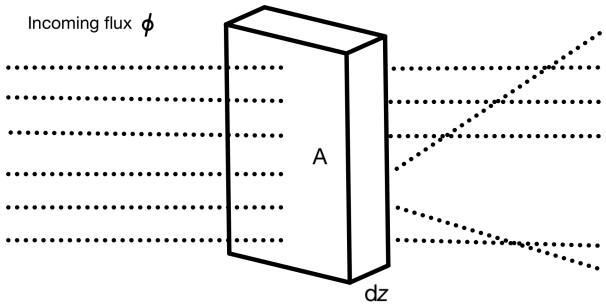


Figure 5.4: Beam colliding with a thin slab of material in fixed-target configuration.

This can be visualized as the total area of the incident beam that is scattered by the target. Alternatively, this is the area that a single particle has to cross to be scattered by the target. In our hard-sphere scattering example, the cross-section is just the geometric cross-sectional area of the sphere, as expected

$$\sigma = \frac{R^2}{4}\int \mathrm{d}\Omega = \pi R^2. \tag{5.8}$$

The flux, ϕ, of a beam of particles is defined as

$$\phi \equiv \text{number of incident particles per unit area, per unit time.} \tag{5.9}$$

The number of particles entering an area $\mathrm{d}\sigma$ per unit time is $\mathrm{d}N = \phi\,\mathrm{d}\sigma = \phi\, D(\theta)\,\mathrm{d}\Omega$, therefore we have

$$D(\theta) = \frac{1}{\phi}\frac{\mathrm{d}N}{\mathrm{d}\Omega}, \qquad \phi\int D(\theta)\,\mathrm{d}\Omega = N. \tag{5.10}$$

This formula shows, in practical terms, how a scattering cross-section is measured in the laboratory: count the number of particles scattered per unit of time into a given element of solid angle, and normalize to the flux of the incident beam.

5.2 Interaction probability.

5.2.1 Fixed-target configuration.

Let us consider a beam of particles sent to collide with a fixed target made of a thin slab of material, which has area A and very small thickness $\mathrm{d}z$, as shown in Fig. (5.4). The scattering centers are assumed to be uniformly distributed in the target slab, with an average density per unit volume ρ. We associate to each scattering center a small area given by the cross-section relevant to the physics process we wish to study, σ. If the flux of incoming beam particles is uniform, then the probability that a particle of the beam interacts with a scattering center in the target slab is equal to the ratio of the total area covered by the scattering centers and the total area of the target (assuming the incident beam illuminates all the target),

$$\text{Prob. of int. in } \mathrm{d}z \;\equiv\; w = \frac{\sigma\rho A\mathrm{d}z}{A} = \sigma\rho\mathrm{d}z. \tag{5.11}$$

This equation is valid in the so-called *thin target approximation*, which assumes the areas covered by the scattering centers do not overlap. The number of interactions per unit time, or interaction rate, can be obtained using the beam flux, ϕ, as

$$R_{\text{int}} = \phi A w = \underbrace{A\mathrm{d}z}_{V}\rho\phi\sigma = \underbrace{V\rho}_{N_{\text{c}}}\phi\sigma = N_{\text{c}}\phi\sigma, \tag{5.12}$$

where V is the total volume of the target and N_c is the total number of scattering centers in that volume. If the process under study is scattering off nuclei, then the rate can be expressed using the Avogadro number as

$$R_{\rm int} = N_A M \phi \sigma, \tag{5.13}$$

where M is the mass of the target in grams and $N_A = 6.02214076 \times 10^{23}\ \mathrm{mol}^{-1}$ (exact).

Because for thin targets the areas of the scattering centers do not overlap, the quantity N_c in the relation $R_{\rm int} = N_{\rm c}\phi\sigma$ can in fact be regarded as the number of scattering centers per unit of target area, having taken into account the actual target thickness and density. Thus, a common way to specify the properties of a thin target is just to provide the number of scatterers per unit area. This is the reason why in practical applications thin target densities are quoted in units of cm^{-2} ($\mathrm{gr}\cdot\mathrm{cm}^{-2}$, $\mathrm{nuclei}\cdot\mathrm{cm}^{-2}$, etc.).

For thick targets we define $N(z)$ as the number of particles *surviving* (i.e., that do not suffer any scattering process) after travelling a distance z inside the target. The probability of a particle interacting in any thin piece of the target between z and $z + \mathrm{d}z$ is given by w, therefore the number of particles interacting between z and $z + \mathrm{d}z$ is

$$N(z)w = N(z)\sigma\rho\mathrm{d}z. \tag{5.14}$$

The number of incident beam particles at a distance $z + \mathrm{d}z$ is equal to their number at z minus those that have interacted, thus

$$N(z + \mathrm{d}z) = N(z) - N(z)\sigma\rho\mathrm{d}z \implies \mathrm{d}N(z) = -N(z)\sigma\rho\mathrm{d}z. \tag{5.15}$$

It is convenient to define the *mean free path* of the incoming particles as the inverse of $\sigma\rho$, which has units of distance

$$\lambda \equiv \frac{1}{\sigma\rho}. \tag{5.16}$$

To find the number of surviving incident particles as a function of the distance travelled inside the target we just need to integrate

$$\int_0^z \frac{\mathrm{d}N(z')}{N(z')} = -\int_0^z \sigma\rho\,\mathrm{d}z' = -\sigma\rho\,z = -\frac{z}{\lambda}, \qquad \log\big(N(z)\big) - \log\big(N(0)\big) = -\frac{z}{\lambda}.$$

Therefore

$$N(z) = N(0)\,e^{-z/\lambda}. \tag{5.17}$$

The initial number of particles before entering the target is $N(0)$. The probability of an incident particle to not suffer an interaction after travelling a distance z inside the target decreases exponentially. The mean free path, λ, is the distance at which the fraction of surviving particles is equal to $1/e \approx 0.37$.

5.2.2 Collinear and head-on collisions.

In particle colliders we need to take into account the movement of two beams of particles that collide head-on. Let's consider a flux of incoming particles, labelled a, travelling in the positive direction of the z axis with velocity $\boldsymbol{\beta}_a$. It faces another beam of particles, labelled b, travelling in the negative direction of the same axis with velocity $\boldsymbol{\beta}_b$, as shown in Fig. (5.5). The densities of particles per unit volume are denoted as n_a and n_b, respectively. The two beams have the same cross-sectional area, A, with a perfect overlapping. When the two beams cross, a particle of beam a travels in a time interval δt a distance

$$\delta\ell = |\boldsymbol{\beta}_a - \boldsymbol{\beta}_b|\delta t = v\delta t$$

inside the beam b, where $v = |\boldsymbol{\beta}_a - \boldsymbol{\beta}_b|$ is the relative velocity between the two beams. This means that particle a crosses a volume containing

$$n_b A\delta\ell = n_b A v\delta t$$

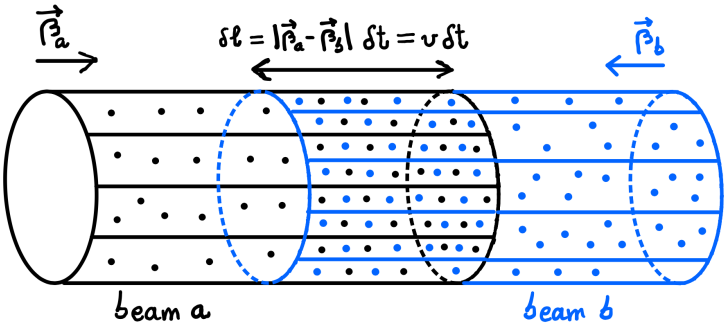


Figure 5.5: Two beam of particles colliding head-on.

particles of type b. The same reasoning used to derive Eq. (5.11) leads to the following probability of interaction of particle a

$$\text{Prob. of int. head-on} \;=\; P_{\text{int}} \;=\; \frac{n_b A v\sigma\delta t}{A} = n_b v\sigma\delta t. \tag{5.18}$$

The number of interactions per unit time (interaction rate) of $n_a V$ particles of type a in a volume V is then given by

$$R_{\text{int}} = \frac{n_a V P_{\text{int}}}{\delta t} = n_a n_b v V \sigma. \tag{5.19}$$

There are $n_b V$ particle of type b in the volume V, therefore the interaction rate *per particle of type* b is given by

$$R_{\text{int,b}} = \frac{R_{\text{int}}}{n_b V} = n_a v\sigma. \tag{5.20}$$

The quantity $\phi_a \equiv n_a v$ can be taken to be flux (units $\text{cm}^{-2}\text{s}^{-1}$) of particles a seen by a particle of type b. Hence $R_{\text{int,b}}$ can be written as

$$R_{\text{int,b}} = \phi_a \sigma. \tag{5.21}$$

The total rate, R_{int}, can be written in terms of a total relative flux factor F defined as[2]

$$F = n_a n_b v. \tag{5.22}$$

Thus

$$R_{\text{int}} = VF\sigma \qquad \text{(head-on collisions)}. \tag{5.23}$$

It should be noted that this derivation is only valid in a non-relativistic regime, which is seldom the case in practical applications of particle physics colliders.

5.3 Quantum scattering theory.

To describe the scattering of a particle by a potential using quantum mechanics we should in principle study the time evolution of a wave packet, which presumably represents an incident particle with some energy width and spatial localization. However, most of the important physics results can be more easily obtained using plane waves. This is rather satisfactory, as long as the dimension of the wave packet is much larger than the size of the scatterer or the range of the potential. Let's consider then an incident plane wave traveling in the z direction, $\psi(z) = Ae^{ikz}$, that scatters off a potential. This is a non-normalizable state, infinite in extent in both space and time, but we can also interpret the real parameter $|A|^2$ as the number of incident beam particles per unit volume. After the scattering has taken place, we expect to have a superposition of the initial plane wave and a scattered spherical wave, as shown schematically in this two-dimensional graph

[2] This relative flux, despite its name, does not have units of $\text{cm}^{-2}\text{s}^{-1}$.

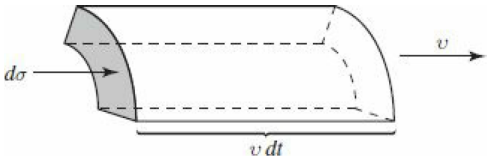
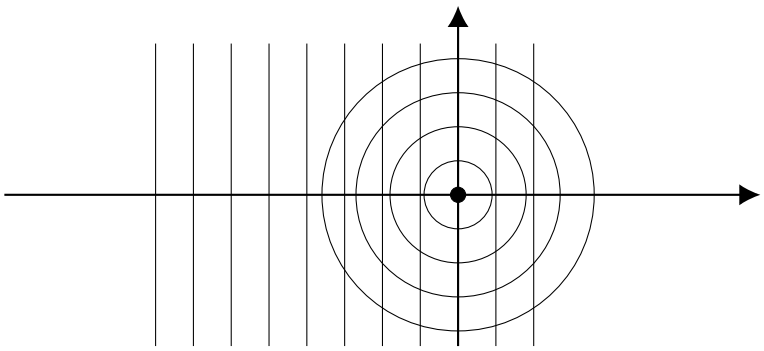


Figure 5.6: Graphical representation of the volume $\mathrm{d}V = v\mathrm{d}t\,\mathrm{d}\sigma$. (taken from Griffiths [7]).



On physical grounds, then, the solutions of the Schrödinger equation for large values of r ought to be of the general form

$$\psi(r,\theta) \approx A\left[e^{ikz} + f(\theta)\frac{e^{ikr}}{r}\right]. \tag{5.24}$$

Note that the spherical wave carries a factor of $1/r$ because, at large values of r, $|\psi|^2$ must go as $1/r^2$ to conserve probability.[3] The wave number k is related to the energy of the incident particles through

$$k = \frac{\sqrt{2mE}}{\hbar}. \tag{5.25}$$

The function $f(\theta)$ is called the *scattering amplitude*, and for the time being we will assume it depends only on the polar angle θ, although in general it could have also an azimuthal dependence on φ. It is not difficult to see that $|f(\theta)|^2$ is in fact the probability of scattering in a given direction θ. Indeed, looking at figure (5.6), we realize that the number of particles (traveling at speed v) that cross the infinitesimal area $\mathrm{d}\sigma$ in a time $\mathrm{d}t$ is given by the product of the differential volume $\mathrm{d}V = v\mathrm{d}t\,\mathrm{d}\sigma$ and the modulus squared of the incident plane wave,

$$|\psi_{\text{incident}}|^2\,\mathrm{d}V = |A|^2(v\mathrm{d}t)\,\mathrm{d}\sigma. \tag{5.26}$$

By probability (or flux) conservation, this must be equal to the number of particles that scatter into the corresponding solid angle $\mathrm{d}\Omega$, which is given by the product of the spherical wave amplitude and the differential element $r^2\,\mathrm{d}\Omega$

$$|\psi_{\text{scattered}}|^2\,\mathrm{d}V = \frac{|A|^2|f|^2}{r^2}(v\mathrm{d}t)\,r^2\,\mathrm{d}\Omega = |A|^2|f|^2(v\mathrm{d}t)\,\mathrm{d}\Omega. \tag{5.27}$$

Equating both expressions we find $\mathrm{d}\sigma = |f|^2\mathrm{d}\Omega$, therefore

$$D(\theta) = \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f(\theta)|^2. \tag{5.28}$$

[3]This decreasing as $1/r$ of the scattered wave was at the origin of the probabilistic interpretation made by Max Born in 1926. At first, Schrödinger and others thought that the wave functions represent particles that are spread out, like pressure disturbances in a fluid, and that *most of the particle* is located where the wave function is large. However, the $1/r$ dependence contradicts the common experience that, though a particle striking a target may be scattered in any direction, it does not break up and go in *all directions*. Born concluded that $\psi(\mathbf{r},t)$ does not tell us how much of the particle is at position $\mathbf{r}$ at time t, but rather the probability that the particle is at or near $\mathbf{r}$ at time t.

The main goal of scattering theory is to compute the scattering amplitude $f(\theta)$ by solving the appropriate Schrödinger equation. In the remainder of this chapter we will study two techniques to attack this problem: partial wave analysis and the Born approximation.

5.4 Partial wave analysis.

Partial wave analysis is especially convenient for spherically symmetric potentials, $V(\mathbf{r}) = V(r)$. It will allow us to examine how states with definite angular momentum are affected by the scatterer. For spherical symmetric potentials, we have already seen in section 1.4 that the solutions of the Schrödinger equation can be written as a product of a radial function and a spherical harmonic,

$$\psi(r, \theta, \varphi) = R(r)\, Y_\ell^m(\theta, \varphi), \tag{5.29}$$

where the function $u(r) = rR(r)$ satisfies the radial equation

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2 u}{\mathrm{d}r^2} + \left[V(r) + \frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}\right] u = Eu, \tag{5.30}$$

which we have already found before as equation (1.98). Here we will only consider solutions with positive energy ($E > 0$), because negative-energy solutions correspond to bound states. At very large r we assume both $V(r)$ and the centrifugal contribution (which goes as $1/r^2$) to be negligible, so the previous equation becomes

$$\frac{\mathrm{d}^2 u}{\mathrm{d}r^2} \approx -k^2 u, \tag{5.31}$$

where $k = \sqrt{2mE}/\hbar$ (note that $E > 0$). Its general solution can be written as

$$u(r) = Ce^{ikr} + De^{-ikr}. \tag{5.32}$$

The first term represents an outgoing spherical wave, and the second represents an incoming one. We want to describe the scattered wave, so we must impose the condition $D = 0$. At very large values of r we have then

$$R(r) = \frac{1}{r}u(r) \sim \frac{e^{ikr}}{r}. \tag{5.33}$$

This was already hinted at before on physical grounds (equation 5.24), and holds for $kr \gg 1$.

To get more information on $R(r)$ we can examine the *intermediate* region, where the centrifugal term must be taken into account but the potential $V(r)$ can still be neglected.[4] In such a case we have

$$\frac{\mathrm{d}^2 u}{\mathrm{d}r^2} - \frac{\ell(\ell+1)}{r^2}u = -k^2 u. \tag{5.34}$$

The general solution to this equation, which you have probably seen when studying the infinite spherical well in three dimensions, is a linear combination of spherical Bessel functions of the first (j_ℓ) and second (or Neumann) class (n_ℓ), of order ℓ,

$$R(r) = \frac{1}{r}u(r) = Aj_\ell(kr) + Bn_\ell(kr), \qquad u(r) = Arj_\ell(kr) + Brn_\ell(kr). \tag{5.35}$$

These special functions have the following explicit form (z can be a complex number here)

$$j_\ell(z) = z^\ell \left(-\frac{1}{z}\frac{\mathrm{d}}{\mathrm{d}z}\right)^\ell \frac{\sin z}{z}, \qquad n_\ell(z) = -z^\ell \left(-\frac{1}{z}\frac{\mathrm{d}}{\mathrm{d}z}\right)^\ell \frac{\cos z}{z}. \tag{5.36}$$

[4] The Coulomb potential does not satisfy this condition, since $1/r$ goes to zero more slowly than the centrifugal term when $r \to \infty$. Partial wave analysis is not applicable to the Coulomb potential, which is not *localized enough*.

Their asymptotic behaviour for large $|z|$ (in our case for large kr) is

$$j_\ell(kr) \approx \frac{1}{kr}\sin\left(kr-\frac{\ell\pi}{2}\right)+\mathcal{O}\left(\frac{1}{(kr)^2}\right), \qquad n_\ell(kr) \approx -\frac{1}{kr}\cos\left(kr-\frac{\ell\pi}{2}\right)+\mathcal{O}\left(\frac{1}{(kr)^2}\right),$$

or, in exponential form,

$$j_\ell(kr) \approx \frac{1}{2ikr}\left[e^{i\left(kr-\frac{\ell\pi}{2}\right)}-e^{-i\left(kr-\frac{\ell\pi}{2}\right)}\right], \qquad n_\ell(kr) \approx -\frac{1}{2kr}\left[e^{i\left(kr+\frac{\ell\pi}{2}\right)}+e^{-i\left(kr-\frac{\ell\pi}{2}\right)}\right].$$

It is clear that neither $j_\ell(kr)$ nor $n_\ell(kr)$ can asymptotically represent only an outgoing or incoming spherical wave. For that purpose we need to choose appropriate linear combinations of them, in a similar way that the spherical functions e^{ikr} and e^{-ikr} are built from linear combinations of sine and cosine functions. The appropriate linear combinations are known as *spherical Hankel functions*,[5]

$$h_\ell^{(1)}(z) = j_\ell(z) + i\, n_\ell(z), \qquad h_\ell^{(2)}(z) = j_\ell(z) - i\, n_\ell(z). \tag{5.37}$$

As a reference, the first two spherical Hankel functions of the first and second type are

$$h_0^{(1)}(z) = -i\frac{e^{iz}}{z}, \qquad h_0^{(2)}(z) = i\frac{e^{-iz}}{z}, \tag{5.38}$$

$$h_1^{(1)}(z) = \left(-\frac{i}{z^2}-\frac{1}{z}\right)e^{iz}, \qquad h_1^{(2)}(z) = \left(\frac{i}{z^2}-\frac{1}{z}\right)e^{-iz}. \tag{5.39}$$

The asymptotic behaviour of the spherical Hankel functions for large $|z|$ is[6]

$$h_\ell^{(1)} \longrightarrow \frac{1}{z}(-i)^{\ell+1}e^{iz}, \qquad h_\ell^{(2)} \longrightarrow \frac{1}{z}(i)^{\ell+1}e^{-iz}, \qquad (\text{for } |z| \gg 1). \tag{5.40}$$

According to this, it is clear that to describe outgoing waves we only need spherical Hankel functions of the first kind, so the radial part of the wave function has to go as

$$R(r) \sim h_\ell^{(1)}(kr). \tag{5.41}$$

The *exact* solution to the Schrödinger equation outside the scattering region, where $V(r) \approx 0$ but $\ell(\ell+1)/r^2$ is not negligible, can then be written as

$$\psi(r,\theta,\varphi) = A\left[e^{ikz} + \sum_{\ell,m} C_{\ell,m}\, h_\ell^{(1)}(kr) Y_\ell^m(\theta,\varphi)\right]. \tag{5.42}$$

The first term, Ae^{ikz}, is the incident plane wave, and the second term represents the scattered wave. Since we are assuming the potential to be spherically symmetric, the wave function cannot depend on φ, which means that only terms with $m = 0$ are present, because $Y_\ell^m \sim e^{im\varphi}$. The spherical harmonics with $m = 0$ are

$$Y_\ell^0(\theta,\varphi) = \sqrt{\frac{2\ell+1}{4\pi}}\, P_\ell(\cos\theta), \tag{5.43}$$

where P_ℓ is the ℓ-th Legendre polynomial. Instead of the $C_{\ell,0}$ coefficient it is customary to introduce a new parameter a_ℓ, called the ℓ-th *partial wave amplitude*, such that[7]

$$C_{\ell,0} \equiv i^{\ell+1}k\sqrt{4\pi(2\ell+1)}\, a_\ell. \tag{5.44}$$

[5] The spherical Hankel functions, or spherical Bessel functions of the third class, are themselves also a pair of linearly independent solutions of the Bessel equation.

[6] We use the *modulus* of z, $|z|$, because in general the argument of these special functions is a complex number, but we will be using them only with real arguments.

[7] The parameter a_ℓ is also called *partial scattering amplitude*. If you happen to be reading also a collection of lecture notes on nuclear and particle physics written some years ago for courses at the USC by Juan J. Saborido, note that the a_ℓ used there has a different meaning. Calling a_ℓ^{old} the parameter used on those old lecture notes, the relation to the one introduced here is $a_\ell^{\text{old}} = 1 + 2ika_\ell$.

With this notation the wave function can be written as

$$\psi(r,\theta,\varphi) = A\left[e^{ikz} + k\sum_{\ell=0}^{\infty} i^{\ell+1}(2\ell+1)\,a_\ell\,h_\ell^{(1)}(kr)\,P_\ell(\cos\theta)\right]. \tag{5.45}$$

For very large r the Hankel function goes like

$$h_\ell^{(1)}(kr) \;\longrightarrow\; \frac{1}{kr}(-i)^{\ell+1}e^{ikr}, \tag{5.46}$$

therefore we have

$$\psi(r,\theta,\varphi) \approx A\left[e^{ikz} + f(\theta)\frac{e^{ikr}}{r}\right], \tag{5.47}$$

where

$$f(\theta) = \sum_{\ell=0}^{\infty}(2\ell+1)\,a_\ell\,P_\ell(\cos\theta). \tag{5.48}$$

We have arrived again at equation (5.24) following a more rigorous route. Nevertheless, it should be noted that this equation is only valid for a potential $V(r)$ that verifies the condition $rV(r) \to 0$ when $r \to \infty$ (it has a shorter range than the Coulomb potential).

Since this is the asymptotic behaviour of the solution for $r \to \infty$, it may seem surprising that the plane-wave term is not depleted by the scattering potential. Actually, in the forward direction there is an interference between the two terms in equation (5.47), which does decrease the amplitude of the plane wave beyond the scattering center, as required by the conservation of probability. This implies that there must be a relation between the forward scattering amplitude, $f(\theta = 0)$, and the total scattering cross section. This relation is known as the *optical theorem*. A simplified form of this theorem will be proved in section 5.6 for central potentials (equation 5.95).

To make more explicit the physical meaning of the partial wave amplitude, a_ℓ, let us write the plane wave in terms of spherical waves (Rayleigh's formula)

$$e^{ikz} = \sum_{\ell=0}^{\infty} i^\ell(2\ell+1)\,j_\ell(kr)\,P_\ell(\cos\theta). \tag{5.49}$$

For large kr this is (remember that $i^\ell = e^{i(\pi/2)\ell}$)

$$\begin{aligned} e^{ikz} &\approx \sum_{\ell=0}^{\infty} i^\ell(2\ell+1)\,\frac{1}{2ikr}\left[e^{i\left(kr-\frac{\ell\pi}{2}\right)} - e^{-i\left(kr-\frac{\ell\pi}{2}\right)}\right]P_\ell(\cos\theta) \\ &= \sum_{\ell=0}^{\infty}\frac{(2\ell+1)}{2ikr}\left[e^{ikr} - e^{-i(kr-\ell\pi)}\right]P_\ell(\cos\theta). \end{aligned} \tag{5.50}$$

Therefore (also for large kr)

$$\begin{aligned} \psi(r,\theta,\varphi) &\approx A\left[e^{ikz} + f(\theta)\frac{e^{ikr}}{r}\right] \\ &\approx A\left[\sum_\ell(2\ell+1)P_\ell(\cos\theta)\left(\frac{e^{ikr} - e^{-i(kr-\ell\pi)}}{2ikr}\right) + \sum_\ell(2\ell+1)\,a_\ell\,P_\ell(\cos\theta)\frac{e^{ikr}}{r}\right] \\ &= A\sum_\ell\frac{(2\ell+1)P_\ell(\cos\theta)}{2ik}\left[(1+2ika_\ell)\frac{e^{ikr}}{r} - \frac{e^{-i(kr-\ell\pi)}}{r}\right]. \end{aligned} \tag{5.51}$$

Comparing equations (5.50) and (5.51) we clearly see now the physical meaning of a_ℓ. When the scattering potential is absent we can analyze the plane wave as the sum of a spherically

outgoing wave behaving like e^{ikr}/r and a spherical incoming wave behaving like $-e^{-i(kr-\ell\pi)}/r$ for each ℓ. The presence of the scattering potential changes only the coefficient of the outgoing wave as

$$1 \longrightarrow (1 + 2ika_\ell), \tag{5.52}$$

while the incoming wave is completely unaffected.

It is possible to derive a simple expression for the total cross-section in terms of the partial wave amplitudes. The Legendre polynomials satisfy the following orthogonality relation,

$$\int_{-1}^{+1} P_\ell(x)\, P_{\ell'}(x)\, \mathrm{d}x \;=\; \left(\frac{2}{2\ell+1}\right) \delta_{\ell\ell'}. \tag{5.53}$$

Therefore

$$\begin{aligned}\sigma = \int |f(\theta)|^2\, \mathrm{d}\Omega &= \int \sum_{\ell=0}^{\infty}\sum_{\ell'=0}^{\infty}(2\ell+1)(2\ell'+1)\, a_\ell a_{\ell'}^*\, P_\ell(\cos\theta)\, P_{\ell'}(\cos\theta)\, \mathrm{d}\Omega \\ &= 2\pi \sum_{\ell=0}^{\infty}\sum_{\ell'=0}^{\infty} a_\ell a_{\ell'}^* \int (2\ell+1)(2\ell'+1)\, P_\ell(\cos\theta)\, P_{\ell'}(\cos\theta)\, \mathrm{d}(\cos\theta). \end{aligned} \tag{5.54}$$

And finally

$$\sigma = 4\pi \sum_{\ell=0}^{\infty}(2\ell+1)\, |a_\ell|^2. \tag{5.55}$$

Our problem has been reduced then to compute the partial wave amplitudes, a_ℓ. This can be done by solving the Schrödinger equation in the small r region, where $V(r) \neq 0$, and matching the solution to the one obtained for intermediate r (equation (5.45)) using appropriate boundary conditions.

The ortoghonality condition of the Legendre Polynomials can also be used to express directly the partial wave amplitudes in terms the scattering amplitude. Indeed, multiply $f(\theta)$ by $P_\ell(\cos\theta)\sin\theta$ and integrate over θ to get, using equation (5.48),

$$\begin{aligned}\int_0^\pi f(\theta) P_{\ell'}(\cos\theta)\sin\theta\, \mathrm{d}\theta &= \sum_{\ell=0}^{\infty}\int_0^\pi (2\ell+1) a_\ell P_{\ell'}(\cos\theta) P_\ell(\cos\theta)\sin\theta\, \mathrm{d}\theta \\ &= \sum_{\ell=0}^{\infty}\int_{-1}^{1}(2\ell+1) a_\ell P_{\ell'}(x) P_\ell(x)\, \mathrm{d}x = \sum_{\ell=0}^{\infty}(2\ell+1) a_\ell \frac{2\delta_{\ell\ell'}}{2\ell+1} = 2a_{\ell'}. \end{aligned} \tag{5.56}$$

Therefore

$$a_\ell = \frac{1}{2}\int_0^\pi f(\theta) P_\ell(\cos\theta)\, \sin\theta\, \mathrm{d}\theta. \tag{5.57}$$

The solution for the wave function in equation (5.45) is written in a somewhat *hybrid* manner. The incident plane wave contains implicitly contributions from states with all possible values of the angular momentum ℓ, while the outgoing spherical wave shows explicitly the contributions of all partial waves. We can get a more useful and homogeneous expression, only in terms of r and θ, using the Rayleigh's formula (5.49) for the plane wave. The result is

$$\psi(r,\theta,\varphi) = A\sum_{\ell=0}^{\infty} i^\ell (2\ell+1)\, \left[j_\ell(kr) + ik\, a_\ell\, h_\ell^{(1)}(kr)\right]\, P_\ell(\cos\theta). \tag{5.58}$$

We will need at some point the asymptotic behaviour of this wave function for large kr. This can be easily worked out using the asymptotic behaviour of the spherical functions

$$j_\ell(z) = \frac{1}{2}\left[h_\ell^{(1)}(z) + h_\ell^{(2)}(z)\right] \approx \frac{1}{2z}\left[(-i)^{\ell+1}e^{iz} + (i)^{\ell+1}e^{-iz}\right], \qquad \text{for } |z| \gg 1. \tag{5.59}$$

The result is

$$\psi(r,\theta,\varphi) \approx \frac{A}{2ikr}\sum_{\ell=0}^{\infty}(2\ell+1)\left[(1+2ik\, a_\ell)\, e^{ikr} - (-1)^\ell e^{-ikr}\right] P_\ell(\cos\theta), \qquad \text{for } kr \gg 1. \tag{5.60}$$

5.4.1 Quantum hard-sphere scattering.

Let's apply the foregoing discussion to the hard-sphere scattering problem, where the potential is taken to be

$$V(r) = \begin{cases} \infty, & \text{if } r \le R \\ 0, & \text{if } r > R \end{cases} \tag{5.61}$$

Here the boundary condition forces the wave function to vanish everywhere inside the sphere of radius R. In particular,

$$\psi(R, \theta) = 0, \tag{5.62}$$

for all values of θ. It then follows that

$$\sum_{\ell=0}^{\infty} i^\ell (2\ell+1) \left[j_\ell(kR) + ik\, a_\ell\, h_\ell^{(1)}(kR) \right] P_\ell(\cos\theta) = 0. \tag{5.63}$$

Multiply now this equation by $P_{\ell'}(\cos\theta)\, \sin\theta\, \mathrm{d}\theta$, integrate from 0 to π and use the orthogonality relation (5.53) to get

$$\begin{aligned} \sum_{\ell=0}^{\infty} i^\ell (2\ell+1) \left[j_\ell(kR) + ik\, a_\ell\, h_\ell^{(1)}(kR) \right] \int_0^\pi P_\ell(\cos\theta)\, P_{\ell'}(\cos\theta)\, \sin\theta\, \mathrm{d}\theta \\ = \sum_{\ell=0}^{\infty} i^\ell (2\ell+1) \left[j_\ell(kR) + ik\, a_\ell\, h_\ell^{(1)}(kR) \right] \left(\frac{2}{2\ell+1} \right) \delta_{\ell\ell'} \\ = 2i^\ell \left[j_\ell(kR) + ik\, a_\ell\, h_\ell^{(1)}(kR) \right] = 0. \end{aligned} \tag{5.64}$$

From this we conclude that

$$a_\ell = i\, \frac{j_\ell(kR)}{k\, h_\ell^{(1)}(kR)}. \tag{5.65}$$

And the total cross-section is

$$\sigma = 4\pi \sum_{\ell=0}^{\infty} (2\ell+1)\, |a_\ell|^2 = \frac{4\pi}{k^2} \sum_{\ell=0}^{\infty} (2\ell+1) \left| \frac{j_\ell(kR)}{h_\ell^{(1)}(kR)} \right|^2. \tag{5.66}$$

It is instructive to examine this *exact* answer for some limiting cases. For low energy scattering where $kR \ll 1$, which means that the wavelength is much larger that the sphere radius (remember that $k = 2\pi/\lambda$), we need the asymptotic behaviour of the spherical Bessel and Neumann functions,

$$j_\ell \to \frac{2^\ell\, \ell!}{(2\ell+1)!}\, z^\ell, \qquad n_\ell \to -\frac{(2\ell)!}{2^\ell\, \ell!}\, \frac{1}{z^{\ell+1}}, \qquad \text{for } \; z \ll 1. \tag{5.67}$$

For small z, $n_\ell(z)$ is much larger than $j_\ell(z)$, so

$$\begin{aligned} \frac{j_\ell(z)}{h_\ell^{(1)}(z)} = \frac{j_\ell(z)}{j_\ell(z) + i\, n_\ell(z)} \approx -i \frac{j_\ell(z)}{n_\ell(z)} \\ \approx i \frac{2^\ell\, \ell!}{(2\ell+1)!}\, z^\ell\, \frac{2^\ell\, \ell!}{(2\ell)!}\, \frac{z^{\ell+1}}{1} = \frac{i}{(2\ell+1)} \left(\frac{2^\ell\, \ell!}{(2\ell)!} \right)^2 z^{2\ell+1}. \end{aligned} \tag{5.68}$$

Therefore

$$\sigma \approx \frac{4\pi}{k^2} \sum_{\ell=0}^{\infty} \frac{1}{(2\ell+1)} \left(\frac{2^\ell\, \ell!}{(2\ell)!} \right)^4 (kR)^{4\ell+2}. \tag{5.69}$$

For $kR \ll 1$ we can afford to neglect higher powers of kR. This tells us that in the low energy approximation the scattering is dominated by the $\ell = 0$ term and the differential cross-section is independent of θ, as in the classical case. Retaining only $\ell = 0$ we find

$$\sigma \approx 4\pi R^2. \tag{5.70}$$

Remember that for the classical hard-sphere scattering problem we obtained a total cross-section which is equal to the geometrical cross-section of the sphere (equation 5.8), whereas in the quantum domain the result is four times larger. This is a characteristic result of long-wavelength scattering, and arises in optics as well.

One might think that going to the high enery limit case we should get the geometric cross section. Let's see if this is true. For $kR \gg 1$ the asymptotic behaviour of the Bessel and Neumann spherical functions is

$$j_\ell(kR) \approx \frac{1}{kR} \sin\left(kR - \frac{\ell\pi}{2}\right), \qquad \text{and} \qquad n_\ell(kR) \approx -\frac{1}{kR} \cos\left(kR - \frac{\ell\pi}{2}\right).$$

Therefore

$$\begin{aligned} |a_\ell|^2 = \left| i \frac{j_\ell(kR)}{k\, h_\ell^{(1)}(kR)} \right|^2 &= \frac{1}{k^2} \left| \frac{j_\ell(kR)}{j_\ell(kR) + i n_\ell(kR)} \right|^2 \\ &\approx \frac{1}{k^2} \left| \frac{\sin\left(kR - \frac{\ell\pi}{2}\right)}{\sin\left(kR - \frac{\ell\pi}{2}\right) - i \cos\left(kR - \frac{\ell\pi}{2}\right)} \right|^2 = \frac{1}{k^2} \sin^2\left(kR - \frac{\ell\pi}{2}\right). \end{aligned} \tag{5.71}$$

At high energies, many values of ℓ contribute to the cross-section, but there is a maximum value above which the cross-section vanishes. This can be estimated to be $\ell_{\max} \approx kR$, since $\hbar kR$ is the magnitude of the angular momentum of the incident particle with respect to the scattering center when it is coming with an impact parameter of $b \approx R$, while for $b > R$ the particle misses the target. The cross-section is then

$$\sigma \approx \frac{4\pi}{k^2} \sum_{\ell=0}^{kR} (2\ell + 1) \sin^2\left(kR - \frac{\ell\pi}{2}\right) = \frac{4\pi}{k^2} \sum_{\ell=0}^{kR} (2\ell + 1) \sin^2 \delta_\ell. \tag{5.72}$$

The parameter $\delta_\ell = (kR - \ell\pi/2)$ is called a *phase shift*. The physical meaning of δ_ℓ is discussed in the next section, but for the time being take it just as a shorthand for $(kR - \ell\pi/2)$. Note that δ_ℓ decreases by $\pi/2$ every time ℓ increases by one unit, therefore for an adjacent pair of partial waves we have

$$\sin^2 \delta_\ell + \sin^2 \delta_{\ell+1} = \sin^2 \delta_\ell + \sin^2\left(\delta_\ell - \frac{\pi}{2}\right) = \sin^2 \delta_\ell + \cos^2 \delta_\ell = 1. \tag{5.73}$$

$$\sum_{\ell=0}^{kR} \sin^2 \delta_\ell = \sum_{\ell=0}^{\frac{kR}{2}-1} \left(\sin^2 \delta_{2\ell} + \sin^2 \delta_{2\ell+1}\right) = \frac{kR}{2}. \tag{5.74}$$

Also, since there are so many values of ℓ contributing to the cross-section, it is permissible to replace $\sin^2 \delta_\ell$ by its average value $(1/2)$ in the following sum[8]

$$\sum_{\ell=0}^{kR} 2\ell \sin^2 \delta_\ell \approx \sum_{\ell=0}^{kR} \ell = \frac{1}{2} kR(kR+1) = \frac{(kR)^2}{2} + \frac{kR}{2}. \tag{5.75}$$

[8] In the limit that a is very large we have

$$\frac{1}{a} \int_0^a \sin^2 x \, \mathrm{d}x = \frac{1}{a} \left(\frac{a}{2} - \frac{1}{2} \sin a \cos a\right) \longrightarrow \frac{1}{2}.$$

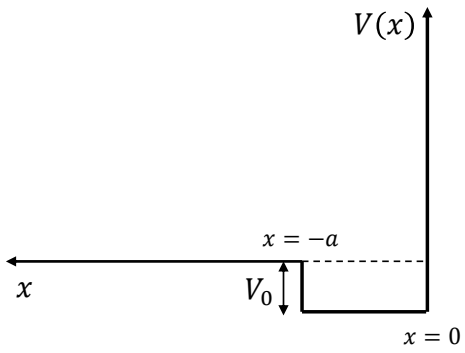


Figure 5.7: One-dimensional scattering from a localized potential bounded by an infinite wall.

Finally, neglecting terms linear in kR,

$$\sigma \approx \frac{4\pi}{k^2}\sum_{\ell=0}^{kR}(2\ell+1)\,\sin^2\delta_\ell \approx \frac{4\pi}{k^2}\left[\frac{(kR)^2}{2}+kR\right] \approx 2\pi R^2. \tag{5.76}$$

This is twice the geometric cross-section, which looks a bit surprising. One might have expected that high-energy spheres in the center-of-mass frame experience some reaction only if they approach each other along lines separated by no more than a distance R, the range of their interaction. In this case, the asymptotic value of the total cross section would be πR^2, not $2\pi R^2$. However, this larger cross-section can be attributed to elastic scattering due to the diffraction that takes place when two particles approach each other at distances a little larger than R. If you think twice, this is not so surprising after all, since we know that objects in the quantum domain exhibit a wavelike behaviour and a sort of Fraunhofer diffraction has to occur because each sphere is an *obstacle* for the other.[9]

5.5 Phase shifts.

Consider a one-dimensional problem with the potential

$$V(x) = \begin{cases} 0, & \text{if } x < 0 \\ \infty, & \text{if } x \geq 0 \end{cases} \tag{5.77}$$

Conservation of probability tells us that a wave incident from the left, $\psi_i(x) = Ae^{ikx}$, will be entirely reflected into $\psi_r(x) = Be^{-ikx} = -Ae^{-ikx}$. The condition $B = -A = e^{i\pi}A$ ensures hat the total wave function (incident plus reflected) vanishes at the origin (as it should, because there is a wall of infinite potential),

$$\psi_0(x=0) \;=\; A\left(e^{ikx} - e^{-ikx}\right) = 0. \tag{5.78}$$

Suppose now that there is a small localized region with a non-zero potential between $x = -a$ and $x = 0$, such as (see figure 5.7)

$$V(x) = \begin{cases} 0, & \text{if } x \leq -a \\ -V_0, & \text{if } -a < x < 0 \\ \infty, & \text{if } x \geq 0 \end{cases} \tag{5.79}$$

[9] Fraunhofer diffraction occurs when light waves are directed to a small hole in a piece of paper, or just to a small obstacle.

Conservation of probability again forces the amplitude of the reflected wave for $x < -a$ to be the same as that of the incident wave ($|B| = |A|$), but the phase-change will depend on the exact form of the little potential well, so in general we should write $B = Ae^{i2\delta}$. Indeed, changing the phase of the wave is the only thing the potential can do in elastic scattering. The total wave function for $x < -a$ takes now the form

$$\psi(x) \;=\; A\left(e^{ikx} - e^{-i(2\delta - kx)}\right). \tag{5.80}$$

The δ parameter is called *phase shift*, and its determination for all relevant partial waves completely solves the scattering problem. This is done by solving the Schrödinger equation in the scattering region ($-a < x < 0$) and imposing suitable boundary conditions. The factor of 2 in $B = Ae^{i2\delta}$ is purely conventional. We may think of the incident wave as being phase shifted once on the way in, and again in the way out, so that δ would be a *one-way* phase shift and 2δ the total phase shift.

In the three-dimensional case the incident plane wave Ae^{ikz}, although it carries no angular momentum in the z direction (there are no terms with $m \neq 0$ in the Rayleigh's formula), includes all values of the *total* angular momentum ($\ell = 0, 1, 2, \ldots$). In a spherically symmetric potential the angular momentum of the scattered particle does not change, therefore each partial wave (labelled by a particular ℓ) scatters independently with no change in amplitude (only in phase). If there were no potential then we would only have the original plane wave function, $\psi_0 = Ae^{ikz}$, and according to equation (5.49) the ℓ-th partial wave would be

$$\psi_0^{(\ell)} = Ai^{\ell}(2\ell + 1)\, j_\ell(kr)\, P_\ell(\cos\theta). \tag{5.81}$$

Let's find the asymptotic behaviour of $\psi_0^{(\ell)}$ for large kr. From equations (5.37) and (5.40) we see that

$$j_\ell(z) = \frac{1}{2}\left[h_\ell^{(1)}(z) + h_\ell^{(2)}(z)\right] \approx \frac{1}{2z}\left[(-i)^{\ell+1}e^{iz} + (i)^{\ell+1}e^{-iz}\right], \qquad \text{for } |z| \gg 1. \tag{5.82}$$

Therefore for large kr and *no potential* we would have

$$\psi_0^{(\ell)} \approx A\frac{(2\ell+1)}{2ikr}\left[e^{ikr} - (-1)^{\ell}e^{-ikr}\right] P_\ell(\cos\theta), \qquad (V(r) = 0). \tag{5.83}$$

Now we should think about how the addition of a scattering potential modifies this expression. The second term in the square brackets, which goes as $(-1)^{\ell}e^{-ikr}$, represents an incoming spherical wave and cannot be affected by the presence of a potential. The first term is an outgoing spherical wave. This is the term that gets changed by the scattering potential, which introduces a phase shift δ_ℓ. Therefore we should write

$$\psi^{(\ell)} \approx A\frac{(2\ell+1)}{2ikr}\left[e^{i(kr+2\delta_\ell)} - (-1)^{\ell}e^{-ikr}\right] P_\ell(\cos\theta), \qquad (V(r) \neq 0). \tag{5.84}$$

To find a relation between the partial wave amplitudes, a_ℓ, and the phase shifts, δ_ℓ, we need to compare this last equation with the asymptotic behaviour of equation (5.60) for large kr. Paying attention only to the ℓ-wave, this is

$$\psi^{(\ell)} \approx \frac{A(2\ell+1)}{2ikr}\left[(1 + 2ika_\ell)e^{ikr} - (-1)^{\ell}e^{-ikr}\right] P_\ell(\cos\theta). \tag{5.85}$$

By comparing the coefficients that go with the e^{ikr} term we see that $e^{i2\delta_\ell} = 1 + 2ika_\ell$, or

$$a_\ell = \frac{1}{2ik}\left(e^{i2\delta_\ell} - 1\right) = \frac{1}{k}e^{i\delta_\ell}\sin\delta_\ell = \frac{1}{k\cot\delta_\ell - ik}. \tag{5.86}$$

Alternatively, we could have arrived at this by using directly (5.52) and making the identification $(1 + 2ika_\ell) = e^{i2\delta_\ell}$. Note that both a_ℓ and δ_ℓ are just parameters[10] and the relation obtained

[10] They have an energy dependency (functions of k) but are independent of r.

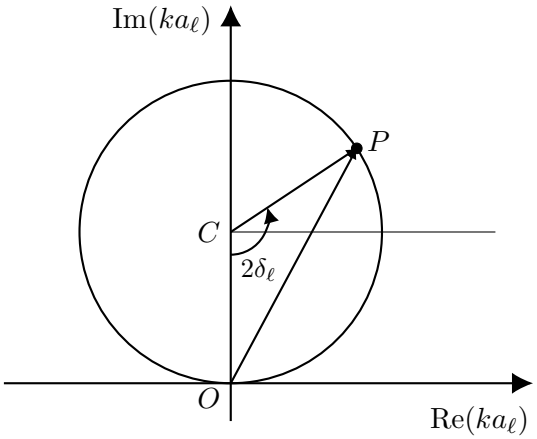


Figure 5.8: Argand diagram for ka_ℓ. The length of the segment OP is the magnitude of ka_ℓ. The center of the circle is at $C = (0, 1/2)$. Figure adapted from Sakurai and Napolitano [13].

between them is exact even though we have used the asymptotic form of the wave functions to derive it. Inserting this in equation (5.48), the scattering amplitude becomes

$$f(\theta) = \frac{1}{k}\sum_{\ell=0}^{\infty}(2\ell+1)\, e^{i\delta_\ell}\sin(\delta_\ell)\, P_\ell(\cos\theta), \tag{5.87}$$

which can also be written as

$$f(\theta) = \frac{1}{2ik}\sum_{\ell=0}^{\infty}(2\ell+1)\,(e^{2i\delta_\ell}-1)\, P_\ell(\cos\theta). \tag{5.88}$$

The total cross-section is

$$\sigma = 4\pi\sum_{\ell=0}^{\infty}(2\ell+1)\,|a_\ell|^2 = \frac{4\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell+1)\,\sin^2(\delta_\ell). \tag{5.89}$$

The phase shift formalism exploits conservation of angular momentum to reduce the complex parameter a_ℓ to a single real one, δ_ℓ.

As a function of energy (or k) both $\delta(k)$ and $a_\ell(k)$ change, but they do it in a precisely related manner. This can be conveniently shown by drawing an Argand diagram for ka_ℓ in the complex plane. From the relation $a_\ell = (e^{i2\delta_\ell}-1)/2ik$ we have

$$ka_\ell = \frac{i}{2} + \frac{1}{2}e^{-i\left(\frac{\pi}{2}+2\delta_\ell\right)}. \tag{5.90}$$

This is a circle of radius $1/2$, known as the *unitary circle*,[11] centered at $i/2$ (see figure 5.8). Note that if δ_ℓ is small, then a_ℓ must stay near the bottom of the circle. It can be positive or negative, but a_ℓ is almost purely real because

$$a_\ell = \frac{1}{k}e^{i\delta_\ell}\sin\delta_\ell \approx \frac{1}{k}(1+i\delta_\ell)\delta_\ell \approx \frac{\delta_\ell}{k}. \tag{5.91}$$

On the other hand, if δ_ℓ is near $\pi/2$, then ka_ℓ is almost purely imaginary, $ka_\ell \approx i$, with maximal magnitude. Under such condition the ℓ-partial wave may be in resonance, a concept which we discuss in the next section. Note that equation (5.89) tells us that the maximum partial cross-section is achieved when $\sin^2\delta_\ell = 1$,

$$\sigma^{(\ell)}_{\max} = \frac{4\pi}{k^2}(2\ell+1). \tag{5.92}$$

[11] It is given this name because it can be shown to follow from the unitarity condition (or probability conservation) of the scattering process.

5.6 The optical theorem.

Since the Lengendre polynomials satisfy $P_\ell(1) = 1$, from equation (5.87) we see that the forward ($\theta = 0$) scattering amplitude is

$$f(0) = \frac{1}{k}\sum_{\ell=0}^{\infty}(2\ell+1)\, e^{i\delta_\ell}\sin(\delta_\ell), \tag{5.93}$$

therefore

$$\mathrm{Im}\big[f(0)\big] = \frac{1}{k}\sum_{\ell=0}^{\infty}(2\ell+1)\,\sin^2(\delta_\ell). \tag{5.94}$$

Comparing this result with equation (5.89) we arrive at a relation between the total cross-section and the imaginary part of the forward scattering amplitude, known as the *optical theorem*,[12]

$$\sigma = \frac{4\pi}{k}\mathrm{Im}\big[f(0)\big]. \tag{5.95}$$

Another instructive derivation of this theorem can be given in terms of the probability density current, which is defined below and satisfies a continuity equation

$$\mathbf{j}(\mathbf{r},t) = \frac{\hbar}{2mi}\left[\Psi^*(\boldsymbol{\nabla}\Psi) - (\boldsymbol{\nabla}\Psi^*)\Psi\right], \qquad \boldsymbol{\nabla}\mathbf{j} + \frac{\partial|\Psi|^2}{\partial t} = 0. \tag{5.96}$$

If Ψ is a stationary state, $\Psi(\mathbf{r},t) = \psi(\mathbf{r})\exp(-iEt/\hbar)$, that verifies the Schrödinger equation with a *real potential*, it is easy to show that the divergence of the current vanishes[13]

$$\boldsymbol{\nabla}\mathbf{j} = \frac{\hbar}{2mi}\boldsymbol{\nabla}\left[\psi^*(\boldsymbol{\nabla}\psi) - (\boldsymbol{\nabla}\psi^*)\psi\right] = 0. \tag{5.97}$$

Using Gauss's theorem over a spherical surface we must have then

$$\int \boldsymbol{\nabla}\mathbf{j}\,\mathrm{d}^3\mathbf{r} = \oint \mathbf{j}\,\mathrm{d}^2\mathbf{s} = \oint \mathbf{j}\,\hat{\mathbf{r}}\,\mathrm{d}^2 s = r^2\int_0^\pi \sin\theta\,\mathrm{d}\theta\int_0^{2\pi}\mathbf{j}\,\hat{\mathbf{r}}\,\mathrm{d}\varphi = 0, \tag{5.98}$$

where $\mathrm{d}^2\mathbf{s}$ is a two-dimensional differential element and $\hat{\mathbf{r}}$ is a unit vector orthogonal to the surface. This result, together with the continuity equation, means that the outgoing flux of particles must be equal to the incoming flux. To calculate $\mathbf{j}\,\hat{\mathbf{r}}$ we need the gradient operator in spherical coordinates

$$\boldsymbol{\nabla} = \frac{\partial}{\partial r}\hat{\mathbf{r}} + \frac{1}{r}\frac{\partial}{\partial\theta}\hat{\boldsymbol{\theta}} + \frac{1}{r\sin\theta}\frac{\partial}{\partial\varphi}\hat{\boldsymbol{\varphi}}. \tag{5.99}$$

Since $\hat{\mathbf{r}}\hat{\boldsymbol{\theta}} = \hat{\mathbf{r}}\hat{\boldsymbol{\varphi}} = 0$, only the radial derivative of the gradient operator contributes, and

$$\mathbf{j}\,\hat{\mathbf{r}} = \frac{\hbar}{2mi}\left[\psi^*(\boldsymbol{\nabla}\psi) - (\boldsymbol{\nabla}\Psi^*)\psi\right]\hat{\mathbf{r}} = \frac{\hbar}{2mi}\left(\psi^*\frac{\partial\psi}{\partial r} - \frac{\partial\psi^*}{\partial r}\psi\right). \tag{5.100}$$

Therefore, for a sphere of any radius r,

$$0 = r^2\int_0^\pi \sin\theta\,\mathrm{d}\theta\int_0^{2\pi}\mathrm{d}\varphi\left(\psi^*\frac{\partial\psi}{\partial r} - \frac{\partial\psi^*}{\partial r}\psi\right). \tag{5.101}$$

In particular, we can take r large enough to use for ψ the asymptotic formula (5.47),

$$\psi = A\left(e^{ikz} + f(\theta)\frac{e^{ikr}}{r}\right). \tag{5.47}$$

[12] This theorem was first found by Lord Rayleigh in classical electrodynamics as a relation between the absorption of light and the imaginary part of the index of refraction.

[13] In fact, for a stationary state $\partial|\Psi|^2/\partial t = 0$, so the continuity equation implies $\boldsymbol{\nabla}\mathbf{j} = 0$.

Upon inserting this expression into (5.101) and performing the integrals, it follows the optical theorem.[14]

This discussion shows that the optical theorem is closely related to the conservation of probability.[15] In elastic scattering the total number of particles entering the scattering region per unit time must be equal to the number leaving it per unit time in all directions. It is clear that, because of the scattering process, the intensity (flux) of the incident beam must decrease by an amount proportional to the total cross-section. As stated before, this decrease is due to the interference between the two terms in equation (5.47), which is proportional to $\mathrm{Im}\big[f(0)\big]$. The exact relation between the total cross-section and $\mathrm{Im}\big[f(0)\big]$ is given by the optical theorem.

5.7 Resonance scattering.

In some cases, the scattering cross-section for a given partial wave exhibits a pronounced peak in a certain energy interval. This is usually indicative of an underlying resonance process. There are numerous examples of such processes in atomic, nuclear and particle physics. For a brief elementary description consider a finite-range potential $V(r)$, where the appropriate effective potential for the radial wave function of the ℓ-th partial wave is

$$V_{\text{eff}} = V(r) + \frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}. \tag{5.102}$$

The centrifugal barrier, given by the second term, is always repulsive (positive). Therefore, when $V(r)$ itself is attractive (in some region) we may have a situation where the effective potential has an attractive well followed by a repulsive barrier at larger distances. Figure (5.9) shows a cross-section along the z-axis of such a potential (not to scale). The potential has a high value much larger than the energy of the incident particle E in a thick shell around the origin, surrounding an inner region where the potential is much smaller. If the barrier were *infinitely* high we could have a particle trapped inside forming a stationary bound state. In practice, of course, the barrier is finite and the particle can escape through a quantum-mechanical tunnel effect. We may call it a *quasi-bound* state of *resonance* energy E_r.

Now, if we send particles to the scattering center at various energies E, the corresponding scattering phase shifts δ_ℓ will rise through the value $\pi/2$ as the incident particle energy rises through that of the quasi-bound state E_r. At the same time, the partial-wave cross-section passes through its maximum possible value of $4\pi(2\ell+1)/k^2$. It can be seen that such a sharp rise in the phase shift is, in the time-dependent Shcrödinger equation, associated with a delay of the emergence of the trapped particle.

Let's see how the scattering amplitudes vary in the vicinity of the resonance energy. If the scattering cross section $\sigma^{(\ell)} = 4\pi(2\ell+1)\,|a_\ell|^2$ for the partial wave ℓ becomes large at energies close to E_r, then the phase shift δ_ℓ must go through $\pi/2$. Since $a_\ell = 1/(k\cot\delta_\ell - ik)$ and $\cot\delta_\ell \approx 0$ near the resonance energy (in fact $\cot\delta_\ell(E_r) = \cot\pi/2 = 0$), we may attempt to expand δ_ℓ as follows

$$\cot\delta_\ell(E) = \cot\delta_\ell(E_r) + (E-E_r)\left(\frac{\mathrm{d}\cot\delta_\ell}{\mathrm{d}E}\right)_{E=E_r} + \ldots \approx (E-E_r)\left(\frac{\mathrm{d}\cot\delta_\ell}{\mathrm{d}E}\right)_{E=E_r}.$$

The derivative of $\cot\delta_\ell$ evaluated at $E=E_r$ is just the negative of the phase shift derivative,

$$\left(\frac{\mathrm{d}\cot\delta_\ell}{\mathrm{d}E}\right)_{E=E_r} = \left(-\frac{1}{\sin^2\delta_\ell}\frac{\mathrm{d}\delta_\ell}{\mathrm{d}E}\right)_{E=E_r} = -\left(\frac{\mathrm{d}\delta_\ell}{\mathrm{d}E}\right)_{E=E_r} = -\frac{2}{\Gamma},$$

[14] Although the calculation is not very lengthy, we omit the details here. The interested reader is referred to section 7.3 of [15].

[15] In quantum mechanics a process that conserves probability is also said that conserves or satisfies unitarity, since transformations in the Hilbert space that keep transition probabilities invariant are implemented by unitary operators.

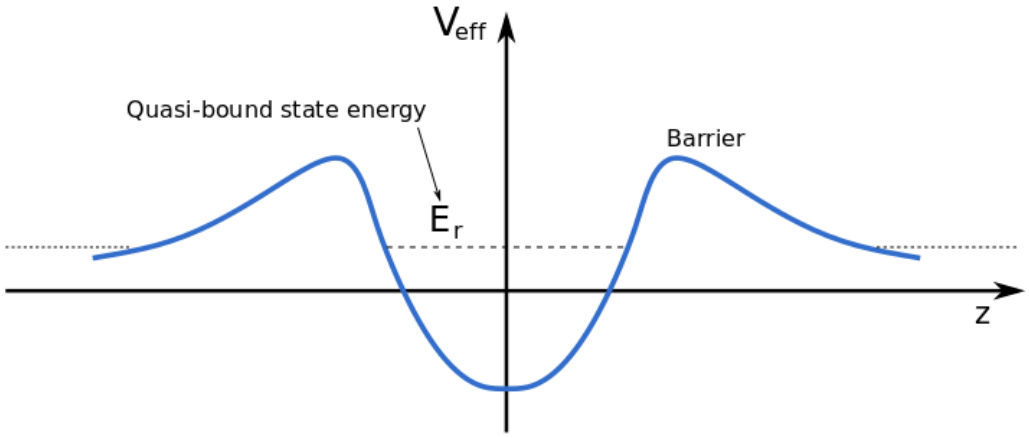


Figure 5.9: Example of $V_{\rm eff}$ versus z. For $\ell \neq 0$ the barrier can be due to $\frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2}$. For $\ell = 0$ the barrier must be due to $V(r)$ itself.

where we have defined the *width* Γ through the relation

$$\frac{1}{\Gamma} = \frac{1}{2}\left(\frac{\mathrm{d}\delta_\ell}{\mathrm{d}E}\right)_{E=E_r}. \tag{5.103}$$

So

$$\cot\delta_\ell(E) \approx -\frac{2(E-E_r)}{\Gamma} \tag{5.104}$$

For a given partial wave the scattering cross-section becomes

$$\sigma^{(\ell)} = 4\pi(2\ell+1)\,|a_\ell|^2 = \frac{4\pi(2\ell+1)}{k^2}\left|\frac{1}{\cot\delta_\ell - i}\right|^2 = \frac{4\pi(2\ell+1)}{k^2}\frac{1}{\cot^2\delta_\ell + 1}$$
$$\approx \frac{4\pi(2\ell+1)}{k^2}\frac{(\Gamma/2)^2}{(E-E_r)^2+(\Gamma/2)^2}.$$

$$\sigma^{(\ell)} \approx \frac{4\pi(2\ell+1)}{k^2}\frac{(\Gamma/2)^2}{(E-E_r)^2+(\Gamma/2)^2}. \tag{5.105}$$

This formula tells us that we are justified to have defined Γ as the full width at half-maximum, provided the resonance is reasonably narrow so that the variation in $1/k^2$ can be ignored.

There are situations in nuclear physics where the potential barrier is so thick that the decay rate of the *quasi-bound* states Γ is very small. So small that nuclei in these states can be found in nature rather than as resonances in scattering processes. A prototypical example is provided by nuclei that are unstable against the emission of alpha particles. In transitions in which the α particle is emitted in an s wave, such as $^{238}\mathrm{U} \to {}^{234}\mathrm{Th} + \alpha$ or $^{226}\mathrm{Ra} \to {}^{222}\mathrm{Rn} + \alpha$, the barrier arises purely from the Coulomb potential, which in alpha decay is $V(r) = 2Ze^2/r$, where Z is the atomic number of the final nucleus. For the ^{238}U alpha decay the barrier is so high and thick that leads to a lifetime of 4.47×10^9 years.

5.8 The Born approximation.

The Born approximation uses the integral form of the Schrödinger equation, which we will derive now. First, note that the time-independent Schrödinger equation,

$$-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi = E\psi, \tag{5.106}$$

can be written in a compact way as

$$\left(\nabla^2 + k^2\right)\psi = Q, \tag{5.107}$$

where

$$k = \frac{\sqrt{2mE}}{\hbar}, \qquad \text{and} \qquad Q = \frac{2m}{\hbar^2} V\psi. \tag{5.108}$$

This looks like a Helmholtz equation, but the inhomogeneous term itself, Q, depends on ψ. The solution can be written as

$$\psi(\mathbf{r}) = \int G(\mathbf{r} - \mathbf{r}_0) Q(\mathbf{r}_0)\, \mathrm{d}^3\mathbf{r}_0, \tag{5.109}$$

where $G(\mathbf{r})$ is the Green function that satisfies the Helmholtz equation with a delta source

$$\left(\nabla^2 + k^2\right) G(\mathbf{r}) = \delta^3(\mathbf{r}). \tag{5.110}$$

This can be easily verified by showing that $\left(\nabla^2 + k^2\right)\psi = Q$ holds. Indeed,

$$\begin{aligned}\left(\nabla^2 + k^2\right)\psi(\mathbf{r}) &= \int \Big[\left(\nabla^2 + k^2\right) G(\mathbf{r} - \mathbf{r}_0)\Big] Q(\mathbf{r}_0)\, \mathrm{d}^3\mathbf{r}_0 \\ &= \int \delta^3(\mathbf{r} - \mathbf{r}_0)\, Q(\mathbf{r}_0)\, \mathrm{d}^3\mathbf{r}_0 = Q(\mathbf{r}).\end{aligned} \tag{5.111}$$

To find $G(\mathbf{r})$ it is easier to work with its Fourier transform, so that the differential equation (5.110) transforms into an algebraic equation. Let

$$G(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int e^{i\mathbf{s}\cdot\mathbf{r}} g(\mathbf{s})\, \mathrm{d}^3\mathbf{s}. \tag{5.112}$$

Then

$$\left(\nabla^2 + k^2\right) G(\mathbf{r}) = \frac{1}{(2\pi)^{3/2}} \int \Big[\left(\nabla^2 + k^2\right) e^{i\mathbf{s}\cdot\mathbf{r}}\Big] g(\mathbf{s})\, \mathrm{d}^3\mathbf{s}. \tag{5.113}$$

Using now the relations

$$\nabla^2 e^{i\mathbf{s}\cdot\mathbf{r}} = -s^2 e^{i\mathbf{s}\cdot\mathbf{r}}, \qquad \text{and} \qquad \delta^3(\mathbf{r}) = \frac{1}{(2\pi)^3} \int e^{i\mathbf{s}\cdot\mathbf{r}}\, \mathrm{d}^3\mathbf{s}, \tag{5.114}$$

the equation $\left(\nabla^2 + k^2\right) G(\mathbf{r}) = \delta^3(\mathbf{r})$ transforms into

$$\frac{1}{(2\pi)^{3/2}} \int \left(-s^2 + k^2\right) e^{i\mathbf{s}\cdot\mathbf{r}} g(\mathbf{s})\, \mathrm{d}^3\mathbf{s} = \frac{1}{(2\pi)^3} \int e^{i\mathbf{s}\cdot\mathbf{r}}\, \mathrm{d}^3\mathbf{s}. \tag{5.115}$$

From this we conclude that

$$g(\mathbf{s}) = \frac{1}{(2\pi)^{3/2}(k^2 - s^2)}, \tag{5.116}$$

and inserting this in the Fourier transform we have

$$G(\mathbf{r}) = \frac{1}{(2\pi)^3} \int \frac{e^{i\mathbf{s}\cdot\mathbf{r}}}{(k^2 - s^2)}\, \mathrm{d}^3\mathbf{s}. \tag{5.117}$$

We do this integral using spherical coordinates (s, θ, φ), with the polar axis running along $\mathbf{r}$ (see figure 5.10), so that $\mathbf{s}\cdot\mathbf{r} = sr\cos\theta$ ($\mathbf{r}$ is fixed in the integral over $\mathbf{s}$). The φ integral is just 2π and the one over θ is

$$\int_0^\pi e^{isr\cos\theta} \sin\theta\, \mathrm{d}\theta = -\frac{e^{isr\cos\theta}}{isr}\bigg|_0^\pi = \frac{1}{isr}\left(e^{isr} - e^{-isr}\right) = \frac{2\sin(sr)}{sr}. \tag{5.118}$$

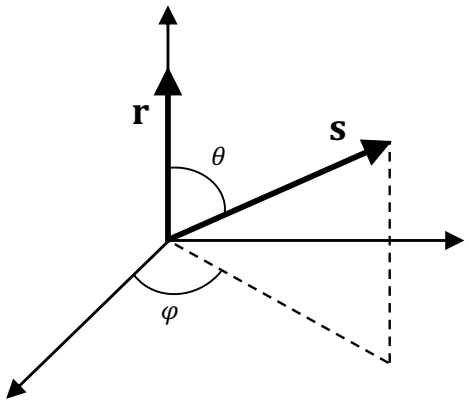


Figure 5.10: Coordinate frame to do the integral (5.117).

Therefore

$$G(\mathbf{r}) = \frac{2\pi}{(2\pi)^3}\frac{2}{r}\int_0^\infty \frac{\sin(sr)}{s(k^2-s^2)}\, s^2 \mathrm{d}s = \frac{1}{4\pi^2 r}\int_{-\infty}^{\infty} \frac{s\sin(sr)}{k^2-s^2}\,\mathrm{d}s. \tag{5.119}$$

We have extended the integration range to $(-\infty, +\infty)$ and divided by 2, since the function to be integrated is even. The last expression can be written in the following form

$$G(\mathbf{r}) = \frac{i}{8\pi^2 r}\left[\int_{-\infty}^{\infty} \frac{se^{isr}}{(s-k)(s+k)}\,\mathrm{d}s \;-\; \int_{-\infty}^{\infty} \frac{se^{-isr}}{(s-k)(s+k)}\,\mathrm{d}s\right] = \frac{i}{8\pi^2 r}(I_1 - I_2). \tag{5.120}$$

These two integrals can be evaluated using Cauchy's integral formula, which says that

$$\oint_{\mathcal{C}} \frac{f(z)}{(z-z_0)}\,\mathrm{d}z \;=\; 2\pi i\, f(z_0), \tag{5.121}$$

for any $f(z)$ analytic inside and on a simple closed curve $\mathcal{C}$ on the complex plane that encloses the pole $z = z_0$. In our case, the integrations are along the real axis and go over two pole singularities at $z_0 = \pm k$. To evaluate I_1 we choose the contour $\mathcal{C}_1$ shown in the first plot of figure (5.11) and let the radius of the semicircle go to infinity. In that case, the factor e^{isr} goes to zero when s has a large positive imaginary part and the integral in the semicircle becomes zero. The contour $\mathcal{C}_1$ encloses only the singularity at $s = +k$, so

$$I_1 = \oint_{\mathcal{C}_1} \left(\frac{se^{isr}}{s+k}\right)\frac{1}{s-k}\,\mathrm{d}s = 2\pi i \left(\frac{se^{isr}}{s+k}\right)_{s=k} = i\pi e^{ikr}. \tag{5.122}$$

For the I_2 integral we choose the contour $\mathcal{C}_2$, which encloses the pole $s = -k$ and goes around in the clockwise direction, thereby picking up a minus sign, so

$$I_2 = -\oint_{\mathcal{C}_2} \left(\frac{se^{-isr}}{s-k}\right)\frac{1}{s+k}\,\mathrm{d}s = -2\pi i \left(\frac{se^{-isr}}{s-k}\right)_{s=-k} = -i\pi e^{ikr}. \tag{5.123}$$

Finally, the Green function for the Helmholtz equation is

$$G(\mathbf{r}) = \frac{i}{8\pi^2 r}\left[i\pi e^{ikr} - \left(-i\pi e^{ikr}\right)\right] = -\frac{e^{ikr}}{4\pi r}. \tag{5.124}$$

This is a solution to the equation $\left(\nabla^2 + k^2\right) G(\mathbf{r}) = \delta^3(\mathbf{r})$. Other acceptable solutions would be found by choosing different ways of going around the poles in the contour integrations. Remember that we can add to $G(\mathbf{r})$ any function that satisfies the homogeneous Helmholtz equation, $\left(\nabla^2 + k^2\right) G_0(\mathbf{r}) = 0$, and the result $(G + G_0)$ will still be an acceptable solution.

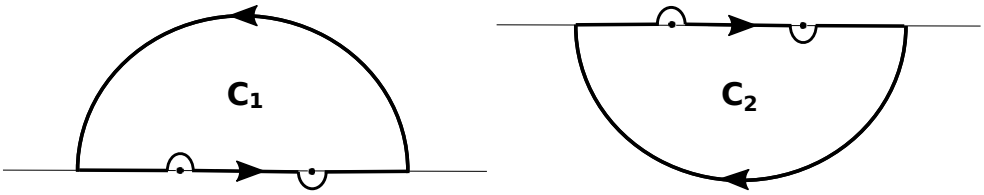


Figure 5.11: Closing the contour for the integrations in equation (5.120).

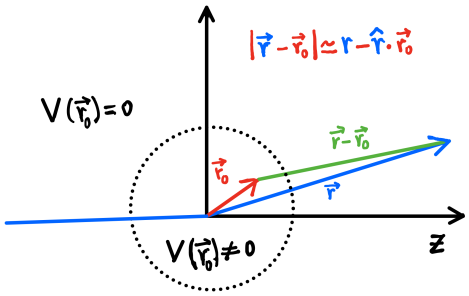


Figure 5.12: The blue line indicates the trajectory of a particle scattered by a localized potential.

Now we can go back to equation (5.109), $\psi(\mathbf{r}) = \int G(\mathbf{r}-\mathbf{r}_0)Q(\mathbf{r}_0)\,\mathrm{d}^3\mathbf{r}_0$, and write the general solution to the Schrödinger equation as

$$\psi(\mathbf{r}) = \psi_0(\mathbf{r}) - \frac{m}{2\pi\hbar^2}\int \frac{e^{ik|\mathbf{r}-\mathbf{r}_0|}}{|\mathbf{r}-\mathbf{r}_0|}V(\mathbf{r}_0)\psi(\mathbf{r}_0)\,\mathrm{d}^3\mathbf{r}_0, \tag{5.125}$$

where ψ_0 satisfies the free-particle Schrödinger equation, $\left(\nabla^2 + k^2\right)\psi_0 = 0$. Equation (5.125) is the integral form of the Schrödinger equation and it is entirely equivalent to the more familiar differential form. Note that the wave function we are looking for appears in both sides of the equation ($\psi(\mathbf{r})$ and $\psi(\mathbf{r}_0)$), so we cannot do explicitly the integral unless we already know the solution. Nevertheless, this integral equation can be solved using an iterative procedure in a way that is very well suited to study scattering problems.

5.8.1 The first Born approximation.

In a scattering problem the potential $V(\mathbf{r}_0)$ is localized about some given point, for example around $\mathbf{r}_0 = 0$, and goes to zero far away. Our goal is to calculate $\psi(\mathbf{r})$ in the asymptotic region where $|\mathbf{r}| \gg |\mathbf{r}_0|$ and $|\mathbf{r}-\mathbf{r}_0|$ can be approximated as

$$|\mathbf{r}-\mathbf{r}_0|^2 = r^2 + r_0^2 - 2\mathbf{r}\cdot\mathbf{r}_0 \approx r^2 - 2\mathbf{r}\cdot\mathbf{r}_0 = r^2\left(1 - 2\frac{\mathbf{r}\cdot\mathbf{r}_0}{r^2}\right) = r^2\left(1 - 2\frac{\hat{\mathbf{r}}\cdot\mathbf{r}_0}{r}\right). \tag{5.126}$$

Therefore[16]

$$|\mathbf{r}-\mathbf{r}_0| \approx r - \hat{\mathbf{r}}\cdot\mathbf{r}_0 = r - \frac{\mathbf{k}}{k}\cdot\mathbf{r}_0, \tag{5.127}$$

where $\mathbf{k} \equiv k\hat{\mathbf{r}}$ is the wave vector (momentum $\hbar\mathbf{k}$) of the scattered particle. See Figs. 5.12 and 5.13. We have then

$$e^{ik|\mathbf{r}-\mathbf{r}_0|} \approx e^{ikr}e^{-i\mathbf{k}\cdot\mathbf{r}_0}, \qquad \frac{e^{ik|\mathbf{r}-\mathbf{r}_0|}}{|\mathbf{r}-\mathbf{r}_0|} \approx \frac{e^{ikr}}{r}e^{-i\mathbf{k}\cdot\mathbf{r}_0}. \tag{5.128}$$

[16] Using $\sqrt{1-2x} \approx 1 - x$ for small x.

In the denominator of the last expression we have made the stronger approximation $|\mathbf{r}-\mathbf{r}_0| \approx r$. This is appropriate for an expansion in powers of the small quantity r_0/r, while keeping only the lowest order term. Now we take $\psi_0(\mathbf{r}) = Ae^{ikz}$ to be the incident plane wave and use equation (5.125) to write, for large r,

$$\psi(\mathbf{r}) \approx Ae^{ikz} - \frac{m}{2\pi\hbar^2}\frac{e^{ikr}}{r}\int e^{-i\mathbf{k}\cdot\mathbf{r}_0}V(\mathbf{r}_0)\psi(\mathbf{r}_0)\,\mathrm{d}^3\mathbf{r}_0. \tag{5.129}$$

By comparing this with $\psi \approx A[e^{ikz} + f(\theta,\varphi)e^{ikr}/r]$ we can identify the scattering amplitude as

$$f(\theta,\varphi) = -\frac{m}{2\pi\hbar^2 A}\int e^{-i\mathbf{k}\cdot\mathbf{r}_0}V(\mathbf{r}_0)\psi(\mathbf{r}_0)\,\mathrm{d}^3\mathbf{r}_0. \tag{5.130}$$

This last expression is *exact* in the sense that provides the scattering amplitude for the outgoing wave far away from the scattering center without assuming anything else than a localized potential. Now we introduce the *Born approximation*, which is obtained by treating the potential $V(\mathbf{r}_0)$ as a small perturbation and assuming that the incoming plane wave is not substantially altered by the potential. In such a case we may write inside the integral

$$\psi(\mathbf{r}_0) \approx \psi_0(\mathbf{r}_0) = Ae^{ikz_0} = Ae^{i\mathbf{k}'\cdot\mathbf{r}_0}, \qquad \text{where} \quad \mathbf{k}' \equiv k\hat{\mathbf{z}}. \tag{5.131}$$

This represents a *weak potential* approximation compared to the incident energy, so that the deflection of the particle by the potential is very small. The scattering amplitude in the Born approximation is then proportional to the Fourier transform of the potential

$$f(\theta,\varphi) \approx -\frac{m}{2\pi\hbar^2}\int e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}_0}\,V(\mathbf{r}_0)\,\mathrm{d}^3\mathbf{r}_0. \tag{5.132}$$

Both $\mathbf{k}'$ and $\mathbf{k}$ have magnitude k, but $\mathbf{k}'$ points in the direction of the incident beam, while $\mathbf{k}$ points in the direction of the scattered particle (figure 5.13). Denoting $\boldsymbol{\kappa} = \mathbf{k}'-\mathbf{k}$, the momentum transfer in the process is $\hbar\boldsymbol{\kappa} = \hbar(\mathbf{k}'-\mathbf{k})$. Note that for $\theta = 0$ we have $\mathbf{k}' = \mathbf{k}$, therefore $e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}_0} = 1$ and Eq. (5.132) implies that $\mathrm{Im}\big[f(0)\big] = 0$. According to the optical theorem, this can only be strictly true if the total cross-section vanishes. In fact, the Born approximation violates unitarity and the optical theorem, but the size of the violation is negligible in the cases where the Born approximation is applicable.

The Born approximation is valid when either of the two following conditions holds:

1. $\langle V\rangle \ll \dfrac{\hbar^2}{ma^2}$.
2. $\langle V\rangle \ll \dfrac{\hbar^2}{ma^2}\,ka$.

Where a is the range of the potential and $\langle V\rangle$ is the averaged potential defined as

$$\langle V\rangle = \frac{1}{4\pi a^2}\int \frac{V(r)}{r}\,\mathrm{d}^3\mathbf{r}. \tag{5.133}$$

When the first condition is satisfied, the Born approximation is valid for all energies. The second condition shows that the approximation is always applicable for sufficiently high-energy particles. Since this condition is weaker than the first one, if the potential can be regarded as a small perturbation at low energies, then at high-energies it can still be assumed to be a small perturbation and the Born approximation to hold. The opposite is not necessarily true: the validity of the Born approximation at high energy for a specific potential does not imply its validity at low energy.

For low energy (long wavelength) scattering, the exponential factor inside the integral can be taken to be constant over the scattering region, which is very small compared to wavelength

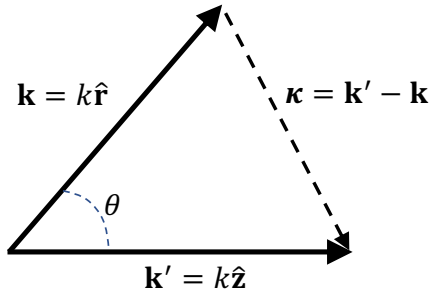


Figure 5.13: The wave vector $\mathbf{k}$ pointing in the scattered direction, and $\mathbf{k}'$ pointing in the incident direction. The momentum transfer in the process is $\hbar\boldsymbol{\kappa} = \hbar(\mathbf{k}' - \mathbf{k})$. Since $|\mathbf{k}'| = |\mathbf{k}| = k$, we have $\kappa = 2k\sin(\theta/2)$.

associated with the incident particle. In this case, we have $e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}_0} \approx 1$ and the Born approximation simplifies to

$$f(\theta,\varphi) \approx -\frac{m}{2\pi\hbar^2}\int V(\mathbf{r})\,\mathrm{d}^3\mathbf{r}, \qquad \text{(low energy)}. \tag{5.134}$$

Let's take, as an example, low-energy *soft-sphere*[17] scattering,

$$V(\mathbf{r}) = \begin{cases} V_0, & \text{if } r \le a, \\ 0, & \text{if } r > a. \end{cases} \tag{5.135}$$

where V_0 is a small parameter. The low-energy scattering amplitude would be

$$f(\theta,\varphi) \approx -\frac{m}{2\pi\hbar^2}V_0\left(\frac{4}{3}\pi a^3\right). \tag{5.136}$$

The differential and the total cross-sections are

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f(\theta,\varphi)|^2 \approx \left(\frac{2mV_0a^3}{3\hbar^2}\right)^2, \qquad \sigma \approx 4\pi\left(\frac{2mV_0a^3}{3\hbar^2}\right)^2. \tag{5.137}$$

If the potential is spherically symmetric, $V(\mathbf{r}) = V(r)$, the Born approximation can be written in a simpler form. Recalling that $\boldsymbol{\kappa} = \mathbf{k}' - \mathbf{k}$, and taking the polar axis for the $\mathbf{r}_0$ integral to lie along $\boldsymbol{\kappa}$ we have

$$(\mathbf{k}' - \mathbf{k})\cdot\mathbf{r}_0 = \kappa r_0\cos\theta_0. \tag{5.138}$$

Therefore

$$f(\theta) \approx -\frac{m}{2\pi\hbar^2}\int e^{i\kappa r_0\cos\theta_0}\,V(r_0)\,r_0^2\sin\theta_0\,\mathrm{d}r_0\,\mathrm{d}\theta_0\,\mathrm{d}\varphi_0. \tag{5.139}$$

The integral over φ gives a factor of 2π, and that over θ_0 has already appeared in equation (5.118). Dropping the subscript on r_0 we are left with the following scattering amplitude for spherically symmetric potentials in the Born approximation (but not necessarily low energy),

$$f(\theta) \approx -\frac{2m}{\hbar^2\kappa}\int_0^\infty rV(r)\sin(\kappa r)\,\mathrm{d}r. \tag{5.140}$$

Note that the angular dependence goes in the κ factor, $\kappa = 2k\sin(\theta/2)$.

[17] Of course, hard-sphere scattering, where $V_0 = \infty$, is not a weak potential and cannot be treated with the Born approximation.

A historically important example of spherical potential is the *Yukawa potential*, which was first used as a crude model for the binding force in atomic nuclei. It is written as

$$V(r) = \beta \frac{e^{-\mu r}}{r}, \tag{5.141}$$

where β and μ are constants. The Born approximation gives[18]

$$f(\theta) \approx -\frac{2m\beta}{\hbar^2 \kappa} \int_0^\infty e^{-\mu r} \sin(\kappa r)\, \mathrm{d}r = -\frac{2m\beta}{\hbar^2(\mu^2 + \kappa^2)}. \tag{5.142}$$

If we set $\mu = 0$ and $\beta = q_1 q_2/(4\pi\varepsilon_0)$, the Yukawa potential reduces to the Coulomb potential. The scattering amplitude would be

$$f(\theta) \approx -\frac{2mq_1q_2}{4\pi\varepsilon_0 \hbar^2 \kappa^2}. \tag{5.143}$$

Recalling that $\kappa = 2k\sin(\theta/2)$ and $k = \sqrt{2mE}/\hbar$, this can be written as

$$f(\theta) = -\frac{q_1 q_2}{16\pi\varepsilon_0 E \sin^2(\theta/2)}. \tag{5.144}$$

The differential cross-section is

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f(\theta)|^2 = \left(\frac{q_1 q_2}{16\pi\varepsilon_0 E \sin^2(\theta/2)} \right)^2. \tag{5.145}$$

This is the famous Rutherford formula. It is notorious that for the Coulomb potential, classical mechanics, the Born approximation and quantum field theory all yield the same result for the differential cross-section.

5.8.2 The Born series.

Let us write the integral form of the Schrödinger equation as

$$\psi(\mathbf{r}) = \psi_0(\mathbf{r}) + \int g(\mathbf{r} - \mathbf{r}_0) V(\mathbf{r}_0) \psi(\mathbf{r}_0)\, \mathrm{d}^3\mathbf{r}_0, \tag{5.146}$$

where ψ_0 is the incident wave, V is the scattering potential and

$$g(\mathbf{r}) = -\frac{m}{2\pi\hbar^2} \frac{e^{ikr}}{r} \tag{5.147}$$

is the Green's function. Schematically,

$$\psi = \psi_0 + \int gV\psi. \tag{5.148}$$

We may take, as a first approximation, $\psi \approx \psi_0$ and insert it in the previous equation to get $\psi = \psi_0 + \int gV\psi_0$. After this *first step*, we could try to plug the result again under the integral sign to obtain a (presumably) better approximation:

$$\psi = \psi_0 + \int gV \left(\psi_0 + \int gV\psi_0 \right) = \psi_0 + \int gV\psi_0 + \int\int gVgV\psi_0. \tag{5.149}$$

Iterating this procedure we obtain a formal series for ψ:

$$\psi = \psi_0 + \int gV\psi_0 + \int\int gVgV\psi_0 + \int\int\int gVgVgV\psi_0 + \ldots \tag{5.150}$$

[18] This will be shown in one of the problems of this chapter.

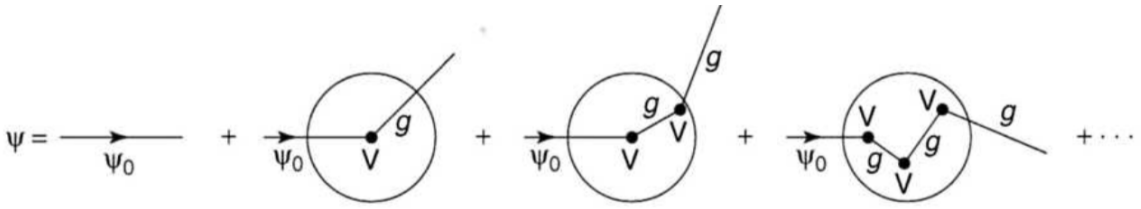


Figure 5.14: Diagrammatic interpretation of the Born series (taken from Griffiths [7]).

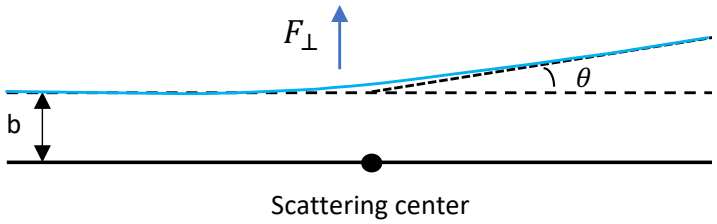


Figure 5.15: The impulse approximation in classical scattering. The actual trajectory followed by the particle is shown by the blue line. It emerges at an angle θ which is estimated as $\theta = \arctan(I/p)$, where p is the momentum, $I = \int F_\perp \, \mathrm{d}t$ is the transverse impulse and $F_\perp$ is the perpendicular component of the force exerted by the potential on the particle. The integral $\int F_\perp \, \mathrm{d}t$ is computed along the dashed straight line that the particle would follow if there were no potential.

This is the *Born series*, which can be diagrammatically represented as shown in figure (5.14). The zeroth order corresponds to *no scattering* at all, so that the incident wave goes by unscathed. In the first-order approximation, the effect of the potential is assumed to originate from a single modification of the wave function at a localized point, and afterward the particle continues in a different direction without interacting again. This holds some similarity with the *impulse approximation* in classical scattering theory. In that approximation, the scattering angle is computed using the transverse impulse that the potential would deliver to the particle if it followed a straight line without deflection (see figure 5.15).

The second Born approximation arises from truncating the Born series (5.150) after the third term. In this case, the particle suffers a first interaction at a given point, it propagates as a free particle to a new location to interact again before leaving the potential region. This procedure can in principle be continued as long as desired with the hope that the emerging series will converge to the exact solution of the problem. In this context the Green's function is often called *the propagator*, because it describes how the disturbance propagates between one interaction and the next. The Born series was the inspiration for Feynman's formulation of relativistic quantum mechanics, which is expressed entirely in terms of vertex factors (V) and propagators (g), connected together in *Feynman diagrams*.

5.9 Scattering by a complex potential. Absorption processes.

In the preceding sections we have considered only the elastic scattering between two spinless particles by a real potential. Remember that the wave number $k = \sqrt{2mE}/\hbar$ was treated always as a constant: the initial particle comes in with the same value of k as it goes scattered away. In general, however, when a particle collides with a target, *non-elastic* scattering may occur. For example, when electrons with sufficient kinetic energy collide with atoms in the ground state, excitations of the atoms are possible in addition to elastic scattering. If an inelastic process such

as this takes place, it means that a fraction of the incident particles are *removed* from the *elastic channel*. The electron looses some energy in the process of exciting the atom and goes out with a different wave number $k' = \sqrt{2mE'}/\hbar$. It is also said that the particles are *absorbed* (partial absorption of the elastic channel), even if they are scattered with a different energy and not really absorbed by the target in a literal sense. The different physical processes that may take place in a scattering collision are referred to as the various *channels* of the reaction/collision. Thus, we have generically the inelastic channel and the elastic channel. The inelastic category can be subdivided into several channels for different inelastic processes.

It is easy to see that the absorption of particles can be represented in a phenomenological way by introducing a complex potential (also called optical potential) having a negative imaginary part

$$V(\mathbf{r}) = V_R(\mathbf{r}) + iV_I(\mathbf{r}). \tag{5.151}$$

As usual, we look for stationary solutions of the time-dependent Schrödinger equation having the form $\Psi(\mathbf{r},t) = \psi(\mathbf{r})\exp(-iEt/\hbar)$, where $E = \hbar^2k^2/(2m)$. The wave function $\psi(\mathbf{r})$ satisfies the time-independent Schrödinger equation

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V_R(\mathbf{r}) + iV_I(\mathbf{r})\right)\psi(\mathbf{r}) = E\psi(\mathbf{r}), \tag{5.152}$$

while its complex conjugate $\psi^*(\mathbf{r})$ satisfies

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V_R(\mathbf{r}) - iV_I(\mathbf{r})\right)\psi^*(\mathbf{r}) = E\psi^*(\mathbf{r}), \tag{5.153}$$

Using these two equations we find that the divergence of the probability current density is proportional to the product of the imaginary part of the potential and the probability density $|\Psi|^2 = |\psi|^2$,

$$\boldsymbol{\nabla}\mathbf{j} = \frac{\hbar}{2mi}\boldsymbol{\nabla}\left[\psi^*(\boldsymbol{\nabla}\psi) - (\boldsymbol{\nabla}\psi^*)\psi\right] = \frac{\hbar}{2mi}\left[\psi^*(\nabla^2\psi) - (\nabla^2\psi^*)\psi\right] = \frac{2}{\hbar}V_I|\psi|^2. \tag{5.154}$$

A non-zero divergence means that, if $V_I < 0$, there is a local absorption of the incident beam. The rate of absorption of incident particles per unit time per unit volume at the point $\mathbf{r}$ is $2V_I(\mathbf{r})/\hbar$. The total number of particles absorbed per unit time within the volume V is

$$N_{\text{abs}} = \frac{2}{\hbar}\int_V V_I|\psi|^2\,\mathrm{d}^3\mathbf{r}.$$

We consider now the case of a central complex potential, $V(r) = V_R(r) + iV_I(r)$. It is possible to make a partial wave analysis of the problem, as we did for a central real potential, but we will skip the details and jump to the main results.[19] It turns out that, for the elastic scattering amplitude, f_{el}, we can directly take equation (5.88) and extend the phase shift δ_ℓ to be a complex number,

$$\delta_\ell \;=\; \mathrm{Re}(\delta_\ell) \;+\; i\,\mathrm{Im}(\delta_\ell).$$

Thus, the elastic amplitude can be written as

$$f_{\text{el}}(k,\theta) = \frac{1}{2ik}\sum_{\ell=0}^{\infty}(2\ell+1)\left[S_\ell(k) - 1\right]P_\ell(\cos\theta), \tag{5.155}$$

where

$$S_\ell(k) = e^{2i\delta_\ell(k)} = e^{2i\mathrm{Re}(\delta_\ell(k))}e^{-2\mathrm{Im}(\delta_\ell(k))}, \tag{5.156}$$

is the so-called S-matrix element and $\delta_\ell(k)$ is a *complex* phase shift for the ℓ-wave. In this context, the S-matrix element is related to the amplitude of the outgoing elastic wave and,

[19] For a more extensive treatment see section 12.8 of [2].

since particles can be removed from the incident (elastic) channel, its modulus must satisfy the relation $|S_\ell(k)| \leq 1$, which implies that $\mathrm{Im}(\delta_\ell(k)) \geq 0$. If only elastic scattering occurs ($V_I = 0$) we have $|S_\ell(k)| = 1$ and $\mathrm{Im}(\delta_\ell(k)) = 0$. It is common to introduce an absorption factor, also called *inelasticity*, defined as

$$\eta_\ell(k) = e^{-2\mathrm{Im}(\delta_\ell(k))}, \tag{5.157}$$

so that the S-matrix element takes the form

$$S_\ell(k) = \eta_\ell(k)\, e^{2i\mathrm{Re}(\delta_\ell(k))}. \tag{5.158}$$

The absorption factor satisfies the relation $0 \leq \eta_\ell(k) \leq 1$. The case $\eta_\ell = 1$ corresponds to pure elastic scattering. The partial amplitude for elastic scattering with angular momentum ℓ is written in terms of S_ℓ as

$$a_\ell = \frac{1}{2ik}\big(e^{2i\delta_\ell} - 1\big) = \frac{1}{2ik}\big(S_\ell - 1\big). \tag{5.159}$$

The total and differential elastic cross-sections are given by

$$\frac{\mathrm{d}\sigma_{\mathrm{el}}}{\mathrm{d}\Omega} = |f_{\mathrm{el}}|^2, \qquad \sigma_{\mathrm{tot}}^{\mathrm{el}} = 4\pi \sum_{\ell=0}^{\infty}(2\ell+1)|a_\ell|^2. \tag{5.160}$$

In terms of S_ℓ, the elastic differential cross-section is

$$\frac{\mathrm{d}\sigma_{\mathrm{el}}}{\mathrm{d}\Omega} = |f_{\mathrm{el}}|^2, \tag{5.161}$$

and the total elastic cross-section is

$$\sigma_{\mathrm{tot}}^{\mathrm{el}} = \frac{\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell+1)\,|1 - S_\ell|^2 = \frac{\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell+1)\left|1 - \eta_\ell\, e^{2i\mathrm{Re}(\delta_\ell)}\right|^2 = \sum_{\ell=0}^{\infty}\sigma_\ell^{\mathrm{el}}, \tag{5.162}$$

where the partial elastic cross-sections are

$$\sigma_\ell^{\mathrm{el}} = \frac{\pi}{k^2}(2\ell+1)\,|1 - S_\ell|^2 \leq \frac{4\pi(2\ell+1)}{k^2}. \tag{5.163}$$

The total inelastic cross-section (also called absorption cross-section) can be obtained from the ratio of the outgoing radial flux and the incident (plane wave) flux. The result is

$$\sigma_{\mathrm{tot}}^{\mathrm{abs}} = \frac{\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell+1)(1 - |S_\ell|^2) = \frac{\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell+1)(1 - \eta_\ell^2) = \sum_{\ell=0}^{\infty}\sigma_\ell^{\mathrm{abs}}, \tag{5.164}$$

where

$$\sigma_\ell^{\mathrm{abs}} = \frac{\pi}{k^2}(2\ell+1)(1 - \eta_\ell^2) \leq \frac{\pi(2\ell+1)}{k^2}. \tag{5.165}$$

The total cross-section is the sum of the elastic and inelastic cross-sections,

$$\begin{aligned}\sigma_{\mathrm{tot}} = \sigma_{\mathrm{tot}}^{\mathrm{el}} + \sigma_{\mathrm{tot}}^{\mathrm{abs}} &= \frac{2\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell+1)\Big[1 - \mathrm{Re}\,\big(S_\ell\big)\Big] \\ &= \frac{2\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell+1)\Big[1 - \eta_\ell\cos(2\mathrm{Re}(\delta_\ell))\Big] = \sum_{\ell=0}^{\infty}\sigma_\ell,\end{aligned} \tag{5.166}$$

where

$$\sigma_\ell = \frac{2\pi}{k^2}(2\ell+1)\Big[1 - \eta_\ell\cos(2\mathrm{Re}(\delta_\ell))\Big] \leq \frac{4\pi}{k^2}(2\ell+1). \tag{5.167}$$

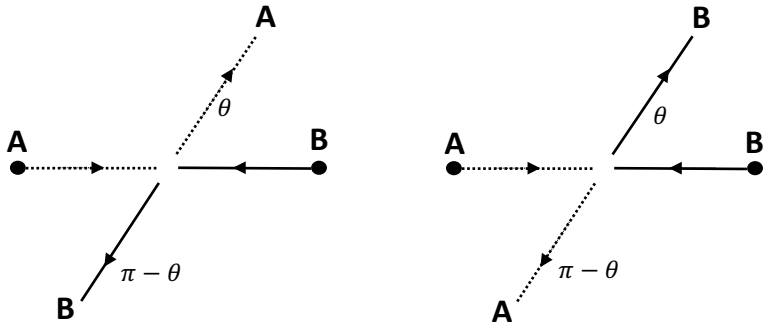


Figure 5.16: Elastic scattering of identical particles. We cannot tell which particle flies away with polar scattering angle θ and which one with $\pi - \theta$. The left and right configurations in this figure are indistinguishable.

Let us collect the three different partial cross-sections for the ℓ wave

$$\sigma_\ell^{\rm el} = \frac{\pi}{k^2}(2\ell+1)\big|1 - S_\ell\big|^2, \tag{5.168}$$

$$\sigma_\ell^{\rm abs} = \frac{\pi}{k^2}(2\ell+1)\Big(1 - |S_\ell|^2\Big), \tag{5.169}$$

$$\sigma_\ell = \frac{2\pi}{k^2}(2\ell+1)\Big[1 - {\rm Re}\,\big(S_\ell\big)\Big]. \tag{5.170}$$

The pure elastic scattering for the ℓ-th partial wave corresponds to $|S_\ell| = \eta_\ell = 1$, in which case $\sigma_\ell^{\rm abs} = 0$. Note that, although the inelastic scattering can be zero, elastic scattering is always present. In other words, we cannot have only inelastic scattering, because this is always accompanied by elastic scattering.

From equations (5.155) and (5.166) it can be seen that the optical theorem takes now the generalised form

$$\sigma_{\rm tot} = \frac{4\pi}{k}{\rm Im}\Big[f_{\rm el}(k, \theta = 0)\Big]. \tag{5.171}$$

This formula relates the total cross-section to the imaginary part of the elastic scattering amplitude in the forward direction.

5.10 Scattering of identical particles.

Let us consider the elastic scattering of two identical particles in the center of mass frame, and assume the cross section depends only on the polar scattering angle θ (central potential). In classical mechanics the cross section would be given by the sum of two terms, $\sigma_{\rm cl}(\theta) + \sigma_{\rm cl}(\pi - \theta)$, but in quantum mechanics it is not possible to distinguish which particle scatters at an angle θ from the one that scatters at $\pi - \theta$ (see Figure 5.16).[20] If the particles are bosons, the scattering amplitude must be symmetric with respect to the interchange of angles θ and $\pi - \theta$,

$$f_{\rm boson}(\theta) = f(\theta) + f(\pi - \theta). \tag{5.172}$$

Therefore the differential cross section is

$$\left.\frac{{\rm d}\sigma}{{\rm d}\Omega}\right|_{\rm boson} = |f(\theta) + f(\pi-\theta)|^2 = |f(\theta)|^2 + |f(\pi-\theta)|^2 + 2\,{\rm Re}\Big[f(\theta)^* f(\pi - \theta)\Big]. \tag{5.173}$$

[20] Don't be confused by the figure in two dimensions, which seems to suggest that one of the scattering angles is $\pi + \theta$. The polar scattering angle is always in the interval $0 \le \theta \le \pi$.

The key difference between this equation and its classical counterpart is the interference term $2\,\mathrm{Re}\Big[f(\theta)^* f(\pi-\theta)\Big]$. In particular, for $\theta=\pi/2$ we have

$$\left.\frac{\mathrm{d}\sigma(\theta=\pi/2)}{\mathrm{d}\Omega}\right|_{\text{boson}} = 4\,|f(\pi/2)|^2\,, \tag{5.174}$$

which is twice as large as the classical equation, since the latter does not have the interference term.

Consider now the elastic scattering of two identical spin 1/2 particles. The total wave function of the system must be antisymmetric, therefore when the two particles are in a spin singlet state (antisymmetric), the scattering amplitude must be symmetric and the differential cross section is given by

$$\frac{\mathrm{d}\sigma_S}{\mathrm{d}\Omega} = |f(\theta)+f(\pi-\theta)|^2 = |f(\theta)|^2+|f(\pi-\theta)|^2+2\,\mathrm{Re}\Big[f(\theta)^* f(\pi-\theta)\Big]. \tag{5.175}$$

However, if the two particles are in a spin triple state (symmetric), the spatial wave function must be antisymmetric and the cross section is

$$\frac{\mathrm{d}\sigma_A}{\mathrm{d}\Omega} = |f(\theta)-f(\pi-\theta)|^2 = |f(\theta)|^2+|f(\pi-\theta)|^2-2\,\mathrm{Re}\Big[f(\theta)^* f(\pi-\theta)\Big]. \tag{5.176}$$

If the incident particles are unpolarized the various spin states are equally likely, so the triplet state will be three times as likely as the singlet, therefore

$$\begin{aligned}\left.\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right|_{\text{fermion}} = \frac{3}{4}\frac{\mathrm{d}\sigma_A}{\mathrm{d}\Omega}+\frac{1}{4}\frac{\mathrm{d}\sigma_S}{\mathrm{d}\Omega} &= \frac{3}{4}\Big|f(\theta)-f(\pi-\theta)\Big|^2+\frac{1}{4}\Big|f(\theta)+f(\pi-\theta)\Big|^2 \\ &= |f(\theta)|^2+|f(\pi-\theta)|^2-\mathrm{Re}\Big[f(\theta)^* f(\pi-\theta)\Big].\end{aligned} \tag{5.177}$$

For $\theta=\pi/2$ this equation gives

$$\left.\frac{\mathrm{d}\sigma(\theta=\pi/2)}{\mathrm{d}\Omega}\right|_{\text{fermion}} = |f(\pi/2)|^2\,, \tag{5.178}$$

which is half the classical expression, $(\mathrm{d}\sigma/\mathrm{d}\Omega)_{\text{cl}} = 2|f(\pi/2)|^2$, and four times smaller than the expression corresponding to the scattering of two identical bosons, $(\mathrm{d}\sigma/\mathrm{d}\Omega)_{\text{boson}} = 4|f(\pi/2)|^2$.

The partial wave analysis of elastic scattering of identical particles can be simplified by noting the following property of the Legendre polynomials

$$P_\ell\big(\cos(\pi-\theta)\big) = P_\ell\big(-\cos\theta\big) = (-1)^\ell P_\ell(\cos\theta), \tag{5.179}$$

which we can insert in Eq. (5.87) to get

$$\begin{aligned}f(\pi-\theta) &= \frac{1}{k}\sum_{\ell=0}^{\infty}(2\ell+1)\,e^{i\delta_\ell}\sin(\delta_\ell)\,P_\ell(\pi-\cos\theta) \\ &= \frac{1}{k}\sum_{\ell=0}^{\infty}(-1)^\ell(2\ell+1)\,e^{i\delta_\ell}\sin(\delta_\ell)P_\ell(\cos\theta).\end{aligned} \tag{5.180}$$

Therefore

$$f(\theta)\pm f(\pi-\theta) = \frac{1}{k}\sum_{\ell=0}^{\infty}\Big[1\pm(-1)^\ell\Big]\,(2\ell+1)\,e^{i\delta_\ell}\sin(\delta_\ell)P_\ell(\cos\theta). \tag{5.181}$$

5.11 Problems

1. **Rutherford scattering.** An incident particle of charge q_1 and kinetic energy E scatters off a heavy stationary particle of charge q_2. Given the formula relating the impact parameter to the scattering angle

$$b = \frac{q_1 q_2}{8\pi\epsilon_0 E}\cot(\theta/2).$$

   (a) Determine the differential scattering cross-section.

   (b) Show that the total cross-section for Rutherford scattering is *infinite*. We say that $1/r$ has *infinite range*; you can not scape from a Coulomb force.

2. In the limit of large r the solution to the Schrödinger equation in a three-dimensional scattering problem has the form

$$\psi(r,\theta) \approx A\left\{e^{ikz} + f(\theta)\frac{e^{ikr}}{r}\right\}, \text{ for large } r.$$

   Construct the analogs of this equation for one-dimensional and two-dimensional scattering.

3. Show that the Wronskian

$$W = j_\ell(x)n'_\ell(x) - n_\ell(x)j'_\ell(x),$$

   where $j_\ell(x)$ and $n_\ell(x)$ are the spherical Bessel and Neumann functions, defined in (5.36), satisfies the differential equation

$$\frac{\mathrm{d}W}{\mathrm{d}x} = -\frac{2}{x}W.$$

   Solve this equation and use the behaviour of $j_\ell(x)$, $n_\ell(x)$ for small x to fix the constant of integration and get

$$W = \frac{1}{x^2}.$$

   The fact that $W \neq 0$ proves that the linear combinations of $j_\ell(kr)$ and $n_\ell(kr)$ are the general solution of the radial equation for a free particle with given k and ℓ.
   Hints:

   - It is helpful to use the differential equations

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{2}{x}\frac{\mathrm{d}}{\mathrm{d}x} + 1 - \frac{\ell(\ell+1)}{x^2}\right] j_\ell(x) = 0,$$

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{2}{x}\frac{\mathrm{d}}{\mathrm{d}x} + 1 - \frac{\ell(\ell+1)}{x^2}\right] n_\ell(x) = 0.$$

   - To determine the constant of integration you might find useful to consider the asymptotic behaviour of the spherical Bessel and Neumann functions for $x \ll 1$

$$j_\ell \to \frac{2^\ell \ell!}{(2\ell+1)!}x^\ell, \qquad n_\ell \to -\frac{(2\ell)!}{2^\ell \ell!}\frac{1}{x^{\ell+1}}.$$

4. Show that the radial contributions to the partial waves in equation (5.58),

$$R_{k\ell}(r) = \frac{u_{k\ell}(r)}{r} = A_\ell \left[j_\ell(kr) + ik\, a_\ell\, h_\ell^{(1)}(kr) \right],$$

that provide the solution to Schrödinger equation outside the scattering region, can also be expressed as

$$R_{k\ell} = B_\ell \left[\cos\delta_\ell\, j_\ell(kr) - \sin\delta_\ell\, n_\ell(kr) \right],$$

where $B_\ell = e^{i\delta_\ell} A_\ell$. Now, consider the asymptotic form of the spherical Bessel and Neumann functions

$$\lim_{x\to\infty} j_\ell(x) = \frac{1}{x}\sin\left(x - \ell\frac{\pi}{2}\right), \qquad \lim_{x\to\infty} n_\ell(x) = -\frac{1}{x}\cos\left(x - \ell\frac{\pi}{2}\right),$$

and deduce that

$$\lim_{r\to\infty} u_{k\ell}(r) \propto \sin\left(kr + \delta_\ell - \ell\frac{\pi}{2}\right).$$

5. Consider the potential

$$V(r) = V_0 \frac{a^2}{r^2}, \qquad V_0 > 0.$$

and take its corresponding radial equation

$$-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2 u}{\mathrm{d}r^2} + \left[V_0 \frac{a^2}{r^2} + \frac{\hbar^2}{2m}\frac{\ell(\ell+1)}{r^2} \right] u = Eu.$$

a) Show that if you propose solutions of the form: $u(r) \propto \sqrt{x}y(x)$, where $x = kr$ and $k^2 = 2mE/(\hbar^2)$, y fulfils the Bessel equation,

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + (x^2 - \alpha^2)y = 0,$$

with

$$\alpha^2 = \frac{2mV_0a^2}{\hbar^2} + \left(\ell + \frac{1}{2}\right)^2.$$

You can find detailed information on the solutions to this equation in https://en.wikipedia.org/wiki/Bessel_function.

b) The solutions to the Schrödinger equation have to be finite at the origin, which in this case means that only Bessel functions of the first kind are admissible. The asymptotic behaviour of these functions for $x \gg |\alpha^2 - 1/4|$ is

$$J_\alpha(x) \approx \sqrt{\frac{2}{\pi x}}\cos\left(x - \alpha\frac{\pi}{2} - \frac{\pi}{4}\right).$$

Use this result and the one derived in problem 4 to conclude that in the scattering of particles by the given potential the phase shifts are

$$\delta_\ell = \left[(\ell + 1/2) - \sqrt{(\ell+1/2)^2 + \frac{2mV_0a^2}{\hbar^2}} \right]\frac{\pi}{2}.$$

c) Show that for a fixed angle the cross section dependence on the energy of the scattered particle is

$$\frac{d\sigma}{d\Omega} \propto \left.\frac{1}{E}\right|_{\theta=\text{const}}.$$

This is a consequence of δ_ℓ being independent on k for this problem potential.

d) Show that, if $2mV_0/\hbar^2 \ll 1$ then

$$\delta_\ell = -\frac{\pi}{2}\frac{mV_0a^2}{\hbar^2\left(\ell+\frac{1}{2}\right)} \ll 1.$$

Use this result and the generating function of Legendre polynomials

$$\sum_{\ell=0}^{\infty} P_\ell(x)t^\ell = \frac{1}{\sqrt{1-2tx+t^2}},$$

to demonstrate that in that regime the differential cross section is

$$\frac{d\sigma}{d\theta} = \frac{\pi^3}{2\hbar^2}\frac{V_0^2ma^4}{E}\cot\left(\frac{\theta}{2}\right).$$

e) For $V(r)$ calculate the differential cross section in the Born approximation and compare it with the result obtained in d). You might need the next definite integral

$$\int_0^\infty \frac{\sin x}{x}dx = \frac{\pi}{2}.$$

6. Scattering of the s and p waves by a hard sphere. We wish to study the phase shifts $\delta_0(k)$ and $\delta_1(k)$ produced by a hard-sphere potential,

$$V(r) = \begin{cases} \infty, & \text{if } r \le R, \\ 0, & \text{if } r > R, \end{cases}$$

on incident s ($\ell = 0$) and p ($\ell = 1$) waves. In particular, we want to verify that $\delta_1(k)$ becomes negligible compared to $\delta_0(k)$ at low energy.

a) Write the radial equation for the function $u_{kl}(r)$ for $r > R$ and consider the particular cases with $\ell = 0, 1$. Show that for $\ell = 0$ the general solution is of the form

$$u_{k0} = B\sin k(r-R),$$

while for $\ell = 1$ the solution is

$$u_{k1} = C\left[\frac{\sin kr}{kr} - \cos kr + a\left(\frac{\cos kr}{kr} + \sin kr\right)\right],$$

where B, C and a are constants.

b) Deduce that $\delta_0(k) = -kR$ and $a = \tan\delta_1(k)$.

Hint: You may wish to use the asymptotic result deduced in problem 4.

c) Determine the constant a from the condition imposed on $u_{k1}(r)$ at $r = R$ and conclude that $\delta_1(k) = \arctan(kR) - kR$.

d) Show that, as k approaches zero, $\delta_1(k)$ behaves like $(kR)^3$, which makes it negligible compared to $\delta_0(k)$. This result is true in general: for any potential of finite range R, the phase shift $\delta_\ell(k)$ behaves like $(kR)^{2\ell+1}$ at low energies.

7. *Square Spherical Well*: Bound States and Scattering Resonances.
Consider the central potential

$$V(r) = \begin{cases} -V_0 & \text{for } r < r_0, \\ 0, & \text{for } r > r_0, \end{cases}$$

where V_0 is a positive constant. Set:

$$k_0 = \sqrt{\frac{2mV_0}{\hbar^2}}.$$

We shall confine ourselves to the study of the s wave ($\ell = 0$).

a) *Bound states* ($E < 0$).

i) Write the radial equation in the two regions $r > r_0$ and $r < r_0$, as well as the condition at the origin. Show that, if one sets

$$\rho = \sqrt{\frac{-2mE}{\hbar^2}} \qquad \text{and} \qquad K = \sqrt{k_0^2 - \rho^2} = \sqrt{\frac{2m}{\hbar^2}(V_0 + E)},$$

the solution of the radial equation, $u_\ell(r)$, takes necessarily the following form for $\ell = 0$,

$$u_0(r) = \begin{cases} Ae^{-\rho r} & \text{for } r > r_0, \\ B \sin Kr & \text{for } r < r_0. \end{cases}$$

ii) Write the matching conditions at $r = r_0$. Deduce from them that the only possible values for ρ are those that satisfy the condition:

$$\tan Kr_0 = -\frac{K}{\rho}.$$

iii) Discuss this equation: indicate the number of bound states, s, as a function of the depth of the well (for fixed r_0) and show, in particular, that there are no bound states if this depth is too small.

b) *Scattering resonances* ($E > 0$).

i) Again write the radial equation, this time setting

$$k = \sqrt{\frac{2mE}{\hbar^2}} \qquad \text{and} \qquad K' = \sqrt{k_0^2 + k^2}.$$

Denoting the solutions as $u_{k\ell}(r)$, show that $u_{k0}(r)$ is of the form:

$$u_{k0}(r) = \begin{cases} A \sin(kr + \delta_0) & \text{for } r > r_0, \\ B \sin K'r & \text{for } r < r_0. \end{cases}$$

ii) Choose $A = 1$. Show, using the continuity conditions at $r = r_0$, that the constant B and the phase shift δ_0 are given by:

$$B^2 = \frac{k^2}{k^2 + k_0^2 \cos^2 K'r_0} \qquad \text{and} \qquad \delta_0 = -kr_0 + \alpha(k),$$

with

$$\tan \alpha(k) = \frac{k}{K'} \tan K'r_0.$$

iii) Plot the curve representing B^2 as a function of k. This curve clearly shows resonances, for which B^2 is a maximum. What are the values of k associated with these resonances? What is then the value of $\alpha(k)$? Show that, if there exists such a resonance for a small energy ($kr_0 \ll 1$), the corresponding contribution of the s wave to the total cross section is practically maximal.

c) *Relation between bound states and scattering resonances.*

Assume that $k_0 r_0$ is very close to $(2n+1)\frac{\pi}{2}$, where n is an integer, and set:

$$k_0 r_0 = (2n+1)\frac{\pi}{2} + \epsilon, \qquad \text{with} \qquad |\epsilon| \ll 1.$$

i) Show that, if ϵ is positive, there exists a bound state whose binding energy $E = -\hbar^2\rho^2/2m$ is given by the condition:

$$\rho \simeq \epsilon k_0$$

ii) Show that if, on the other hand, ϵ is negative, there exists a scattering resonance at energy $E = \hbar^2 k^2/2m$ such that:

$$k^2 \simeq -\frac{2k_0\epsilon}{r_0}.$$

iii) Deduce from this that if the depth of the well is gradually decreased (for fixed r_0), the bound state which disappears when $k_0 r_0$ passes through and odd multiple of $\pi/2$ gives rise to a low energy scattering resonance.

8. Consider the potential given by the spherical delta-function shell

$$V(r) = -V_0\delta(r-a),$$

where V_0 and a are positive constants.

a) Show that the S-wave phase shift satisfies the relation

$$\tan\delta_0 = \frac{\sin^2 ka}{\frac{\hbar^2 k}{2mV_0} - \sin ka\cos ka}.$$

b) Demonstrate that if $\hbar^2/(2m) < V_0 a$ the S-wave presents a resonance. In particular, show that for $V_0 = \hbar^2\pi/(4am)$ the resonance happens at an energy given by

$$E_R = \frac{\hbar^2\pi^2}{32ma^2},$$

and that its lifetime is

$$\tau = \frac{8ma^2}{\hbar\pi^2}.$$

c) In the low energy limit $ka \ll 1$, where only the S-wave has to be considered show that the scattering amplitude, $f(\theta)$, is given by,

$$f(\theta) \approx \frac{a}{\frac{\hbar^2}{2mV_0 a} - 1},$$

and calculate the differential and total cross-sections under this hypothesis.

d) Obtain the scattering amplitude for $V(r)$ in the first Born approximation. Once obtained, extrapolate the result to the low-energy limit, $ka \ll 1$, and show that it coincides with the one obtained in c) if $V_0 a \ll \hbar^2/(2m)$, where no resonances are present.

Hint: Solve exactly the time-independent Schrödinger radial equation for $\ell = 0$ and consider the two cases $r < a$ and $r > a$. By integrating the equation from $r = a - \epsilon$ to $r = a + \epsilon$ (where $0 < \epsilon \ll 1$), show that

$$\left[\left.\frac{du}{dr}\right|_{a+\epsilon} - \left.\frac{du}{dr}\right|_{a-\epsilon} \right] = -\frac{2mV_0}{\hbar^2} u(a).$$

Inserting into this equation the explicit solutions for $u(r)$ in the regions $r < a$ and $r > a$, you should be able to determine the S-wave shift.

9. Consider the following scattering amplitude:

$$f(\theta) = \frac{1}{4k}\left[(2+2i) + 3(\sqrt{3}+i)\cos\theta\right].$$

a) Find the S-wave, P-wave and D-wave amplitudes for this scattering amplitude.

b) Find δ_0 and δ_1, the S- and P-wave phase shifts.

c) Demonstrate that $f(\theta)$ fulfills the optical theorem.

$$\sigma_{\text{tot}} = \frac{4\pi}{k}\text{Im}f(0).$$

10. Neutrons ($m_n c^2 \approx 940\,\text{MeV}$) with kinetic energy $E = 0.2\,\text{MeV}$ are scattered elastically off a repulsive central potential. It is known that the total elastic cross-section is $\sigma = 2.75\,\text{barn}$, and that the s-wave phase shift is $\delta_0 = -0.42\,\text{rad}$. Compute the p-wave phase shift assuming that partial waves with $\ell > 2$ can be neglected.

11. Neutrons with kinetic energy $E = 4\,\text{MeV}$ are scattered by a thin layer of heavy nuclei. Three significant phase shifts are observed: $\delta_0 = 26.1^\text{o}$, $\delta_1 = 7.6^\text{o}$ and $\delta_2 = 0.3^\text{o}$. The neutron flux is $10^{11}\,\text{cm}^{-2}\text{s}^{-1}$, the transversal area of the neutron beam is $A_{\text{be}} = 1\,\text{cm}^2$ and the target density is 10^{21} nuclei cm^{-2} (thin target approximation). Determine:

a) The total cross section.

b) The differential cross section.

c) The average number of neutrons detected by a detector of 10^{-5} sr placed at 90^o with respect to the beam direction.

12. The analysis of the scattering of particles from a potential of characteristic length a shows that the **low-energy** phase shifts are given by

$$\sin(\delta_\ell) = -\frac{\sqrt{\ell}(ak)^{2\ell+1}}{\sqrt{(2\ell+1)}},$$

where $k = \sqrt{2mE}/\hbar$, m is the mass of the particles and E their energy.

a) Prove that the total cross section at low energy is

$$\sigma = 4\pi a^2 \left(\frac{2mEa^2/\hbar^2}{1-(2mEa^2/\hbar^2)^2} \right)^2 . \qquad (1\star)$$

Hint: you will need the following sum, valid for $|x| < 1$,

$$\frac{x}{(1-x)^2} = \sum_{\ell=0}^{\infty} \ell x^\ell .$$

b) Using the phase shifts prove that, up to order $(ak)^4$, the differential cross section is

$$\frac{d\sigma}{d\Omega} = 3a^2 (ak)^4 \cos^2\theta . \qquad (2\star)$$

You will only need to take into account the p wave.

c) Compute the total cross section using equation $(2\star)$. Compare the result with the total cross section given by equation $(1\star)$ and discuss their compatibility.

d) Assume that $(2\star)$ is the appropriate differential cross section for neutrons incident upon a specific target and take $a = 40$ femtometers (fm). A neutron beam of energy 1.3 eV and flux $\mathcal{L} = 10^{36}$ cm^{-2}s^{-1} is aimed at that target. Compute the number of neutrons scattered per second into a cone of $4\pi \times 10^{-3}$ steradians around the incident beam direction.

Data: Neutron mass: $m = 939.6$ MeV/c^2, $\hbar c = 197$ MeV·fm.

13. The analysis of the scattering of particles of mass m and energy E from a potential of characteristic length a shows that the **low energy** cross-section is

$$\sigma = 4\pi a^2 \exp\left(\frac{2a^2 mE}{\hbar^2} \right)^2 = 4\pi a^2 \, e^{(ka)^4} . \qquad (1\star)$$

a) Prove that this cross-section is compatible with the following expression for the phase shifts.

$$\sin(\delta_\ell) = -\frac{(ka)^{2\ell+1}}{\sqrt{(2\ell+1)(\ell!)}} . \qquad (2\star)$$

b) Using those phase-shifts prove that the cross-section up to order $(ka)^2$ is

$$\frac{d\sigma}{d\Omega} = a^2 \left(1 + \sqrt{12}(ka)^2 \cos\theta \right) .$$

You only need to take into account s and p waves.

c) Determine the total cross-section using the result obtained in part b). Compare it with the cross-section given by equation $(1\star)$ and discuss the compatibility of both results.

d) Using the phase-shifts of equation $(2\star)$, prove that the scattering amplitude which accounts for all the partial wave contributions and the cross-section

given in equation (1⋆) verify the optical theorem:

$$\sigma = \frac{4\pi}{k}\mathrm{Im}f(0).$$

Remember that $P_\ell(1) = 1$.

14. The Yukawa potential has the form,

$$V(r) = \beta\frac{e^{-\mu r}}{r},$$

where β and μ are constants. Demonstrate that, in the first Born approximation, the scattering amplitude is given by

$$f(\theta) = -\frac{2m\beta}{\hbar^2(\mu^2 + 4k^2\sin^2(\theta/2))}.$$

By setting $\beta = q_1q_2/(4\pi\epsilon_0)$ and $\mu = 0$ the Yukawa potential becomes Coulomb potential. Check that the differential cross section obtained with these substitutions is precisely Rutherford's formula,

$$\frac{d\sigma}{d\Omega} = \left[\frac{q_1q_2}{16\pi\epsilon_0 E\sin^2(\theta/2)}\right]^2.$$

Although the quoted expression for the scattering amplitude is only valid in the first Born approximation, it is possible to show that higher-order corrections only change $f(\theta)$ by a phase factor when $\mu = 0$ (see [15], section 7.9). Since this has no effect on the differential cross section, it turns out that the Rutherford formula holds even beyond the first Born approximation.

15. Use the Born approximation to determine the total cross-section for scattering from a gaussian potential,

$$V(r) = Ae^{-\mu r^2}.$$

Express your answer in terms of the constants A, μ, and m (the mass of the incident particle), and $k = \sqrt{2mE}/\hbar$, where E is the incident particle energy.

16. Consider the scattering of particles of mass m and energy E by a potential $V(r) = V_0e^{-r/a}$, where V_0 and a have constant value.

    a) Show that in the first Born approximation the scattering amplitude is:

$$f^B(\theta) = -\frac{4mV_0a^3}{\hbar^2(1 + 4a^2k^2\mathrm{sen}^2(\theta/2))^2},$$

    where $k = \sqrt{2mE}/\hbar$.

    b) Employing $f^B(\theta)$ find the S-wave amplitude a_0 and the scattering amplitude considering only this wave.

    c) Compare this last scattering amplitude with $f^B(\theta)$ and discuss the validity of their compatibility.

17. Let's suppose that, in a very simple model, an atom with Z electrons can be viewed as an homogeneous spherical distribution of charge $-Ze$ and radius R, with a point-like nucleus of charge Ze. Consider now an scattering experiment where a beam of electrons of mass m and wave number k is sent onto a target of such simplified atoms.

    - Focus on the scattering of one of the electrons by one of the atoms. Take the origin of coordinates to be at the center of the atom and use the Gauss theorem to prove that the incident electron sees an electrostatic potential energy given by
    $$V(r) = \begin{cases} 0, \text{ if } r > R, \\ -\frac{Ze^2}{8\pi\epsilon_0}\left(\frac{2}{r} - \frac{3}{R} + \frac{r^2}{R^3}\right), \text{ if } r \leq R. \end{cases}$$
    - Prove that the scattering amplitude in the first Born approximation is
    $$f^B(\theta) = \frac{2m}{\hbar^2\kappa}\frac{Ze^2}{4\pi\epsilon_0} R\left[\frac{1}{\kappa R} + 3\frac{\cos(\kappa R)}{(\kappa R)^3} - 3\frac{\sin(\kappa R)}{(\kappa R)^4}\right],$$
    where $\kappa = 2k\sin(\theta/2)$.
    - Prove that for electrons with wavelength much larger than R the previous expression can be approximated as
    $$f^B(\theta) \approx \frac{ZR^2}{5a_0},$$
    where a_0 is the Bohr radius, so that in this case the total cross section is
    $$\sigma \approx 4\pi R^2 \left(\frac{ZR}{5a_0}\right)^2.$$
    Comment on this result for the hydrogen atom taking $Z = 1$ and $R = a_0$.

18. Consider a particle (a) in a bound state $|\varphi_0\rangle$ described by the wave function $\varphi_0(\mathbf{r}_a)$, localized about a point O. Towards this particle is directed a beam of particles (b) of mass m, momentum $\hbar\mathbf{k}_i$, energy $E_i = \hbar\mathbf{k}_i^2/2m$ and wave function $e^{i\mathbf{k}_i\cdot\mathbf{r}_b}/(2\pi)^{3/2}$. Each particle of the beam (b) interacts with the bound particle (a). The potential energy W which describes this interaction depends only on the relative position of the two particles, $\mathbf{r}_b - \mathbf{r}_a$.

    a) Calculate the matrix element
    $$\langle a:\varphi_0;\ b:\mathbf{k}_f|W(\mathbf{r}_b - \mathbf{r}_a)|a:\varphi_0;\ b:\mathbf{k}_i\rangle,$$
    where particle (a) remains in its initial bound state $|\varphi_0\rangle$ and particle (b) transitions from the initial state $|\mathbf{k}_i\rangle$ to the final state $|\mathbf{k}_f\rangle$. The expression for this matrix element should include the Fourier transform $\overline{W}(\mathbf{k})$ of the $W(\mathbf{r}_b - \mathbf{r}_a)$ potential,
    $$W(\mathbf{r}_b - \mathbf{r}_a) = \frac{1}{(2\pi)^{3/2}}\int \overline{W}(\mathbf{k}) e^{i\mathbf{k}\cdot(\mathbf{r}_b - \mathbf{r}_a)} d^3\mathbf{k}.$$

    b) Consider the process in which, under the effect of the interaction W, particle (b) is scattered in a certain direction, with particle (a) remaining in the same

quantum state $|\varphi_0\rangle$ after the scattering process (elastic scattering). Calculate, in the Born approximation, the cross section for the elastic scattering of particle (b) by particle (a) in the state $|\varphi_0\rangle$. You should obtain a result which is equal to the scattering cross section by the potential $W(\mathbf{r})$, multiplied by a term that characterizes the $|\varphi_0\rangle$ bound state, called the *form factor*.

Hint: Start from equation (5.132) in braket notation,

$$f(\theta,\varphi) \approx -\frac{m}{2\pi\hbar^2}\int e^{i(\mathbf{k}_i-\mathbf{k}_f)\cdot\mathbf{r}_0}\,V(\mathbf{r}_0)\,\mathrm{d}^3\mathbf{r}_0 = -\frac{(2\pi)^2 m}{\hbar}\langle\mathbf{k}_f|V(\mathbf{r})|\mathbf{k}_i\rangle,$$

and transform it to get an expression appropriate for the scattering of a free-particle off a bound-particle by simply making the replacement

$$\langle\mathbf{k}_f|V(\mathbf{r})|\mathbf{k}_i\rangle \to \langle a:\varphi_0;\ b:\mathbf{k}_f|W(\mathbf{r}_b-\mathbf{r}_a)|a:\varphi_0;\ b:\mathbf{k}_i\rangle.$$

19. Study the elastic scattering of fast electrons by Hydrogen atoms in the ground state following the next steps.

    (a) The effective (average) potential, $U(r)$, created by a Hydrogen atom in its ground state can be computed using the charge density $e\rho(r) = e\delta(\mathbf{r}) - e|\psi(r)|^2$, where the first term describes the point-like nucleus (at the origin of coordinates) and the second describes the electron *cloud* given by the ground-state wave function $\psi(r) = e^{-r/a_0}/\sqrt{\pi a_0^3}$. Prove that the Poisson equation, $\nabla^2 U = -4\pi e\rho$, takes the following form for $r \neq 0$,

    $$\frac{\mathrm{d}^2\chi}{\mathrm{d}r^2} = \frac{4er}{a_0^3}\,e^{-2r/a_0},$$

    where $\chi(r) = rU(r)$.

    (b) Integrate the previous equation using the boundary conditions $\chi(\infty) = \chi'(\infty) = 0$ to obtain

    $$\chi(r) = \frac{4e}{a_0^3}\int_r^\infty \mathrm{d}r_1 \int_{r_1}^\infty r_2\,e^{-2r_2/a_0}\,\mathrm{d}r_2.$$

    From this you should be able to get

    $$U(r) = \frac{\chi(r)}{r} = e\left(\frac{1}{r}+\frac{1}{a_0}\right)e^{-2r/a_0}.$$

    Study how $U(r)$ behaves for $r \to 0$ and $r \to \infty$. Explain, in physical terms, why the asymptotic expressions you have obtained make sense.

    (c) Use the first Born approximation to compute the scattering amplitude, $f(\theta)$, which describes the elastic scattering of fast electrons by Hydrogen atoms. You should get

    $$f(\kappa) = 2a_0\,\frac{8+\kappa^2a_0^2}{(4+\kappa^2a_0^2)^2},$$

    where $\kappa = 2k\sin(\theta/2)$.

    (d) Prove that the total scattering cross section is

    $$\sigma = \frac{7\pi}{3k^2}\left[1-\frac{3k^4a_0^4+9k^2a_0^2+7}{7(1+k^2a_0^2)^3}\right].$$

    This shows that in the limiting case of $ka_0 \gg 1$ the cross section does not depend on the value of a_0.

e) Compute the cross section, in barns, for $10\,\mathrm{MeV}/c$ electrons.

20. A particle is scattered by a completely absorptive (*black*) sphere of radius a which is large compared to the de Broglie wavelength $\lambda = 2\pi/k$ of the incident particle. Following a semiclassical reasoning, we would expect that if the particle approaches the black sphere with an impact parameter $b > a$, no interaction (scattering) takes place. On the other hand, for $b < a$ the particle is absorbed by the sphere. According to this, compute the total elastic cross-section and the total inelastic cross-section (you may use the approximation $\sum_{n=0}^{N}(2n+1) \approx N^2$ for large N). Give the numerical values of the three cross-sections for $a = 1$ fm.

21. Let $\Psi(\mathbf{r},t) = \psi(\mathbf{r})\exp(-iEt/\hbar)$ be a stationary-state solution of the Schrödinger equation with a real potential V. The current density is

$$\mathbf{j}(\mathbf{r},t) = \frac{\hbar}{2mi}\left[\Psi^*(\boldsymbol{\nabla}\Psi) - (\boldsymbol{\nabla}\Psi^*)\Psi\right] = \frac{\hbar}{2mi}\left[\psi^*(\boldsymbol{\nabla}\psi) - (\boldsymbol{\nabla}\Psi^*)\psi\right].$$

Prove that the divergence of this current vanishes,

$$\boldsymbol{\nabla}\mathbf{j} = \frac{\hbar}{2mi}\boldsymbol{\nabla}\left[\psi^*(\boldsymbol{\nabla}\psi) - (\boldsymbol{\nabla}\psi^*)\psi\right] = 0.$$

Using Gauss's theorem, $\int \boldsymbol{\nabla}\mathbf{j}\,\mathrm{d}^3\mathbf{r} = \oint \mathbf{j}\,\mathrm{d}^2\mathbf{s}$, prove that the following integral is zero for any value of the radius r,

$$0 = r^2\int_0^{\pi}\sin\theta\,\mathrm{d}\theta\int_0^{2\pi}\left(\psi^*\frac{\partial\psi}{\partial r} - \frac{\partial\psi^*}{\partial r}\psi\right)\mathrm{d}\varphi.$$

Next, take the limit of r very large and use the asymptotic form $\psi = A\Big(e^{ikz} + f(\theta)\,(e^{ikr}/r)\Big)$ in the previous equation to derive the optical theorem.

Hints: Use the gradient operator in spherical coordinates. When computing the integrals you will encounter terms containing the exponential factor $e^{\pm ikr(1-\cos\theta)}$, like

$$\int_0^{\pi}\int_0^{2\pi} e^{ikr(1-\cos\theta)}g(\theta)\,\sin\theta\,\mathrm{d}\theta\mathrm{d}\varphi \;=\; 2\pi\int_{-1}^{+1} g(\omega)e^{ikr(1-\omega)}\,\mathrm{d}\omega$$

where $g(\omega)$ is any smooth function of $\omega \equiv \cos\theta$. Use repeatedly the technique of integration by parts in the previous expression to get an asymptotic expansion in powers of $1/r$. Since we are only interested in the limit $r \to \infty$, all terms except the first one in the expansion can be discarded, which then can be taken to be the result of the integral. You may read also pages 133 and 134 of Schiff's textbook on Quantum Mechanics (McGraw-Hill, 1968).

Appendix A

Additional problems.

A.1 Chapter 1.

1. Let $\Psi(\mathbf{r},t)$ be a solution of the (time-dependent) Schrödinger equation for a free particle of charge q,

$$\frac{\mathbf{p}^2}{2m}\Psi = i\hbar\frac{\partial\Psi}{\partial t}.$$

Check that if we make a global change on the phase of Ψ to get $\Psi'(\mathbf{r},t) = e^{iq\alpha/\hbar}\,\Psi(\mathbf{r},t)$, where α is a constant, the new $\Psi'(\mathbf{r},t)$ is still a solution of the same Schrödinger equation. Allow now α to be a function of space and time, so that Ψ transforms as

$$\Psi(\mathbf{r},t) \;\longrightarrow\; \Psi'(\mathbf{r},t) = e^{iq\alpha(\mathbf{r},t)/\hbar}\,\Psi(\mathbf{r},t).$$

Check that this new $\Psi'(\mathbf{r},t)$ is no longer a solution of the Schrödinger equation stated before. It may seem strange to you that Nature should be limited to accept the same phase at all locations and at all times. Perhaps we should ask ourselves how to modify the free particle Schrödinger equation to make it invariant under *local gauge transformations*. The remarkable answer to this question is that we need to introduce an electromagnetic field that interacts with charged particles precisely as described by the Maxwell equations. Consider the Schrödinger equation given by

$$\left[\frac{1}{2m}\left(\mathbf{p}-\frac{q}{c}\mathbf{A}\right)^2 + q\phi\right]\Psi = i\hbar\frac{\partial\Psi}{\partial t},$$

where $\mathbf{A}$ is a vector field and ϕ is a scalar field. Prove that if Ψ is a solution to this equation, then $\Psi'(\mathbf{r},t) = e^{iq\alpha(\mathbf{r},t)/\hbar}\,\Psi(\mathbf{r},t)$ is a solution of

$$\left[\frac{1}{2m}\left(\mathbf{p}-\frac{q}{c}\mathbf{A}'\right)^2 + q\phi'\right]\Psi' = i\hbar\frac{\partial\Psi'}{\partial t},$$

provided the vector field $\mathbf{A}$ and the scalar field ϕ transform as

$$\mathbf{A} \;\longrightarrow\; \mathbf{A}' = \mathbf{A} + \frac{c}{q}\boldsymbol{\nabla}\alpha, \qquad \phi \;\longrightarrow\; \phi' = \phi - \frac{\partial\alpha}{\partial t}.$$

In a sense, Maxwell electrodynamics could be *rediscovered* by demanding our theory to be invariant under space-time-dependent phase transformations. By invoking this *local gauge invariance principle*, properly generalized and implemented in quantum field theory, is how the electroweak and the strong interactions *emerge* in the Standard Model of particle physics.

Hint: the main burden is to prove

$$\left(-i\hbar\boldsymbol{\nabla}-\frac{q}{c}\mathbf{A}-\boldsymbol{\nabla}\alpha\right)^2(e^{iq\alpha/\hbar}\,\Psi)=e^{iq\alpha/\hbar}\left(-i\hbar\boldsymbol{\nabla}-\frac{q}{c}\mathbf{A}\right)^2\Psi.$$

You need to handle carefully the operators. Note, for example, that by applying the chain rule for differentiation you get $(i\hbar\boldsymbol{\nabla})\mathbf{A}\Psi=i\hbar(\boldsymbol{\nabla}\mathbf{A})\Psi+i\hbar\mathbf{A}(\boldsymbol{\nabla}\Psi)$.

Solution: (Disclaimer: some $\hbar$ factors are missing in this solution, since it was first written assuming $\hbar=1$ and the modifications to make the $\hbar$'s explicit are not finished). What you need to prove is that the following equation holds

$$\left[\frac{1}{2m}\left(\mathbf{p}-\frac{q}{c}\mathbf{A}-\boldsymbol{\nabla}\alpha\right)^2+q\phi-q\frac{\partial\alpha}{\partial t}\right]e^{iq\alpha/\hbar}\Psi=i\hbar\frac{\partial}{\partial t}(e^{iq\alpha/\hbar}\Psi).$$

$$\left[\frac{1}{2m}\left(-i\hbar\boldsymbol{\nabla}-\frac{q}{c}\mathbf{A}-\boldsymbol{\nabla}\alpha\right)^2+q\phi-q\frac{\partial\alpha}{\partial t}\right]e^{iq\alpha/\hbar}\Psi=-q\left(\frac{\partial\alpha}{\partial t}\right)e^{iq\alpha/\hbar}\,\Psi+i\hbar\,e^{iq\alpha/\hbar}\left(\frac{\partial\Psi}{\partial t}\right).$$

$$\frac{1}{2m}\left(-i\hbar\boldsymbol{\nabla}-\frac{q}{c}\mathbf{A}-\boldsymbol{\nabla}\alpha\right)^2(e^{iq\alpha/\hbar}\Psi)+q\phi e^{iq\alpha/\hbar}\Psi=i\hbar\,e^{iq\alpha/\hbar}\frac{\partial\Psi}{\partial t}. \tag{A.1}$$

The terms containing the time derivative of α cancel each other out. Now

$$\begin{aligned}\left(-i\hbar\boldsymbol{\nabla}-\frac{q}{c}\mathbf{A}-\boldsymbol{\nabla}\alpha\right)^2&=\left(-i\hbar\boldsymbol{\nabla}-\frac{q}{c}\mathbf{A}-\boldsymbol{\nabla}\alpha\right)\left(-i\hbar\boldsymbol{\nabla}-\frac{q}{c}\mathbf{A}-\boldsymbol{\nabla}\alpha\right)\\&=-\hbar^2\nabla^2+i\hbar\frac{q}{c}(\boldsymbol{\nabla}\mathbf{A})+i\hbar\frac{q}{c}\mathbf{A}\boldsymbol{\nabla}+i\hbar(\nabla^2\alpha)+i\hbar(\boldsymbol{\nabla}\alpha)\boldsymbol{\nabla}\\&\quad+i\hbar\frac{q}{c}\mathbf{A}\boldsymbol{\nabla}+\frac{q^2}{c^2}\mathbf{A}^2+\frac{q}{c}\mathbf{A}(\boldsymbol{\nabla}\alpha)\\&\quad+i\hbar(\boldsymbol{\nabla}\alpha)\boldsymbol{\nabla}+\frac{q}{c}(\boldsymbol{\nabla}\alpha)\mathbf{A}+(\boldsymbol{\nabla}\alpha)^2.\end{aligned}$$

$$\begin{aligned}\left(-i\hbar\boldsymbol{\nabla}-\frac{q}{c}\mathbf{A}-\boldsymbol{\nabla}\alpha\right)^2&=-\hbar^2\nabla^2+\frac{q^2}{c^2}\mathbf{A}^2+(\boldsymbol{\nabla}\alpha)^2\\&\quad+2i\hbar\frac{q}{c}\mathbf{A}\boldsymbol{\nabla}+2i\hbar\frac{q}{c}(\boldsymbol{\nabla}\alpha)\boldsymbol{\nabla}+2\frac{q}{c}(\boldsymbol{\nabla}\alpha)\mathbf{A}+i\hbar\frac{q}{c}(\boldsymbol{\nabla}\mathbf{A})+i\hbar(\nabla^2\alpha).\end{aligned}$$

Before applying this operator to $\Psi'=e^{iq\alpha/\hbar}\Psi$ it is convenient to evaluate separately several parts. First, note that

$$\nabla^2(\Psi_1\Psi_2)=(\nabla^2\Psi_1)\Psi_2+\Psi_1(\nabla^2\Psi_2)+2(\boldsymbol{\nabla}\Psi_1)(\boldsymbol{\nabla}\Psi_2),$$

and

$$\nabla^2e^{iq\alpha/\hbar}=\boldsymbol{\nabla}(\boldsymbol{\nabla}e^{iq\alpha/\hbar})=\boldsymbol{\nabla}\left(\frac{iq}{\hbar}(\boldsymbol{\nabla}\alpha)e^{iq\alpha/\hbar}\right)=\frac{iq}{\hbar}(\nabla^2\alpha)e^{iq\alpha/\hbar}-\frac{q^2}{\hbar^2}(\boldsymbol{\nabla}\alpha)^2e^{iq\alpha/\hbar}$$

Therefore

$$\begin{aligned}-\hbar^2\nabla^2(e^{iq\alpha/\hbar}\Psi)&=-\hbar^2(\nabla^2e^{iq\alpha/\hbar})\Psi-\hbar^2e^{iq\alpha/\hbar}(\nabla^2\Psi)-2\hbar^2(\boldsymbol{\nabla}e^{iq\alpha/\hbar})(\boldsymbol{\nabla}\Psi)\\&=-iq\hbar^2(\nabla^2\alpha)e^{iq\alpha/\hbar}\Psi+q^2\hbar^2(\boldsymbol{\nabla}\alpha)^2e^{iq\alpha/\hbar}\Psi-\hbar^2e^{iq\alpha}(\nabla^2\Psi)-2iq\hbar^2(\boldsymbol{\nabla}\alpha)e^{iq\alpha/\hbar}(\boldsymbol{\nabla}\Psi)\end{aligned}$$

Also

$$\boldsymbol{\nabla}(e^{iq\alpha}\Psi)=(\boldsymbol{\nabla}e^{iq\alpha})\Psi+e^{iq\alpha}\boldsymbol{\nabla}\Psi=iq(\boldsymbol{\nabla}\alpha)e^{iq\alpha}\Psi+e^{iq\alpha}\boldsymbol{\nabla}\Psi$$

We are now ready to compute $(-i\hbar\boldsymbol{\nabla}-q\mathbf{A}-q\boldsymbol{\nabla}\alpha)^2e^{iq\alpha}\Psi$

$$\begin{aligned}(-i\hbar\boldsymbol{\nabla}-q\mathbf{A}-q\boldsymbol{\nabla}\alpha)^2(e^{iq\alpha}\Psi)&=-\hbar^2\nabla^2(e^{iq\alpha}\Psi)+q^2\mathbf{A}^2(e^{iq\alpha}\Psi)+q^2(\boldsymbol{\nabla}\alpha)^2(e^{iq\alpha}\Psi)\\&\quad+2iq\hbar\mathbf{A}\boldsymbol{\nabla}(e^{iq\alpha}\Psi)+2iq\hbar(\boldsymbol{\nabla}\alpha)\boldsymbol{\nabla}(e^{iq\alpha}\Psi)+2q^2(\boldsymbol{\nabla}\alpha)\mathbf{A}(e^{iq\alpha}\Psi)\\&\quad+iq\hbar(\boldsymbol{\nabla}\mathbf{A})(e^{iq\alpha}\Psi)+iq\hbar(\nabla^2\alpha)(e^{iq\alpha}\Psi)\end{aligned}$$

$$
\begin{aligned}
(-i\hbar\boldsymbol{\nabla} - q\mathbf{A} - q\boldsymbol{\nabla}\alpha)^2(e^{iq\alpha}\Psi) = {} & -iq\hbar^2(\nabla^2\alpha)e^{iq\alpha}\Psi + q^2\hbar^2(\boldsymbol{\nabla}\alpha)^2 e^{iq\alpha}\Psi \\
& - \hbar^2 e^{iq\alpha}(\nabla^2\Psi) - 2iq\hbar^2(\boldsymbol{\nabla}\alpha)e^{iq\alpha}(\boldsymbol{\nabla}\Psi) + q^2\mathbf{A}^2(e^{iq\alpha}\Psi) + q^2(\boldsymbol{\nabla}\alpha)^2(e^{iq\alpha}\Psi) \\
& - 2q^2\hbar\mathbf{A}(\boldsymbol{\nabla}\alpha)e^{iq\alpha}\Psi + 2iq\hbar\mathbf{A}e^{iq\alpha}\boldsymbol{\nabla}\Psi - 2q^2\hbar(\boldsymbol{\nabla}\alpha)^2 e^{iq\alpha}\Psi + 2iq\hbar(\boldsymbol{\nabla}\alpha)e^{iq\alpha}\boldsymbol{\nabla}\Psi \\
& + 2q^2(\boldsymbol{\nabla}\alpha)\mathbf{A}(e^{iq\alpha}\Psi) + iq\hbar(\boldsymbol{\nabla}\mathbf{A})(e^{iq\alpha}\Psi) + iq\hbar(\nabla^2\alpha)(e^{iq\alpha}\Psi)
\end{aligned}
$$

$$
\begin{aligned}
e^{-iq\alpha}(-i\hbar\boldsymbol{\nabla} - q\mathbf{A} - q\boldsymbol{\nabla}\alpha)^2(e^{iq\alpha}\Psi) = {} & -iq\hbar^2(\nabla^2\alpha)\Psi + q^2\hbar^2(\boldsymbol{\nabla}\alpha)^2\Psi \\
& - \hbar^2(\nabla^2\Psi) - 2iq\hbar^2(\boldsymbol{\nabla}\alpha)(\boldsymbol{\nabla}\Psi) + q^2\mathbf{A}^2(\Psi) + q^2(\boldsymbol{\nabla}\alpha)^2(\Psi) - 2q^2\hbar\mathbf{A}(\boldsymbol{\nabla}\alpha)\Psi \\
& + 2iq\hbar\mathbf{A}\boldsymbol{\nabla}\Psi - 2q^2\hbar(\boldsymbol{\nabla}\alpha)^2\Psi + 2iq\hbar(\boldsymbol{\nabla}\alpha)\boldsymbol{\nabla}\Psi + 2q^2(\boldsymbol{\nabla}\alpha)\mathbf{A}(\Psi) \\
& + iq\hbar(\boldsymbol{\nabla}\mathbf{A})(\Psi) + iq\hbar(\nabla^2\alpha)(\Psi)
\end{aligned}
$$

Reordering and painting with the same color terms which cancel each other out we get

$$
\begin{aligned}
e^{-iq\alpha}(-i\hbar\boldsymbol{\nabla} - q\mathbf{A} - q\boldsymbol{\nabla}\alpha)^2(e^{iq\alpha}\Psi) = {} & \\
& - \hbar^2(\nabla^2\Psi) + iq\hbar(\boldsymbol{\nabla}\mathbf{A})(\Psi) + 2iq\hbar\mathbf{A}(\boldsymbol{\nabla}\Psi) + q^2\mathbf{A}^2(\Psi) \\
& -iq\hbar^2(\nabla^2\alpha)\Psi + q^2\hbar^2(\boldsymbol{\nabla}\alpha)^2\Psi - 2iq\hbar^2(\boldsymbol{\nabla}\alpha)(\boldsymbol{\nabla}\Psi) + q^2(\boldsymbol{\nabla}\alpha)^2(\Psi) - 2q^2\hbar\mathbf{A}(\boldsymbol{\nabla}\alpha)\Psi \\
& - 2q^2\hbar(\boldsymbol{\nabla}\alpha)^2\Psi + 2iq\hbar(\boldsymbol{\nabla}\alpha)\boldsymbol{\nabla}\Psi + 2q^2(\boldsymbol{\nabla}\alpha)\mathbf{A}\Psi + iq\hbar(\nabla^2\alpha)\Psi
\end{aligned}
$$

$$
\begin{aligned}
(-i\hbar\boldsymbol{\nabla} - q\mathbf{A} - q\boldsymbol{\nabla}\alpha)^2(e^{iq\alpha}\Psi) = {} & \\
& e^{iq\alpha}\Big[- \hbar^2(\nabla^2\Psi) + iq\hbar(\boldsymbol{\nabla}\mathbf{A})(\Psi) + 2iq\hbar\mathbf{A}(\boldsymbol{\nabla}\Psi) + q^2\mathbf{A}^2(\Psi)\Big]
\end{aligned}
$$

But

$$
\begin{aligned}
(-i\hbar\boldsymbol{\nabla} - q\mathbf{A})^2 &= -\hbar^2\nabla^2 + iq\hbar(\boldsymbol{\nabla}\mathbf{A}) + iq\hbar\mathbf{A}\boldsymbol{\nabla} + iq\hbar\mathbf{A}\boldsymbol{\nabla} + q^2\mathbf{A}^2. \\
(-i\hbar\boldsymbol{\nabla} - q\mathbf{A})^2\Psi &= -\hbar^2\nabla^2\Psi + iq\hbar(\boldsymbol{\nabla}\mathbf{A})\Psi + 2iq\hbar\mathbf{A}(\boldsymbol{\nabla}\Psi) + q^2\mathbf{A}^2\Psi.
\end{aligned}
$$

Therefore

$$
\big(- i\hbar\boldsymbol{\nabla} - q\mathbf{A} - q\boldsymbol{\nabla}\alpha\big)^2(e^{iq\alpha}\Psi) = e^{iq\alpha}\big(- i\hbar\boldsymbol{\nabla} - q\mathbf{A}\big)^2\Psi
$$

Taking this into account, the Schrödinger equation (A.1) transforms into

$$
\frac{1}{2m}e^{iq\alpha}\big(- i\hbar\boldsymbol{\nabla} - q\mathbf{A}\big)^2\Psi + q\phi e^{iq\alpha}\Psi = i\hbar\, e^{iq\alpha}\frac{\partial\Psi}{\partial t}.
$$

Now we can eliminate the $e^{iq\alpha}$ phase to get

$$
\frac{1}{2m}\big(- i\hbar\boldsymbol{\nabla} - q\mathbf{A}\big)^2\Psi + q\phi\Psi = i\hbar\frac{\partial\Psi}{\partial t},
$$

which is the original assumption we started from. What follows is a little extra material that might be useful to avoid easy mistakes.

It is usual to define the so-called *covariant* derivatives as (we set now $\hbar = 1$)

$$
\mathbf{D} = \boldsymbol{\nabla} - iq\mathbf{A}, \qquad D^0 = \frac{\partial}{\partial t} + iq\phi.
$$

The Schrödinger equation containing interaction terms with the electromagnetic field can be written as

$$
\frac{1}{2m}(-i\mathbf{D})^2\Psi = iD^0\Psi.
$$

It is not difficult to show that

$$
\begin{aligned}
\mathbf{D}'\Psi' &= e^{iq\alpha}\mathbf{D}\Psi, \qquad & D^{0\prime}\Psi' &= e^{iq\alpha}D^0\Psi. \\
(\mathbf{D}')^2\Psi' &= e^{iq\alpha}\mathbf{D}^2\Psi, \qquad & (D^{0\prime})^2\Psi' &= e^{iq\alpha}(D^0)^2\Psi.
\end{aligned}
$$

However, be careful to note that the extra $iq\mathbf{A}$ term in the covariant derivative spoils the ordinary chain derivative rule that one might be tempted to use

$$
\mathbf{D}(\Psi_1\Psi_2) \neq (\mathbf{D}\Psi_1)\Psi_2 + \Psi_1(\mathbf{D}\Psi_2).
$$

The correct expression is

$$\begin{aligned}\mathbf{D}(\Psi_1\Psi_2) = (\boldsymbol{\nabla} - iq\mathbf{A})(\Psi_1\Psi_2) &= \boldsymbol{\nabla}(\Psi_1\Psi_2) - iq\mathbf{A}\Psi_1\Psi_2 \\ &= (\boldsymbol{\nabla}\Psi_1)\Psi_2 + \Psi_1(\boldsymbol{\nabla}\Psi_2) - iq\mathbf{A}\Psi_1\Psi_2 \\ &= \underbrace{(\boldsymbol{\nabla}\Psi_1)\Psi_2 - iq\mathbf{A}\Psi_1\Psi_2}_{(\mathbf{D}\Psi_1)\Psi_2} + \underbrace{\Psi_1(\boldsymbol{\nabla}\Psi_2) - iq\mathbf{A}\Psi_1\Psi_2}_{\Psi_1(\mathbf{D}\Psi_2)} + iq\mathbf{A}\Psi_1\Psi_2\end{aligned}$$

So

$$\mathbf{D}(\Psi_1\Psi_2) = (\mathbf{D}\Psi_1)\Psi_2 + \Psi_1(\mathbf{D}\Psi_2) + iq\mathbf{A}\Psi_1\Psi_2 \tag{A.2}$$

2. In this problem you are going to try to derive the Hamiltonian for the spin-orbit interaction in the Hydrogen atom using classical electrodynamics. Start by assuming that the electron and the proton can be treated as classical point charged particles describing an orbital circular motion of radius r around each other. In the rest frame of the electron, the proton looks like a circular current of intensity $I = e/T$, where T is the orbital period. According to the Biot–Savart law, the magnetic field generated by such a circular current is $B = \mu_0 I/2r = \mu_0 e/2rT$. By noting that the orbital angular momentum of the electron is $L = rmv = 2\pi r^2/T$, express the magnetic field B in terms of L. Next, consider the electron as if it were a tiny circulating current around a ring of mass m and radius r_e. The magnetic moment of this model can then be written as $\mu = q\pi r_e^2/T_e$, with T_e the rotation period and q the charge of the electron smeared over the ring. From classical mechanics we know that T_e can be related to the electron spinning angular momentum using its moment of inertia, $S = 2\pi I_e/T_e$, ($I_e = mr_e^2$ for a ring). Write then μ in terms of S and multiply it by the magnetic field B to get $\mathcal{H}_{\text{so}} = -\mu B$. Compare your result with equation (1.159) and comment on the differences, if any. (Note that $1/c^2 = \epsilon_0\mu_0$).

Solution: In the electron rest frame the proton could be *pictured* as a continuous circular current loop. According to the Biot-Savart law, the magnetic field generated by such a current is

$$B = \frac{\mu_0 I}{2r} = \frac{\mu_0 e}{2rT} = \frac{1}{2\epsilon_0}\frac{e}{c^2 rT} = \frac{1}{4\pi\epsilon_0}\frac{ev}{c^2 r^2},$$

where $\epsilon_0\mu_0 = 1/c^2$ and the effective current is given by $I = e/T$, being $T = 2\pi r/v$ the period of the orbit and r its radius. We can write the previous expression in terms of the orbital angular momentum $\mathbf{L}$ *of the electron* (in the rest frame of the nucleus) by noting that its classical velocity would be $v = 2\pi r/T$ and $L = rmv = 2\pi mr^2/T$. Furthermore, $\mathbf{B}$ and $\mathbf{L}$ point in the same direction, therefore

$$\mathbf{B} = \frac{1}{4\pi\epsilon_0}\frac{e}{mc^2 r^3}\mathbf{L}.$$

According to classical electrodynamics, a spinning charge creates a magnetic dipole moment proportional to its (spin) angular momentum. Let's see how this comes out for a charge q smeared over a ring of radius r_e and mass m which rotates with a period T_e. In this situation, the magnetic moment is defined in classical electrodynamics as the current (q/T_e) multiplied by the area πr_e^2,

$$\mu = \frac{q\pi r_e^2}{T_e}.$$

From classical mechanics we know that the (spin) angular moment of the ring is obtained by multiplying its moment of inertia (mr_e^2) by its angular velocity $(2\pi/T_e)$,

$$S = \frac{2\pi}{T_e}mr_e^2, \qquad \text{therefore} \qquad \mu = \frac{q}{2m}S, \qquad \text{or} \qquad \boldsymbol{\mu} = \frac{q}{2m}\mathbf{S}.$$

The gyromagnetic ratio for this simple classical model comes out then as $\mu/S = q/2m$, which is independent of r_e and T_e. The preceding reasoning is valid for any charge distribution having a shape given by a figure of revolution. The electron magnetic moment in this model would be then

$$\boldsymbol{\mu}_e = -\frac{e}{2m}\mathbf{S}.$$

Putting everything together, we get the following Hamiltonian for the spin-orbit interaction

$$\mathcal{H} = -\boldsymbol{\mu}\mathbf{B} = \frac{e^2}{4\pi\epsilon_0}\frac{1}{2m^2c^2r^3}\,\mathbf{LS}, \qquad \text{or, in Gaussian units} \qquad \mathcal{H} = \frac{e^2}{2m^2c^2r^3}\,\mathbf{LS}.$$

This is the same result we got from the Dirac equation. However, all our classical reasoning has been made in the rest frame of the electron, which is not an inertial system. This means that we should have applied the appropriate kinematic correction originating from the effect known as *Thomas precession.*[1] It can be shown that in this case the Thomas precession gives an extra factor of 1/2. With this factor, the Hamiltonian for the spin-orbit interaction would be

$$\mathcal{H} = \frac{e^2}{4m^2c^2r^3}\,\mathbf{LS},$$

a result that is half the correct value. Of course, you have probably noticed that by modeling the electron as a classical tiny current loop gives only half of its true magnetic moment, which is $\boldsymbol{\mu}_e = -\frac{e}{m}\mathbf{S}$. Using this expression we would get the correct Hamiltonian for the spin-orbit interaction. It is interesting to note that, coincidentally, the extra factor of 1/2 due to the Thomas precession cancels the factor 2 in the (quantum) modified gyromagnetic ratio for the electron.

3. We have computed the relativistic correction to the energy levels of hydrogenic atoms using the next-to-leading term (the second) in the following power expansion,

$$T = \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \frac{p^6}{16m^5c^4} + \dots$$

Compute now the correction due to the third term, but do it only for the hydrogen ground state. Give also the ratio of the third-term to the second-term correction. Hint: write $p^6/(16m^5c^4)$ using $\mathcal{H}_0 = p^2/(2m) - Ze^2/r$ and proceed as we did when computing the correction due to the second term. At a given point you will need to compute the matrix element $\left\langle \psi \,\middle|\, \frac{1}{r}p^2\frac{1}{r} \,\middle|\, \psi \right\rangle$, where ψ is the hydrogen ground state.

Solution: Let's denote

$$\mathcal{H}_r^{(3)} = \frac{p^6}{16m^5c^4} = \frac{1}{2m^2c^4}\left(\frac{p^2}{2m}\right)^3.$$

Using $\mathcal{H}_0 = \dfrac{p^2}{2m} - \dfrac{Ze^2}{r}$ we may write,

$$\mathcal{H}_r^{(3)} = \frac{1}{2m^2c^4}\left(\mathcal{H}_0 + \frac{Ze^2}{r}\right)^3 = \frac{1}{2m^2c^4}\left[\mathcal{H}_0^2 + Ze^2\left(\mathcal{H}_0\frac{1}{r} + \frac{1}{r}\mathcal{H}_0\right) + \frac{(Ze^2)^2}{r^2}\right]\left(\mathcal{H}_0 + \frac{Ze^2}{r}\right).$$

$$2m^2c^4\mathcal{H}_r^{(3)} = \mathcal{H}_0^3 + Ze^2\mathcal{H}_0\frac{1}{r}\mathcal{H}_0 + Ze^2\frac{1}{r}\mathcal{H}_0^2 + \frac{(Ze^2)^2}{r^2}\mathcal{H}_0 + \mathcal{H}_0^2\frac{Ze^2}{r} + Ze^2\mathcal{H}_0\frac{1}{r}\frac{Ze^2}{r} + Ze^2\frac{1}{r}\mathcal{H}_0\frac{Ze^2}{r} + \frac{(Ze^2)^3}{r^3}.$$

$$2m^2c^4\mathcal{H}_r^{(3)} = \mathcal{H}_0^3 + Ze^2\left(\mathcal{H}_0\frac{1}{r}\mathcal{H}_0\right) + Ze^2\left(\frac{1}{r}\mathcal{H}_0^2\right) + (Ze^2)^2\left(\frac{1}{r^2}\mathcal{H}_0\right) + Ze^2\left(\mathcal{H}_0^2\frac{1}{r}\right) + (Ze^2)^2\left(\mathcal{H}_0\frac{1}{r^2}\right) + (Ze^2)^2\left(\frac{1}{r}\mathcal{H}_0\frac{1}{r}\right) + \frac{(Ze^2)^3}{r^3}.$$

Since $\mathcal{H}_0$ is hermitian,

$$\left\langle \frac{1}{r}\mathcal{H}_0^2 \right\rangle = \left\langle \mathcal{H}_0^2\frac{1}{r} \right\rangle = \left(E_n^{(0)}\right)^2\left\langle \frac{1}{r} \right\rangle, \qquad \left\langle \frac{1}{r^2}\mathcal{H}_0^2 \right\rangle = \left\langle \mathcal{H}_0^2\frac{1}{r^2} \right\rangle = \left(E_n^{(0)}\right)^2\left\langle \frac{1}{r^2} \right\rangle,$$

[1] The Thomas precession is a kinematic effect that arises when making calculations in a non-inertial (accelerated) system. It is due to the fact that the product of two relativistic boosts with non-parallel velocities is equivalent, rather surprisingly, to a single boost plus a rotation.

$$\text{and} \qquad \left\langle \mathcal{H}_0 \frac{1}{r} \mathcal{H}_0 \right\rangle = \left(E_n^{(0)}\right)^2 \left\langle \frac{1}{r} \right\rangle .$$

The expectation value $\left\langle \frac{1}{r} \mathcal{H}_0 \frac{1}{r} \right\rangle$ is more difficult to handle. Still, some simplification can be done since

$$\left\langle \frac{1}{r} \mathcal{H}_0 \frac{1}{r} \right\rangle = \left\langle \frac{1}{r} \left(\frac{p^2}{2m} - \frac{Ze^2}{r} \right) \frac{1}{r} \right\rangle = \frac{1}{2m} \left\langle \frac{1}{r} p^2 \frac{1}{r} \right\rangle - Ze^2 \left\langle \frac{1}{r^3} \right\rangle .$$

So,

$$2m^2c^4 \left\langle \mathcal{H}_r^{(3)} \right\rangle = \left(E_n^{(0)}\right)^3 + 3Ze^2 \left(E_n^{(0)}\right)^2 \left\langle \frac{1}{r} \right\rangle + 2(Ze^2)^2 \left(E_n^{(0)}\right) \left\langle \frac{1}{r^2} \right\rangle + \frac{(Ze^2)^2}{2m} \left\langle \frac{1}{r} p^2 \frac{1}{r} \right\rangle - (Ze^2)^3 \left\langle \frac{1}{r^3} \right\rangle + (Ze^2)^3 \left\langle \frac{1}{r^3} \right\rangle .$$

$$2m^2c^4 \left\langle \mathcal{H}_r^{(3)} \right\rangle = \left(E_n^{(0)}\right)^3 + 3Ze^2 \left(E_n^{(0)}\right)^2 \left\langle \frac{1}{r} \right\rangle + 2(Ze^2)^2 \left(E_n^{(0)}\right) \left\langle \frac{1}{r^2} \right\rangle + \frac{(Ze^2)^2}{2m} \left\langle \frac{1}{r} p^2 \frac{1}{r} \right\rangle .$$

Therefore

$$\left\langle \mathcal{H}_r^{(3)} \right\rangle = \frac{\left(E_n^{(0)}\right)^2}{2m^2c^4} \left[E_n^{(0)} + 3Ze^2 \left\langle \frac{1}{r} \right\rangle + \frac{2(Ze^2)^2}{\left(E_n^{(0)}\right)} \left\langle \frac{1}{r^2} \right\rangle + \frac{(Ze^2)^2}{2m\left(E_n^{(0)}\right)^2} \left\langle \frac{1}{r} p^2 \frac{1}{r} \right\rangle \right] . \tag{A.3}$$

For hydrogenic atoms

$$\left\langle \frac{1}{r} \right\rangle = \frac{Z}{a_0} \frac{1}{n^2}, \qquad \left\langle \frac{1}{r^2} \right\rangle = \frac{Z^2}{a_0^2} \frac{1}{n^3(\ell + \frac{1}{2})} .$$

In particular, for the hydrogen ground state we have $\langle 1/r \rangle = 1/a_0$ and $\langle 1/r^2 \rangle = 2/a_0^2$. We still have to compute $\left\langle \psi \left| \frac{1}{r} p^2 \frac{1}{r} \right| \psi \right\rangle$ for the hydrogen ground state wave function, which is

$$\psi(r) = \left(\frac{1}{a_0} \right)^{3/2} \frac{1}{\sqrt{\pi}} \, e^{-r/a_0} = b \, e^{-r/a_0} .$$

We need the following relations

$$p^2 \left(\frac{1}{r} \psi \right) = -\hbar^2 \nabla^2 \left(\frac{1}{r} \psi \right) = -\hbar^2 \left[\frac{1}{r} \left(\nabla^2 \psi \right) + \psi \left(\nabla^2 \frac{1}{r} \right) + 2 \left(\boldsymbol{\nabla} \frac{1}{r} \right) (\boldsymbol{\nabla} \psi) \right] .$$

$$\boldsymbol{\nabla} \frac{1}{r} = \frac{\partial (1/r)}{\partial r} \, \hat{\mathbf{r}} = -\frac{1}{r^2} \, \hat{\mathbf{r}} = -\frac{1}{r^3} \, \mathbf{r}, \qquad \nabla^2 \frac{1}{r} = -4\pi \delta(\mathbf{r}),$$

$$\boldsymbol{\nabla} \psi = \frac{\partial (b \, e^{-r/a_0})}{\partial r} \, \hat{\mathbf{r}} = -\frac{b}{a_0} \, e^{-r/a_0} \, \hat{\mathbf{r}} = -\frac{b}{a_0} \frac{1}{r} \, e^{-r/a_0} \, \mathbf{r},$$

$$\nabla^2 \psi(r) = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) = \frac{1}{r^2} \left(2r \frac{\partial \psi}{\partial r} + r^2 \frac{\partial^2 \psi}{\partial r^2} \right) = \frac{2}{r} \frac{\partial \psi}{\partial r} + \frac{\partial^2 \psi}{\partial r^2},$$

$$\nabla^2 \left(b \, e^{-r/a_0} \right) = \frac{2}{r} \left(-\frac{b}{a_0} \, e^{-r/a_0} \right) + \frac{b}{a_0^2} \, e^{-r/a_0} .$$

Collecting everything,

$$\begin{aligned} p^2 \left(\frac{1}{r} \psi \right) &= -\hbar^2 \left[-\frac{2b}{r^2 a_0} \, e^{-r/a_0} + \frac{b}{r a_0^2} \, e^{-r/a_0} - 4\pi b \, e^{-r/a_0} \, \delta(\mathbf{r}) + 2 \left(-\frac{1}{r^3} \, \mathbf{r} \right) \left(-\frac{b}{a_0} \frac{1}{r} \, e^{-r/a_0} \, \mathbf{r} \right) \right] \\ &= -\hbar^2 b \, e^{-r/a_0} \left[-\frac{2}{r^2 a_0} + \frac{1}{r a_0^2} - 4\pi \, \delta(\mathbf{r}) + \frac{2}{r^3} \, \mathbf{r} \left(\frac{1}{r a_0} \, \mathbf{r} \right) \right] \\ &= \hbar^2 b \, e^{-r/a_0} \left[\frac{2}{r^2 a_0} - \frac{1}{r a_0^2} + 4\pi \, \delta(\mathbf{r}) - \frac{2}{r^2 a_0} \right] = \hbar^2 b \, e^{-r/a_0} \left[4\pi \, \delta(\mathbf{r}) - \frac{1}{r a_0^2} \right] . \end{aligned}$$

$$\left(\frac{1}{r} \, p^2 \frac{1}{r} \right) \psi = 4\pi b \hbar^2 \, \frac{1}{r} \, e^{-r/a_0} \, \delta(\mathbf{r}) - \frac{\hbar^2 b}{a_0^2} \, \frac{1}{r^2} \, e^{-r/a_0} .$$

So the expectation value we are looking for is

$$\left\langle \psi \left| \frac{1}{r} p^2 \frac{1}{r} \right| \psi \right\rangle = 4\pi b^2 \hbar^2 \int \frac{1}{r} e^{-2r/a_0} \delta(\mathbf{r}) \, \mathrm{d}\tau - \frac{\hbar^2 b^2}{a_0^2} \int \frac{1}{r^2} e^{-2r/a_0} \, \mathrm{d}\tau,$$

where $\mathrm{d}\tau = r^2 \, \mathrm{d}r \, \mathrm{d}\Omega$ is the volume differential element. The angular integrals give just a 4π factor, and for the radial part we get

$$\begin{aligned}\left\langle \psi \left| \frac{1}{r} p^2 \frac{1}{r} \right| \psi \right\rangle &= (4\pi b)^2 \hbar^2 \int_0^\infty \frac{1}{r} e^{-2r/a_0} \delta(\mathbf{r}) \, r^2 \, \mathrm{d}r - \frac{4\pi\hbar^2 b^2}{a_0^2} \int_0^\infty \frac{1}{r^2} e^{-2r/a_0} \, r^2 \, \mathrm{d}r \\ &= -\frac{4\pi\hbar^2 b^2}{a_0^2} \int_0^\infty e^{-2r/a_0} \, \mathrm{d}r = -\frac{4\pi\hbar^2 b^2}{a_0^2} \frac{a_0}{2} = -\frac{2\pi\hbar^2}{a_0} b^2 = -\frac{2\pi\hbar^2}{a_0} \left(\frac{1}{a_0}\right)^3 \frac{1}{\pi} = -\frac{2\hbar^2}{a_0^4}.\end{aligned}$$

Using $a_0 = -e^2/(2E_1^{(0)}) = -e^2/(2E_1)$ (let's write from now on $E_1^{(0)} = E_1$, to simplify the notation), and remembering that $\hbar^2 = a_0 m e^2$, the last term within the brackets of equation (A.3) becomes (for hydrogen)

$$\frac{e^4}{2mE_1^2} \left\langle \frac{1}{r} p^2 \frac{1}{r} \right\rangle = -\frac{e^4}{2mE_1^2} \frac{2\hbar^2}{a_0^4} = -\frac{e^4}{2mE_1^2} \frac{2a_0 m e^2}{a_0^4} = -\frac{e^4}{E_1^2} \frac{e^2}{a_0^3} = -\frac{e^6}{E_1^2} \frac{8E_1^3}{e^6} = -8E_1.$$

Therefore the energy correction for the hydrogen ground state is ($e^2/a_0 = -2E_1$)

$$\begin{aligned}\left\langle \mathcal{H}_r^{(3)} \right\rangle &= \frac{E_1^2}{2m^2c^4} \left[E_1 + 3e^2 \left\langle \frac{1}{r} \right\rangle + \frac{2(e^2)^2}{E_1} \left\langle \frac{1}{r^2} \right\rangle - 8E_1 \right] \\ &= \frac{E_1^2}{2m^2c^4} \left[E_1 + \frac{3e^2}{a_0} + \frac{2(e^2)^2}{E_1} \frac{2}{a_0^2} - 8E_1 \right] = \frac{E_1^2}{2m^2c^4} \left[E_1 + \frac{3e^2}{a_0} + \frac{2(e^2)^2}{E_1} \frac{2}{a_0^2} - 8E_1 \right] \\ &= \frac{E_1^2}{2m^2c^4} \Big[E_1 - 6E_1 + 16E_1 - 8E_1 \Big] = \frac{3E_1^3}{2m^2c^4}.\end{aligned}$$

Equation (1.202) gives us the correction due to the second term of the kinetic energy expansion, which is $-5E_1^2/(2mc^2)$ for the hydrogen ground state. Let's display both corrections

$$\left\langle \mathcal{H}_r^{(2)} \right\rangle = -\frac{5E_1^2}{2mc^2}, \qquad \left\langle \mathcal{H}_r^{(3)} \right\rangle = \frac{3E_1^3}{2m^2c^4}.$$

The ratio we are asked to compute is

$$\left| \frac{\left\langle \mathcal{H}_r^{(3)} \right\rangle}{\left\langle \mathcal{H}_r^{(2)} \right\rangle} \right| = \frac{3E_1^3}{2m^2c^4} \frac{2mc^2}{5E_1^2} = \frac{3}{5} \frac{E_1}{mc^2} = \frac{3}{5} \frac{13.606\,\mathrm{eV}}{0.511 \times 10^6\,\mathrm{eV}} = 1.6 \times 10^{-5}.$$

4. You may probably recall, from your Classical Mechanics course, that energy and angular momentum are not the only constants of motion in a Kepler problem (attractive $1/r$ potential). Since the classical solutions have closed orbits, the vector pointing from the origin along the semi-major axis of the orbit, called the Runge-Lenz vector, is also a constant of motion. This fact can be used to give an algebraic derivation of the hydrogen spectrum, first published by Pauli in 1926, but in this problem we will only prove some properties of the Runge-Lenz vector.
   Consider the Hydrogen atom with its well-known Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m} - \frac{e^2}{r}.$$

   For Hydrogen the Runge-Lenz vector is given by[a]

$$\mathbf{R} = \frac{1}{2m} (\mathbf{p} \times \mathbf{L} - \mathbf{L} \times \mathbf{p}) - e^2 \frac{\mathbf{r}}{r}.$$

   The accidental degeneracy of the Hydrogen atom is due to the fact that this vector,

along with angular momentum, is conserved. In this assignment it will be useful to express the Runge-Lenz vector components in tensor form

$$R_i = \frac{1}{2m}\epsilon_{ijk}\left(p_j L_k - L_j p_k\right) - e^2\frac{x_i}{r},$$

where the Einstein convention of summing over repeated indices is applied. Remember that the Levi-Civita antisymmetric tensor symbol is given by:

$$\epsilon_{ijk} = \begin{cases} 1, & \text{if } ijk \text{ is an even permutation of 123,} \\ -1, & \text{if } ijk \text{ is an odd permutation of 123,} \\ 0, & \text{if any two of } ijk \text{ are equal.} \end{cases}$$

a) Use the tensor form to verify that $\mathbf{R}$ is hermitian.

b) Demonstrate that

$$\left[p_j, \frac{1}{r}\right] = -i\hbar\,\partial_j\left(\frac{1}{r}\right) = i\hbar\,\frac{x_j}{r^3}.$$

You will use this result in c). **Hint:** use a *dummy* wave-function.

c) In preparation for d) show that

$$[R_i, H] = -\frac{e^2}{2m}\left\{\epsilon_{ijk}\left(\left[p_j, \frac{1}{r}\right]L_k - L_j\left[p_k, \frac{1}{r}\right]\right) + \left[\frac{x_i}{r}, p^2\right]\right\}.$$

Hint: The commutators $[p_j, p^2] = 0$, $[L_k, p^2] = 0$ and $[L_k, r] = 0$ will be useful.

d) Starting from the expresion of c), show that $[\mathbf{R}, H] = 0$. This proves that $\mathbf{R}$ is a constant of motion in this Quantum Mechanics problem.

Hint: use the tensor form to demonstrate that $[R_i, H] = 0$ for each component of $\mathbf{R}$. You will need the commutator of item b) and the relation $\epsilon_{ijk}\epsilon_{mnk} = \delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}$.

[a] Note that in classical mechanics there is no difference between $\mathbf{p}\times\mathbf{L}$ and $-\mathbf{L}\times\mathbf{p}$, but it is the average of these operators that it is used in the quantum treatment, because this average is hermitian.

Solution:

a) For any two operators we have $(AB)^\dagger = B^\dagger A^\dagger$. Therefore, since $\mathbf{L}$ and $\mathbf{p}$ are hermitian operators,

$$R_i^\dagger = \frac{1}{2m}\epsilon_{ijk}\left(L_k p_j - p_k L_j\right) - e^2\frac{x_i}{r} = \frac{1}{2m}(-\epsilon_{ikj})\left(L_k p_j - p_k L_j\right) - e^2\frac{x_i}{r}$$
$$= \frac{1}{2m}\epsilon_{ikj}\left(p_k L_j - L_k p_j\right) - e^2\frac{x_i}{r} = R_i$$

b) As suggested, we see how the commutator operates on a generic wave-function:

$$\left[p_j, \frac{1}{r}\right]\psi = \left[-i\hbar\partial_j, \frac{1}{r}\right]\psi = -i\hbar\partial_j\left(\frac{1}{r}\psi\right) + i\hbar\frac{1}{r}\partial_j\left(\psi\right)$$
$$= -i\hbar\psi\partial_j\left(\frac{1}{r}\right) - i\hbar\frac{1}{r}\partial_j\left(\psi\right) + i\hbar\frac{1}{r}\partial_j\left(\psi\right) = -i\hbar\psi\partial_j\left(\frac{1}{r}\right) = i\hbar\,\frac{x_j}{r^3}\psi.$$

c) The commutator $[p_j, p^2] = 0$ is evident. Since the L_k operators in spherical coordinates have only angular derivatives, if follows that $[L_k, r] = 0$. Furthermore, recalling that $[AB, C] = A[B, C] + [A, C]B$, we have

$$[L_k, p^2] = [L_k, p_i p_i] = p_i[L_k, p_i] + [L_k, p_i]p_i.$$

$$[L_k, p_i] = \epsilon_{klm}[x_l p_m, p_i] = \epsilon_{klm}\Big(x_l[p_m, p_i] + [x_l, p_i]p_m\Big) = i\hbar\epsilon_{klm}\delta_{li}p_m = i\hbar\epsilon_{kim}p_m.$$

Thus

$$[L_k, p^2] = i\hbar\Big(\epsilon_{kim}p_i p_m + \epsilon_{kim}p_m p_i\Big) = 0.$$

These commutators lead to

$$[p_j L_k, p^2] = p_j[L_k, p^2] + [p_j, p^2]L_k = 0,$$

$$\left[p_j L_k, \frac{1}{r}\right] = p_j\left[L_k, \frac{1}{r}\right] + \left[p_j, \frac{1}{r}\right]L_k = \left[p_j, \frac{1}{r}\right]L_k, \qquad \text{etc.}$$

Therefore,

$$[R_i, H] = \left[\frac{1}{2m}\epsilon_{ijk}\left(p_j L_k - L_j p_k\right) - e^2\frac{x_i}{r},\ \frac{\mathbf{p}^2}{2m} - \frac{e^2}{r}\right] = \\ - \frac{e^2}{2m}\left\{\epsilon_{ijk}\left(\left[p_j, \frac{1}{r}\right]L_k - L_j\left[p_k, \frac{1}{r}\right]\right) + \left[\frac{x_i}{r}, p^2\right]\right\}.$$

d) Using $\epsilon_{ijk}\epsilon_{mnk} = \delta_{im}\delta_{jn} - \delta_{in}\delta_{jm}$ we can see that:

$$\begin{aligned}
\epsilon_{ijk}\left(\left[p_j, \frac{1}{r}\right]L_k - L_j\left[p_k, \frac{1}{r}\right]\right) &= i\hbar\,\frac{x_j}{r^3}\epsilon_{ijk}\epsilon_{kmn}x_m p_n - i\hbar\,\epsilon_{ijk}\epsilon_{jmn}x_m p_n\frac{x_k}{r^3} \\
&= i\hbar\,\frac{x_j}{r^3}(\delta_{im}\delta_{jn} - \delta_{in}\delta_{jm})x_m p_n - i\hbar\,(\delta_{in}\delta_{km} - \delta_{im}\delta_{kn})x_m p_n\frac{x_k}{r^3} \\
&= i\hbar\,\frac{x_j}{r^3}(x_i p_j - x_j p_i) - i\hbar\,(x_k p_i - x_i p_k)\frac{x_k}{r^3} \\
&= i\hbar\,\frac{x_j}{r^3}(x_i p_j - x_j p_i) + i\hbar\,(x_i p_k - x_k p_i)\frac{x_k}{r^3} \\
&= i\hbar\,\frac{x_j}{r^3}(x_i p_j - x_j p_i) + i\hbar\,(x_i p_j - x_j p_i)\frac{x_j}{r^3} \\
&= i\hbar\,\frac{1}{r^3}x_j x_i p_j - i\hbar\,\frac{1}{r^3}x_j x_j p_i + i\hbar\,x_i p_j x_j\frac{1}{r^3} - i\hbar\,x_j p_i x_j\frac{1}{r^3} \\
&= i\hbar\left(\frac{1}{r^3}x_j x_i p_j - \frac{1}{r}p_i + x_i p_j x_j\frac{1}{r^3} - \overbrace{(p_i r^2 + i\hbar\delta_{ij}x_j)}^{x_j p_i x_j}\frac{1}{r^3}\right) \\
&= i\hbar\left(\frac{1}{r^3}x_j x_i p_j - \frac{1}{r}p_i + x_i p_j x_j\frac{1}{r^3} - p_i\frac{1}{r} - i\hbar\frac{x_i}{r^3}\right) \\
&= i\hbar\left(\frac{1}{r^3}x_j x_i p_j - \frac{1}{r}p_i + \overbrace{(p_j x_j x_i + i\hbar\delta_{ij}x_j)}^{x_i p_j x_j}\frac{1}{r^3} - p_i\frac{1}{r} - i\hbar\frac{x_i}{r^3}\right) \\
&= i\hbar\left(\frac{1}{r^3}x_j x_i p_j - \frac{1}{r}p_i + (p_j x_j x_i + i\hbar x_i)\frac{1}{r^3} - p_i\frac{1}{r} - i\hbar\frac{x_i}{r^3}\right) \\
&= i\hbar\left(\frac{1}{r^3}x_j x_i p_j - \frac{1}{r}p_i + p_j x_j x_i\frac{1}{r^3} - p_i\frac{1}{r}\right). \qquad \text{(A)}
\end{aligned}$$

Besides,

$$\left[\frac{x_i}{r}, p^2\right] = \left[\frac{x_i}{r}, p_j p_j\right] = p_j\left[\frac{x_i}{r}, p_j\right] + \left[\frac{x_i}{r}, p_j\right]p_j,$$

$$\left[\frac{x_i}{r}, p_j\right] = \frac{1}{r}\left[x_i, p_j\right] + \left[\frac{1}{r}, p_j\right]x_i = i\hbar\frac{1}{r}\delta_{ij} - i\hbar\frac{1}{r^3}x_j x_i.$$

So

$$\left[\frac{x_i}{r}, p^2\right] = i\hbar p_j\frac{1}{r}\delta_{ij} - i\hbar p_j\frac{1}{r^3}x_j x_i + i\hbar\frac{1}{r}p_j\delta_{ij} - i\hbar\frac{1}{r^3}x_j x_i p_j \\
= i\hbar\left(p_i\frac{1}{r} - p_j x_j x_i\frac{1}{r^3} + \frac{1}{r}p_i - \frac{1}{r^3}x_j x_i p_j\right). \qquad \text{(B)}$$

It is evident that (A) + (B) = 0 and this cancels the expresion obtained in item c).

5. This problem shows an algebraic method to solve the hydrogen atom. Consider the dimensionless radial equation (1.106) of chapter 1 for $Z = 1$, namely

$$\frac{\mathrm{d}^2 u(r)}{\mathrm{d}r^2} + \left[E + \frac{2}{r} - \frac{\ell(\ell+1)}{r^2}\right] u(r) = 0, \tag{1}$$

and define the following operators

$$a_\ell^- = \frac{\mathrm{d}}{\mathrm{d}r} + \frac{\ell+1}{r} - \frac{1}{\ell+1}, \qquad a_\ell^+ = \frac{\mathrm{d}}{\mathrm{d}r} - \frac{\ell+1}{r} + \frac{1}{\ell+1}. \tag{2}$$

(a) Compute $a_\ell^- a_\ell^+$ and show that the radial equation (1) can be written as

$$\left(a_\ell^- a_\ell^+\right) u_\ell = \left(-E - \frac{1}{(\ell+1)^2}\right) u_\ell. \tag{3}$$

(b) Show that

$$a_\ell^+ a_\ell^- = a_{\ell+1}^- a_{\ell+1}^+ + \frac{1}{(\ell+2)^2} - \frac{1}{(\ell+1)^2}. \tag{4}$$

By multiplying equation (3) by a_ℓ^+, show that $a_\ell^+ u_\ell(r)$ satisfies the radial equation with the same eigenvalue E, but for an angular momentum $\ell' = \ell + 1$.

(c) Similarly, show that $a_{\ell-1}^- u_\ell(r)$ satisfies the radial equation with the same eigenvalue E, but for an angular momentum $\ell' = \ell - 1$.

(d) Integrate by parts the expectation value of $a_\ell^- a_\ell^+$, using the operators defined in (2) and the boundary condition $u(r) \to 0$ for $r \to 0$ and $r \to \infty$, to show that

$$\int_0^\infty u_\ell^* \left(a_\ell^- a_\ell^+\right) u_\ell \, \mathrm{d}r = -\int_0^\infty \left[\frac{\mathrm{d}u_\ell^*}{\mathrm{d}r}\frac{\mathrm{d}u_\ell}{\mathrm{d}r} + \left(\frac{\ell(\ell+1)}{r^2} - \frac{2}{r} + \frac{1}{(\ell+1)^2}\right)|u_\ell|^2\right] \mathrm{d}r.$$

Next, integrate also by parts $\int |a_\ell^+ u_\ell|^2 \, \mathrm{d}r$ using the definition of a_ℓ^+ to find that you obtain the same expression as before with a minus sign. This leads to

$$\int_0^\infty u_\ell^* \left(a_\ell^- a_\ell^+\right) u_\ell \, \mathrm{d}r = -\int_0^\infty |a_\ell^+ u_\ell|^2 \, \mathrm{d}r. \tag{5}$$

Use this result and equation (3) to prove the relation $-E \leq 1/(\ell+1)^2$.

(e) Show that, for a given value of E, there exists a maximum value of the angular momentum ℓ (call it $\ell_{\max}$) such that $E = -1/(\ell_{\max}+1)^2 \equiv -1/n^2$, where we define $n = \ell_{\max} + 1$. In this way we find the hydrogen energy levels. Show that the radial wave function $u_{\ell_{\max}}(r)$ satisfies the following differential equation

$$\left(\frac{\mathrm{d}}{\mathrm{d}r} - \frac{n}{r} + \frac{1}{n}\right) u_{\ell_{\max}}(r) = 0. \tag{6}$$

(f) Using equation (6) and the results proven in this problem, find the hydrogen radial wave functions $u(r)$ for $n = 3$, up to a proportionality factor.

Solution:

(a) Calling $P_\ell = \dfrac{\ell+1}{r} - \dfrac{1}{\ell+1}$, we have

$$\begin{aligned}\left(a_\ell^- a_\ell^+\right) u &= \left(\partial_r + P_\ell\right)\left(\partial_r - P_\ell\right) u = \left(\partial_r + P_\ell\right)\left(\partial_r u - P_\ell u\right) \\ &= \partial_r^2 u - (\partial_r P_\ell)u + P_\ell(\partial_r u) - P_\ell(\partial_r u) - P_\ell^2 u = \left(\partial_r^2 - \partial_r P_\ell - P_\ell^2\right) u \\ &= \left[\partial_r^2 + \frac{(\ell+1)}{r^2} - \frac{(\ell+1)^2}{r^2} - \frac{1}{(\ell+1)^2} + \frac{2}{r}\right] u.\end{aligned}$$

Therefore

$$a_\ell^- a_\ell^+ = \frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{\ell(\ell+1)}{r^2} + \frac{2}{r} - \frac{1}{(\ell+1)^2}. \tag{7}$$

It is now straightforward to show that the radial equation can be written as

$$\left(a_\ell^- a_\ell^+\right) u_\ell = \left(-E - \frac{1}{(\ell+1)^2}\right) u_\ell.$$

(b) Similarly, for $a_\ell^+ a_\ell^-$ we find

$$a_\ell^+ a_\ell^- = \frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{(\ell+1)(\ell+2)}{r^2} + \frac{2}{r} - \frac{1}{(\ell+1)^2}, \tag{8}$$

Therefore

$$a_\ell^+ a_\ell^- = a_{\ell+1}^- a_{\ell+1}^+ + \frac{1}{(\ell+2)^2} - \frac{1}{(\ell+1)^2}.$$

By multiplying equation (3) by a_ℓ^+ we obtain

$$a_\ell^+(a_\ell^- a_\ell^+)u_\ell = (a_\ell^+ a_\ell^-)a_\ell^+ u_\ell = \left(-E - \frac{1}{(\ell+1)^2}\right) a_\ell^+ u_\ell(r),$$

which, using (8), can also be written as

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{(\ell+1)(\ell+2)}{r^2} + \frac{2}{r} - \frac{1}{(\ell+1)^2}\right] a_\ell^+ u_\ell(r) = \left(-E - \frac{1}{(\ell+1)^2}\right) a_\ell^+ u_\ell(r),$$

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{(\ell+1)(\ell+2)}{r^2} + \frac{2}{r}\right] a_\ell^+ u_\ell(r) = -E a_\ell^+ u_\ell(r).$$

Therefore it is found that $a_\ell^+ u_\ell(r)$ satisfies the radial equation with the same eigenvalue, E, but for an angular momentum $\ell' = \ell + 1$.

(c) From expression (8) we find that the radial equation can be written as

$$(a_{\ell-1}^+ a_{\ell-1}^-)u_\ell = \left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{\ell(\ell+1)}{r^2} + \frac{2}{r} - \frac{1}{\ell^2}\right] u_\ell = \left(-E - \frac{1}{\ell^2}\right) u_\ell.$$

Multiplying the previous equation by $a_{\ell-1}^-$ we have

$$a_{\ell-1}^-(a_{\ell-1}^+ a_{\ell-1}^-)u_\ell = (a_{\ell-1}^- a_{\ell-1}^+)a_{\ell-1}^- u_\ell = \left(-E - \frac{1}{\ell^2}\right) a_{\ell-1}^- u_\ell,$$

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{\ell(\ell-1)}{r^2} + \frac{2}{r} - \frac{1}{\ell^2}\right] a_{\ell-1}^- u_\ell = \left(-E - \frac{1}{\ell^2}\right) a_{\ell-1}^- u_\ell,$$

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{\ell(\ell-1)}{r^2} + \frac{2}{r}\right] a_{\ell-1}^- u_\ell = -E a_{\ell-1}^- u_\ell.$$

Therefore $a_{\ell-1}^- u_\ell(r)$ satisfies the radial equation with the same eigenvalue E, but for an angular momentum $\ell' = \ell - 1$.

(d) Using expression (7), the expectation value of $\left(a_\ell^- a_\ell^+\right)$ computed with the radial function $u(r)$ can be written as

$$\int_0^\infty u_\ell^* \left(a_\ell^- a_\ell^+\right) u_\ell \, \mathrm{d}r = \int_0^\infty u_\ell^* \left(\frac{\mathrm{d}^2}{\mathrm{d}r^2} - \frac{\ell(\ell+1)}{r^2} + \frac{2}{r} - \frac{1}{(\ell+1)^2}\right) u_\ell \, \mathrm{d}r.$$

The integration by parts of the first term gives

$$\int_0^\infty u_\ell^* \frac{\mathrm{d}^2 u_\ell}{\mathrm{d}r^2} \, \mathrm{d}r = u_\ell^* \frac{\mathrm{d}u_\ell}{\mathrm{d}r}\bigg|_0^\infty - \int_0^\infty \frac{\mathrm{d}u_\ell^*}{\mathrm{d}r} \frac{\mathrm{d}u_\ell}{\mathrm{d}r} \, \mathrm{d}r = -\int_0^\infty \frac{\mathrm{d}u_\ell^*}{\mathrm{d}r} \frac{\mathrm{d}u_\ell}{\mathrm{d}r} \, \mathrm{d}r,$$

where we have used that, because of boundary conditions, $\lim_{r\to\infty} u_\ell = 0$ and $\lim_{r\to 0} \frac{u_\ell}{r}$ is finite (therefore $\lim_{r\to 0} u_\ell = 0$). We can then conclude that

$$\int_0^\infty u_\ell^* \left(a_\ell^- a_\ell^+\right) u_\ell \, \mathrm{d}r = \int_0^\infty \left[-\frac{\mathrm{d}u_\ell^*}{\mathrm{d}r} \frac{\mathrm{d}u_\ell}{\mathrm{d}r} + \left(-\frac{\ell(\ell+1)}{r^2} + \frac{2}{r} - \frac{1}{(\ell+1)^2}\right) |u_\ell|^2\right] \mathrm{d}r$$

Let us now calculate

$$\int_0^\infty |a_\ell^+ u_\ell|^2 \, \mathrm{d}r = \int_0^\infty (a_\ell^+ u_\ell)^* (a_\ell^+ u_\ell) \, \mathrm{d}r =$$
$$\int_0^\infty \left(\frac{\mathrm{d}u_\ell}{\mathrm{d}r} - \frac{\ell+1}{r} u_\ell + \frac{1}{\ell+1} u_\ell\right)^* \left(\frac{\mathrm{d}u_\ell}{\mathrm{d}r} - \frac{\ell+1}{r} u_\ell + \frac{1}{\ell+1} u_\ell\right) \mathrm{d}r =$$
$$\int_0^\infty \left[\frac{\mathrm{d}u_\ell^*}{\mathrm{d}r} \frac{\mathrm{d}u_\ell}{\mathrm{d}r} - \left(\frac{\ell+1}{r} - \frac{1}{\ell+1}\right) \left(\frac{\mathrm{d}u_\ell^*}{\mathrm{d}r} u_\ell + u_\ell^* \frac{\mathrm{d}u_\ell}{\mathrm{d}r}\right) + \left(\frac{\ell+1}{r} - \frac{1}{\ell+1}\right)^2 |u_\ell|^2\right] \mathrm{d}r =$$
$$\int_0^\infty \left[\frac{\mathrm{d}u_\ell^*}{\mathrm{d}r} \frac{\mathrm{d}u_\ell}{\mathrm{d}r} - \left(\frac{\ell+1}{r} - \frac{1}{\ell+1}\right) \frac{\mathrm{d}|u_\ell|^2}{\mathrm{d}r} + \left(\frac{\ell+1}{r} - \frac{1}{\ell+1}\right)^2 |u_\ell|^2\right] \mathrm{d}r.$$

The middle term is

$$\int_0^\infty \left(\frac{\ell+1}{r} - \frac{1}{\ell+1}\right) \frac{\mathrm{d}|u_\ell|^2}{\mathrm{d}r} \, \mathrm{d}r =$$
$$\left(\frac{\ell+1}{r} - \frac{1}{\ell+1}\right) |u_\ell|^2\bigg|_0^\infty + \int_0^\infty \left(\frac{\ell+1}{r^2}\right) |u_\ell|^2 \, \mathrm{d}r = \int_0^\infty \left(\frac{\ell+1}{r^2}\right) |u_\ell|^2 \, \mathrm{d}r,$$

where we have used integration by parts and, again, the fact that $\lim_{r\to 0} \frac{u_\ell}{r}$ is finite. Going back to the previous expression we have that

$$\int_0^\infty |a_\ell^+ u_\ell|^2 \, \mathrm{d}r = \int_0^\infty \left[\frac{\mathrm{d}u_\ell^*}{\mathrm{d}r} \frac{\mathrm{d}u_\ell}{\mathrm{d}r} - \left(\frac{\ell+1}{r^2}\right) |u_\ell|^2 + \left(\frac{\ell+1}{r} - \frac{1}{\ell+1}\right)^2 |u_\ell|^2\right] \mathrm{d}r =$$
$$\int_0^\infty \left[\frac{\mathrm{d}u_\ell^*}{\mathrm{d}r} \frac{\mathrm{d}u_\ell}{\mathrm{d}r} + \left(\frac{\ell(\ell+1)}{r^2} - \frac{2}{r} + \frac{1}{(\ell+1)^2}\right) |u_\ell|^2\right] \mathrm{d}r.$$

Therefore

$$\int_0^\infty u_\ell^* \left(a_\ell^- a_\ell^+\right) u_\ell \, \mathrm{d}r = -\int_0^\infty |a_\ell^+ u_\ell|^2 \, \mathrm{d}r \;\leq\; 0,$$

which is what we were asked to prove.

From equation (3) we see that the expectation value of $\left(a_\ell^- a_\ell^+\right)$ can also be written as

$$\int_0^\infty u_\ell^* \left(a_\ell^- a_\ell^+\right) u_\ell \, \mathrm{d}r = \left(-E_\ell - \frac{1}{(\ell+1)^2}\right) \int_0^\infty |u_\ell|^2 \, \mathrm{d}r.$$

Comparing the previous two equation we deduce

$$\left(-E_\ell - \frac{1}{(\ell+1)^2}\right) \leq 0 \implies -E \leq \frac{1}{(\ell+1)^2}.$$

(e) By repeatedly applying a_ℓ^+ to $u(r)$ we increase the value of ℓ by one integer. This process is limited from above because we must have $-E_\ell \leq 1/(\ell+1)^2$, so there must be a maximum value of ℓ, call it $\ell_{\max}$, such that

$$-E_\ell = \frac{1}{(\ell_{\max}+1)^2} \equiv \frac{1}{n^2}.$$

The function $a^+_{\ell_{\max}} u_{\ell_{\max}}$ must be identically zero, therefore $u_{\ell_{\max}}$ satisfies

$$\left(\frac{\mathrm{d}}{\mathrm{d}r} - \frac{\ell_{\max}+1}{r} + \frac{1}{\ell_{\max}+1}\right) u_{\ell_{\max}}(r) = \left(\frac{\mathrm{d}}{\mathrm{d}r} - \frac{n}{r} + \frac{1}{n}\right) u_{\ell_{\max}}(r) = 0.$$

(f) The energy levels are $E_n = -1/n^2$ in Rydberg units. The radial equation for $u_{\ell_{\max}}(r)$ can be solved doing

$$\frac{\mathrm{d}u}{\mathrm{d}r} = \left(\frac{n}{r} - \frac{1}{n}\right) u \implies \frac{\mathrm{d}u}{u} = \left(\frac{n}{r} - \frac{1}{n}\right) \mathrm{d}r,$$

$$\log(u) = n\log(r) - \frac{r}{n} + C \implies u_{\ell_{\max}} = A_{n\ell_{\max}}\, r^n\, e^{-r/n},$$

where $A_{n\ell_{\max}}$ is a normalization constant for the state with n and $\ell_{\max}$ quantum numbers. For $n=3$ we have $\ell = 0,1,2$, so $\ell_{\max} = 2$ and

$$u_{\ell_{\max}} = u_2 = A_{32}\, r^3\, e^{-r/3} = A_{32}\, r^3\, e^{-r\sqrt{E}}.$$

By repeatedly applying the $a^-_{\ell-1}$ operator we can work out the other solutions of the same energy, that is for $\ell = 1, 0$.

$$u_{\ell-1} = a^-_{\ell-1} u_\ell = \left(\frac{\mathrm{d}}{\mathrm{d}r} + \frac{(\ell-1)+1}{r} - \frac{1}{(\ell-1)+1}\right) u_\ell = \left(\frac{\mathrm{d}}{\mathrm{d}r} + \frac{\ell}{r} - \frac{1}{\ell}\right) u_\ell.$$

$$\begin{aligned} u_1 = a_1^- u_2 = \left(\frac{\mathrm{d}}{\mathrm{d}r} + \frac{2}{r} - \frac{1}{2}\right) u_2 &= \left(\frac{\mathrm{d}}{\mathrm{d}r} + \frac{2}{r} - \frac{1}{2}\right) A_{32}\, r^3\, e^{-r/3} \\ &= 5A_{32}\, r^2 \left(1 - \frac{r}{6}\right) e^{-r/3} = A_{31}\, r^2 \left(1 - \frac{r}{6}\right) e^{-r/3}. \end{aligned}$$

$$\begin{aligned} u_0 = a_0^- u_1 = \left(\frac{\mathrm{d}}{\mathrm{d}r} + \frac{1}{r} - 1\right) u_1 &= \left(\frac{\mathrm{d}}{\mathrm{d}r} + \frac{1}{r} - 1\right) A_{31}\, r^2 \left(1 - \frac{r}{6}\right) e^{-r/3} \\ &= A_{31}\, 3r \left(1 - \frac{2}{3}r + \frac{2}{27}r^2\right) e^{-r/3} = A_{30}\, r \left(1 - \frac{2}{3}r + \frac{2}{27}r^2\right) e^{-r/3}. \end{aligned}$$

Let's collect all results we got:

$$u_{32} = A_{32}\, r^3\, e^{-r/3}, \qquad u_{31} = A_{31}\, r^2 \left(1 - \frac{r}{6}\right) e^{-r/3}, \qquad u_{30} = A_{30}\, r \left(1 - \frac{2r}{3} + \frac{2r^2}{27}\right) e^{-r/3},$$

and compare them with the corresponding expressions quoted in chapter 1 of the lecture notes,

$$u_{32}(r) = \left(\frac{Z}{a_0}\right)^{3/2} \frac{4}{81\sqrt{30}} \frac{Z^2 r^3}{a_0^2}\, e^{-Zr/3a_0},$$

$$u_{31}(r) = \left(\frac{Z}{a_0}\right)^{3/2} \frac{8}{27\sqrt{6}} \frac{Zr^2}{a_0} \left(1 - \frac{Zr}{6a_0}\right) e^{-Zr/3a_0},$$

$$u_{30}(r) = \left(\frac{Z}{a_0}\right)^{3/2} \frac{2r}{3\sqrt{3}} \left[1 - \frac{2Zr}{3a_0} + \frac{2}{27}\left(\frac{Zr}{a_0}\right)^2\right] e^{-Zr/3a_0}.$$

It is clear that, setting $Z = a_0 = 1$, and leaving aside the normalization constants, we have reproduced the same radial functions.

A.2 Chapter 2.

1. Check that for a uniform magnetic field $\mathbf{B}$ we can choose the vector potential to be $\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r}$. (Hint: all you need to do is to verify that $\boldsymbol{\nabla} \times \mathbf{A} = \mathbf{B}$).

Solution: all we need to do is verify that $\boldsymbol{\nabla} \times \mathbf{A} = \mathbf{B}$. Incidentally, we can show that the relation $\boldsymbol{\nabla}\mathbf{A} = 0$ also holds. Using the rule $\boldsymbol{\nabla}(\mathbf{A} \times \mathbf{B}) = \mathbf{B}(\boldsymbol{\nabla} \times \mathbf{A}) - \mathbf{A}(\boldsymbol{\nabla} \times \mathbf{B})$, we have

$$\boldsymbol{\nabla}\mathbf{A} = \frac{1}{2}\boldsymbol{\nabla}(\mathbf{B} \times \mathbf{r}) = \frac{1}{2}\Big[\mathbf{r}(\boldsymbol{\nabla} \times \mathbf{B}) - \mathbf{B}(\boldsymbol{\nabla} \times \mathbf{r})\Big] = 0,$$

because $\boldsymbol{\nabla} \times \mathbf{B} = 0$ ($\mathbf{B}$ is uniform) and $\boldsymbol{\nabla} \times \mathbf{r} = 0$. Next, we need to use the rule $\boldsymbol{\nabla} \times (\mathbf{A} \times \mathbf{B}) = (\mathbf{B}\boldsymbol{\nabla})\mathbf{A} - (\mathbf{A}\boldsymbol{\nabla})\mathbf{B} + \mathbf{A}(\boldsymbol{\nabla}\mathbf{B}) - \mathbf{B}(\boldsymbol{\nabla}\mathbf{A})$, to find

$$\boldsymbol{\nabla} \times \mathbf{A} = \frac{1}{2}\boldsymbol{\nabla} \times (\mathbf{B} \times \mathbf{r}) = \frac{1}{2}\Big[(\mathbf{r}\boldsymbol{\nabla})\mathbf{B} - (\mathbf{B}\boldsymbol{\nabla})\mathbf{r} + \mathbf{B}(\boldsymbol{\nabla}\mathbf{r}) - \mathbf{r}(\boldsymbol{\nabla}\mathbf{B})\Big].$$

But, since $\mathbf{B}$ is uniform, $(\mathbf{r}\boldsymbol{\nabla})\mathbf{B} = 0$ and $\mathbf{r}(\boldsymbol{\nabla}\mathbf{B}) = 0$. Furthermore $\boldsymbol{\nabla}\mathbf{r} = \frac{\partial x}{\partial x} + \frac{\partial y}{\partial y} + \frac{\partial z}{\partial z} = 3$. And for the remaining term we have

$$(\mathbf{B}\boldsymbol{\nabla})\mathbf{r} = \left(B_x\frac{\partial}{\partial x} + B_y\frac{\partial}{\partial y} + B_z\frac{\partial}{\partial z} +\right)(x\hat{\mathbf{e}}_x + y\hat{\mathbf{e}}_y + z\hat{\mathbf{e}}_z) = B_x\hat{\mathbf{e}}_x + B_y\hat{\mathbf{e}}_y + B_z\hat{\mathbf{e}}_z = \mathbf{B}.$$

So finally

$$\boldsymbol{\nabla} \times \mathbf{A} = \frac{1}{2}\Big(-\mathbf{B} + 3\mathbf{B}\Big) = \mathbf{B}.$$

2. The muonium is an electromagnetic bound state of an electron and a positive muon, μ^+e^-. This system has received considerable attention (theoretical and experimental) to test the validity of quantum electrodynamics. In particular, very precise information can be obtained from the study of the hyperfine structure of its $1s$ ground state. Here we will perform a simplified analysis.

   a) Compute the hyperfine interaction for the muonium ground state in the presence of a constant external magnetic field along the z-axis, $\mathbf{B} = (0, 0, B)$. Use the uncoupled basis and show all algebraic steps in the process of getting the matrix elements.

   b) Repeat the calculations of the previous point in the coupled basis, showing again all needed algebraic steps.

   c) Study the energy shifts as a function the magnetic field amplitude, paying attention to the low and high magnetic field regimes. Draw a sketch (not necessarily to scale) showing the energy of the four levels versus B.

   d) Find the possible points where the curves $E_1(B)$, $E_2(B)$, $E_3(B)$, $E_4(B)$ intersect with one another. Taking $\hbar^2 A_F = 1.846 \times 10^{-5}$ eV for muonium, compute the values of B (in Gauss and Tesla) of the intersection points (ignore the intersection point at $B = 0$ and take always $B > 0$).

   Hints: Take equation (2.180), namely

$$\mathcal{H} \;=\; \frac{g_e\mu_{\rm B}}{\hbar}\mathbf{BS} \;+\; A_F\mathbf{IS} \;-\; \frac{g_{\rm N}\mu_{\rm N}}{\hbar}\mathbf{BI},$$

   and make the substitutions

$$\mathbf{S} \to \mathbf{S}_1, \quad \mathbf{I} \to \mathbf{S}_2, \quad \mu_{\rm N} \to \frac{m_e}{m_\mu}\mu_{\rm B}, \quad g_{\rm N} \to g_e \approx 2, \quad \frac{2\mu_{\rm B}}{\hbar}B = \omega_1, \quad \frac{m_e}{m_\mu}\frac{2\mu_{\rm B}}{\hbar}B = \omega_2,$$

to get

$$\mathcal{H} \;=\; \frac{2\mu_{\rm B}}{\hbar}\mathbf{B}\mathbf{S}_1 - \frac{m_e}{m_\mu}\frac{2\mu_{\rm B}}{\hbar}\mathbf{B}\mathbf{S}_2 \;+\; A_F\mathbf{S}_1\mathbf{S}_2 \;=\; \omega_1 S_{1z} - \omega_2 S_{2z} \;+\; A_F\mathbf{S}_1\mathbf{S}_2,$$

where $\mathbf{S}_1$ and $\mathbf{S}_2$ are the spin operators for the electron and muon, respectively. The states of the uncoupled basis are $|S_1, S_2, m_{s1}, m_{s2}\rangle \equiv |m_{s1}, m_{s2}\rangle$, while the states of the coupled basis are $|S_1, S_2, F, m_F\rangle \equiv |F, m_F\rangle$, where $\mathbf{F} = \mathbf{S}_1 + \mathbf{S}_2$. Use a table of CG-coefficients to relate the elements of one basis to the elements of the other.

Solution:

a) To compute the $\mathcal{H}$ matrix elements in the uncoupled basis, $|m_{s1}, m_{s2}\rangle$, write $\mathcal{H}$ as

$$\mathcal{H} = \omega_1 S_{1z} - \omega_2 S_{2z} + A_F\big[-(S_1)_{+1}(S_2)_{-1} + (S_1)_0(S_2)_0 - (S_1)_{-1}(S_2)_{+1}\big],$$

and use the relations

$$J_{\pm 1}\,|jm\rangle = \mp\hbar\sqrt{\tfrac{1}{2}\left[j(j+1) - m(m\pm 1)\right]}\;|j,(m\pm 1)\rangle\,, \qquad J_0\,|jm\rangle = m\hbar\,|jm\rangle\,,$$

to find the following matrix elements (we omit here the detailed calculations that the student should do)

	$\left\vert\frac{1}{2}\frac{1}{2}\right\rangle$	$\left\vert\frac{1}{2}-\frac{1}{2}\right\rangle$	$\left\vert-\frac{1}{2}\frac{1}{2}\right\rangle$	$\left\vert-\frac{1}{2}-\frac{1}{2}\right\rangle$
$\left\langle\frac{1}{2}\frac{1}{2}\right\vert$	$\frac{\hbar}{2}(\omega_1-\omega_2)+\frac{\hbar^2}{4}A_F$			
$\left\langle\frac{1}{2}-\frac{1}{2}\right\vert$		$\frac{\hbar}{2}(\omega_1+\omega_2)-\frac{\hbar^2}{4}A_F$	$\frac{\hbar^2}{2}A_F$	
$\left\langle-\frac{1}{2}\frac{1}{2}\right\vert$		$\frac{\hbar^2}{2}A_F$	$-\frac{\hbar}{2}(\omega_1+\omega_2)-\frac{\hbar^2}{4}A_F$	
$\left\langle-\frac{1}{2}-\frac{1}{2}\right\vert$				$-\frac{\hbar}{2}(\omega_1-\omega_2)\frac{\hbar^2}{4}+A_F$

Two of the eigenvalues are obvious

$$E_1 = \frac{\hbar}{2}(\omega_1-\omega_2)+\frac{\hbar^2}{4}A_F, \qquad E_2 = -\frac{\hbar}{2}(\omega_1-\omega_2)+\frac{\hbar^2}{4}A_F.$$

The remaining two are obtained solving a quadratic equation. They are

$$E_3 = -\frac{\hbar^2}{4}A_F + \sqrt{\frac{1}{4}\hbar^4A_F^2 + \frac{1}{4}\hbar^2(\omega_1+\omega_2)^2},$$

$$E_4 = -\frac{\hbar^2}{4}A_F - \sqrt{\frac{1}{4}\hbar^4A_F^2 + \frac{1}{4}\hbar^2(\omega_1+\omega_2)^2}.$$

b) To compute the matrix elements in the coupled basis, $|Fm_F\rangle = \{|1,1\rangle, |1,-1\rangle, |1,0\rangle, |0,0\rangle\}$, it is convenient to use the relations

$$F^2 = S_1^2 + S_2^2 + 2\mathbf{S}_1\mathbf{S}_2, \qquad \mathbf{S}_1\mathbf{S}_2 = \frac{1}{2}(F^2 - S_1^2 - S_2^2),$$

to write $\mathcal{H}$ as

$$\mathcal{H} = \omega_1 S_{1z} - \omega_2 S_{2z} + \frac{1}{2}A_F\Big(F^2 - S_1^2 - S_2^2\Big).$$

Using the CG-table of appendix B.5.3 we find

$$|1,1\rangle = \left|\frac{1}{2},\frac{1}{2}\right\rangle, \qquad |1,-1\rangle = \left|-\frac{1}{2},-\frac{1}{2}\right\rangle.$$

$$|1,0\rangle = \frac{1}{\sqrt{2}}\left|\frac{1}{2},-\frac{1}{2}\right\rangle + \frac{1}{\sqrt{2}}\left|-\frac{1}{2},\frac{1}{2}\right\rangle, \qquad |0,0\rangle = \frac{1}{\sqrt{2}}\left|\frac{1}{2},-\frac{1}{2}\right\rangle - \frac{1}{\sqrt{2}}\left|-\frac{1}{2},\frac{1}{2}\right\rangle.$$

Therefore

$$(\omega_1 S_{1z} - \omega_2 S_{2z})\,|1,1\rangle = \frac{\hbar}{2}(\omega_1-\omega_2)\,|1,1\rangle\,.$$

$$(\omega_1 S_{1z} - \omega_2 S_{2z})\,|1,-1\rangle = -\frac{\hbar}{2}(\omega_1 - \omega_2)\,|1,-1\rangle\,.$$

$$(\omega_1 S_{1z} - \omega_2 S_{2z})\,|1,0\rangle = \frac{\hbar}{2\sqrt{2}}(\omega_1+\omega_2)\left|\frac{1}{2},-\frac{1}{2}\right\rangle - \frac{\hbar}{2\sqrt{2}}(\omega_1+\omega_2)\left|-\frac{1}{2},\frac{1}{2}\right\rangle = \frac{\hbar}{2}(\omega_1+\omega_2)\,|0,0\rangle\,.$$

$$(\omega_1 S_{1z} - \omega_2 S_{2z})\,|0,0\rangle = \frac{\hbar}{2\sqrt{2}}(\omega_1+\omega_2)\left|\frac{1}{2},-\frac{1}{2}\right\rangle + \frac{\hbar}{2\sqrt{2}}(\omega_1+\omega_2)\left|-\frac{1}{2},\frac{1}{2}\right\rangle = \frac{\hbar}{2}(\omega_1+\omega_2)\,|1,0\rangle\,.$$

So the matrix elements of this part of the Hamiltonian are

	$\|1,1\rangle$	$\|1,-1\rangle$	$\|1,0\rangle$	$\|0,0\rangle$
$\langle 1,1\|$	$\frac{\hbar}{2}(\omega_1-\omega_2)$			
$\langle 1,-1\|$		$-\frac{\hbar}{2}(\omega_1-\omega_2)$		
$\langle 1,0\|$				$\frac{\hbar}{2}(\omega_1+\omega_2)$
$\langle 0,0\|$			$\frac{\hbar}{2}(\omega_1+\omega_2)$	

The matrix elements of the remaing part are diagonal

$$\left\langle 1,1\left|\frac{1}{2}A_F\Big(F^2-S_1^2-S_2^2\Big)\right|1,1\right\rangle = \frac{\hbar^2}{2}A_F\left(2-\frac{3}{4}-\frac{3}{4}\right) = \frac{\hbar^2}{4}A_F.$$

$$\left\langle 1,-1\left|\frac{1}{2}A_F\Big(F^2-S_1^2-S_2^2\Big)\right|1,-1\right\rangle = \left\langle 1,0\left|\frac{1}{2}A_F\Big(F^2-S_1^2-S_2^2\Big)\right|1,0\right\rangle = \frac{\hbar^2}{4}A_F.$$

$$\left\langle 0,0\left|\frac{1}{2}A_F\Big(F^2-S_1^2-S_2^2\Big)\right|0,0\right\rangle = \frac{\hbar^2}{2}A_F\left(0-\frac{3}{4}-\frac{3}{4}\right) = -\frac{3\hbar^2}{4}A_F.$$

	$\|1,1\rangle$	$\|1,-1\rangle$	$\|1,0\rangle$	$\|0,0\rangle$
$\langle 1,1\|$	$\frac{\hbar^2}{4}A_F$			
$\langle 1,-1\|$		$\frac{\hbar^2}{4}A_F$		
$\langle 1,0\|$			$\frac{\hbar^2}{4}A_F$	
$\langle 0,0\|$				$-\frac{3\hbar^2}{4}A_F$

The matrix elements of the full Hamiltonian are then

	$\|1,1\rangle$	$\|1,-1\rangle$	$\|1,0\rangle$	$\|0,0\rangle$
$\langle 1,1\|$	$\frac{\hbar}{2}(\omega_1-\omega_2)+\frac{\hbar^2}{4}A_F$			
$\langle 1,-1\|$		$-\frac{\hbar}{2}(\omega_1-\omega_2)+\frac{\hbar^2}{4}A_F$		
$\langle 1,0\|$			$\frac{\hbar^2}{4}A_F$	$\frac{\hbar}{2}(\omega_1+\omega_2)$
$\langle 0,0\|$			$\frac{\hbar}{2}(\omega_1+\omega_2)$	$-\frac{3\hbar^2}{4}A_F$

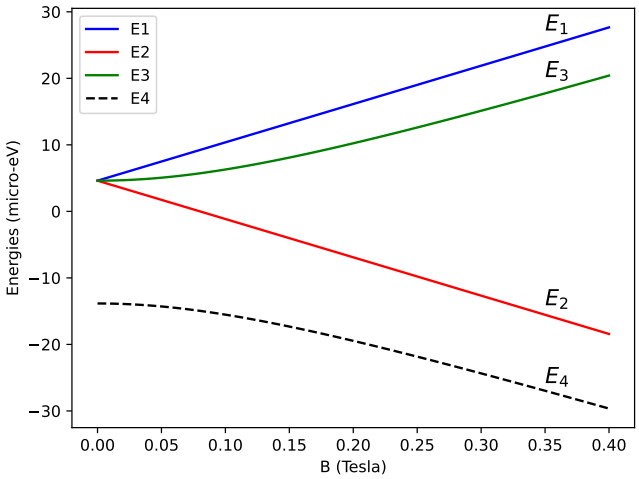
It is easy to see that we get the same eigenvalues as before. They can be written in a more convenient way defining the ratio $\kappa = m_e/m_\mu$ and $\omega \equiv \omega_1$,

$$\omega \equiv \omega_1 = \frac{2\mu_{\rm B}}{\hbar}B, \qquad \omega_2 = \frac{m_e}{m_\mu}\frac{2\mu_{\rm B}}{\hbar}B = \frac{m_e}{m_\mu}\omega_1 = \kappa\omega.$$

$$E_1 = \frac{\hbar^2}{4}A_F + \frac{\hbar}{2}\omega(1-\kappa), \qquad E_2 = \frac{\hbar^2}{4}A_F - \frac{\hbar}{2}\omega(1-\kappa),$$

$$E_3 = -\frac{\hbar^2}{4}A_F + \frac{1}{2}\sqrt{(\hbar^2 A_F)^2 + \hbar^2\omega^2(1+\kappa)^2}, \qquad E_4 = -\frac{\hbar^2}{4}A_F - \frac{1}{2}\sqrt{(\hbar^2 A_F)^2 + \hbar^2\omega^2(1+\kappa)^2}.$$

c) Since ω is proportional to B, we can just as well study the dependency of the previous energies with respect to ω. Both E_1 and E_2 are linear functions of ω, the first one increases with a slope given by $\hbar(1-\kappa)/2$, and the second decreases with the same slope. E_3 and E_4 are not linear functions of $\omega(B)$ at small values of $\omega(B)$, but become linear at high values. The following plot shows the B-dependence of the energy levels for $B < 0.4$ Tesla.



d) Given the sign of the slopes, the only combination of curves that can intersect are $E_1(\omega)$ with $E_3(\omega)$, and $E_2(\omega)$ with $E_4(\omega)$. Let's study first $E_1(\omega)$ with $E_3(\omega)$.
Solving the equation which arises from imposing $E_1 = E_3$,

$$\frac{\hbar^2}{4}A_F + \frac{\hbar}{2}\omega(1-\kappa) = -\frac{\hbar^2}{4}A_F + \frac{1}{2}\sqrt{(\hbar^2 A_F)^2 + \hbar^2\omega^2(1+\kappa)^2},$$

we get

$$\omega = \frac{\hbar}{2}A_F\left(\frac{1}{\kappa} - 1\right), \quad \text{or} \quad B = \frac{\hbar^2 A_F}{4\mu_B}\left(\frac{m_\mu}{m_e} - 1\right).$$

Putting numbers,

$$B = \frac{1.846\times10^{-5}\,\text{eV}}{4\times5.788\times10^{-5}\,\text{eV}\cdot\text{Tesla}^{-1}}\left(\frac{105.66}{0.511} - 1\right) = 16.41\,\text{Tesla} = 1.641\times10^5\,\text{Gauss}.$$

For the condition $E_2 = E_4$ we get

$$B = \frac{\hbar^2 A_F}{4\mu_B}\left(1 - \frac{m_\mu}{m_e}\right) < 0,$$

but this is discarded since we are assuming $B > 0$.

3. In this exercise we shall study the modification of the energy spectrum of a hydrogen atom placed in crossed static electric and magnetic fields in perturbation theory. To simplify things **we will consider that the electron is spinless and neglect also fine structure effects**.
In a famous 1925–1926 work on the hydrogen atom, Wolfgang Pauli was able to calculate the splitting of the levels in an electric field (Stark effect) or in a magnetic field (Zeeman effect). Pauli also noticed that he could obtain a simple and compact formula for the level splitting in a superposition of a magnetic field **B** and an electric field $\mathcal{E}$ both static and uniform, and perpendicular to each other. In this case, he found that a level with principal quantum number n is split into $2n-1$ sublevels $E_n + \delta E_n^{(k)}$ with

$$\delta E_n^{(k)} = \hbar k(\omega_m^2 + n^2\omega_e^2)^{1/2}, \tag{1}$$

where k is an integer ranging from $-(n-1)$ to $n-1$, while ω_m and ω_e are proportional to B and $\mathcal{E}$, respectively .

   (a) Consider the $n = 2$ levels of the hydrogen atom (remember to neglect spin effects). Assume **B** is along the z axis. Use first-order perturbation theory to

obtain the energy levels and the corresponding eigenstates in the presence of **B** only. Check that (1) is valid in this case and give the value of ω_m.

(b) Now assume that, alongside **B** along the z axis, $\mathcal{E}$ is established along the x axis and find, in first-order perturbation theory for the $n = 2$ levels of the hydrogen atom, the corresponding energy shifts. Check that (1) is also valid in this case and give the value of ω_e.

Solution:

(a) For a spinless particle the interaction with an external magnetic field aligned with the z-axis, given by eq. (2.17) of the handouts, reduces to

$$\mathcal{H}_{\rm m} = \frac{1}{\hbar}\mu_{\rm B} L_z B = \frac{e}{2m} L_z B.$$

Since $[\mathcal{H}_{\rm m}, L_z] = 0$ both $\mathcal{H}_{\rm m}$ and L_z are diagonal in the uncoupled basis and the calculation of the splitting of the energy levels in this basis (*Paschen-Bach* regime) is trivial:

$$\Delta E = \langle 2\ell m_\ell | \mathcal{H}_{\rm m} | 2\ell m_\ell \rangle = \frac{e}{2m} B \, \langle 2\ell m_\ell | L_z | 2\ell m_\ell \rangle = \frac{e\hbar}{2m} B m_\ell = \hbar \omega_m m_\ell,$$

where we can identify that

$$\omega_m = \frac{eB}{2m}.$$

The eigenstates and energy levels are those of the uncoupled basis and, of course, equation (1) is accomplished with $\omega_e = 0$ and $k = m_\ell$.

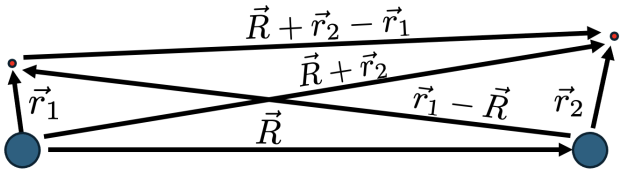
For $n = 2$ the hamiltonian matrix is:

$$\langle 2\ell' m'_\ell | \mathcal{H}_m | 2\ell m_\ell \rangle = \hbar\omega_m \begin{array}{c} \\ \langle 00| \\ \langle 1\bar{1}| \\ \langle 10| \\ \langle 11| \end{array} \begin{array}{c} \begin{array}{cccc} |00\rangle & |1\bar{1}\rangle & |10\rangle & |11\rangle \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

(b) In the presence of $\mathcal{E}$ only, the perturbing Hamiltonian is the electric dipole term $\mathcal{H}_s = e\mathcal{E}\mathbf{r} = e\mathcal{E}x$ (remember the electric field is parallel to the x-axis). Write the matrix representing $\mathcal{H}_s$ in the 4-dimensional $n = 2$ subspace under consideration.

A.3 Chapter 3.

1. Consider two hydrogen atoms with the relative positions of the two protons and the two electrons as shown in this figure,



Assume the protons to be static in their positions. The hamiltonian of the system is then

$$\begin{aligned}\mathcal{H} &= \mathcal{H}_0 + \mathcal{H}',\\ \mathcal{H}_0 &= \frac{p_1^2}{2m} + \frac{p_2^2}{2m} - \frac{e^2}{r_1} - \frac{e^2}{r_2},\\ \mathcal{H}' &= \frac{e^2}{R} + \frac{e^2}{|\mathbf{R}+\mathbf{r}_2-\mathbf{r}_1|} - \frac{e^2}{|\mathbf{R}-\mathbf{r}_1|} - \frac{e^2}{|\mathbf{R}+\mathbf{r}_2|}.\end{aligned}$$

a) Show that for $R \gg r_1$ and $R \gg r_2$

$$\mathcal{H}' \approx \left(\frac{e^2}{R^3}\right)\left(\mathbf{r}_1\cdot\mathbf{r}_2 - 3\left(\mathbf{r}_1\cdot\hat{\mathbf{R}}\right)\left(\mathbf{r}_2\cdot\hat{\mathbf{R}}\right)\right),$$

where $\hat{\mathbf{R}} = \mathbf{R}/R$.

Consider that the approximation to $\mathcal{H}'$ given in part a) can be treated as a perturbation to $\mathcal{H}_0$ and to its ground state, $|\psi_0^{(0)}\rangle = |\phi_{1s}\phi_{1s}\rangle$, with unperturbed energy of $E_0^{(0)} = -e^2/a_0 = -2$ Ry.[a]

b) Let P_k be the parity operator that, acting on a n-particle state, applies parity transformation only to the k-th particle state and leaves all the other particle states unchanged. The total parity operator can then be written as $P = \prod_{k=1}^{n} P_k$. Consider an operator $F(\mathbf{r}_1, ..., \mathbf{r}_k, ..., \mathbf{r}_n)$ such that

$$P_k F P_k = F(\mathbf{r}_1, ..., -\mathbf{r}_k, ..., \mathbf{r}_n) = (-1)^{\pi_{Fk}} F(\mathbf{r}_1, ..., \mathbf{r}_k, ..., \mathbf{r}_n).$$

Show that for initial $|\phi_i\rangle = |\phi_{i1}, ..., \phi_{ik}, ..., \phi_{in}\rangle$ and final $|\phi_f\rangle = |\phi_{f1}, ..., \phi_{fk}, ..., \phi_{fn}\rangle$ states verifying $P_k|\phi_i\rangle = (-1)^{\pi_{ik}}|\phi_i\rangle$ and $P_k|\phi_f\rangle = (-1)^{\pi_{fk}}|\phi_f\rangle$, we have the identity $\langle\phi_f|F|\phi_i\rangle = 0$ unless $1 = (-1)^{\pi_{fk}+\pi_{Fk}+\pi_{ik}}$, for $k = 1, ..., n$. This means that for the F operator there are conservation laws not only for the global parity but also for the parity of k-particle state. Use this to demonstrate that, to first order in perturbation theory, the energy shift is zero, $\Delta E_0^{(1)} = \langle\psi_0^{(0)}|\mathcal{H}'|\psi_0^{(0)}\rangle = 0$.

c) Show, using the demonstration of part b), that if $|\psi_n^{(0)}\rangle$ is an excited state of $\mathcal{H}_0$ with principal quantum numbers $n_1 = 1$ and $n_2 = n$, or $n_1 = n$ and $n_2 = 1$, we have $\langle\psi_n^{(0)}|\mathcal{H}'|\psi_0^{(0)}\rangle = 0$. This means that the lowest-energy excited states for which $\langle\psi_e^{(0)}|\mathcal{H}'|\psi_0^{(0)}\rangle \neq 0$ are those with $n_1 = n_2 = 2$ and energy $E_e^{(0)} = E_0^{(0)}/4$.

d) In the discussion of the quadratic Stark effect, eq. (2.82) to eq. (2.86) in the lecture notes, a lower bound was found for the energy shift of the hydrogen ground state. For the two hydrogen atoms system use the same strategy to demonstrate that a lower bound of the energy shift produced by $\mathcal{H}'$ is

$$\Delta E_0^{(2)} = \sum_{k\neq 0} \frac{\left|\left\langle \psi_k^{(0)} \,\middle|\, \mathcal{H}' \,\middle|\, \psi_0^{(0)} \right\rangle\right|^2}{E_0^{(0)} - E_k^{(0)}} \;>\; \frac{1}{E_0^{(0)} - E_e^{(0)}} \sum_{k\neq 0} \left|\left\langle \psi_k^{(0)} \,\middle|\, \mathcal{H}' \,\middle|\, \psi_0^{(0)} \right\rangle\right|^2,$$

$$8\left(\frac{a_0}{R}\right)^6 E_0^{(0)} < \Delta E_0^{(2)}.$$

[a] The prescription that the order in the ket applies to the dynamical variables is used. For instance, for $|\phi_{i1}\phi_{i2}\rangle$ the corresponding spatial wave function is $\phi_{i1}(\mathbf{r}_1)\phi_{i2}(\mathbf{r}_2)$.

Solution:

a) We can make the usual approximation, which coincides with the multipolar expansion, up to $\frac{1}{R}\mathcal{O}\left(\frac{r^3}{R^3}\right)$,

$$\frac{1}{|\mathbf{R}+\mathbf{r}|} = \frac{1}{\sqrt{(\mathbf{R}+\mathbf{r})^2}} = \frac{1}{\sqrt{R^2+r^2+2\mathbf{R}\cdot\mathbf{r}}} = \frac{1}{R}\frac{1}{\sqrt{1+\frac{r^2}{R^2}+2\frac{\mathbf{R}\cdot\mathbf{r}}{R^2}}} \approx \frac{1}{R}\frac{1}{\sqrt{1+\frac{r^2}{R^2}+2\frac{\mathbf{R}\cdot\mathbf{r}}{R^2}}}$$
$$\approx \frac{1}{R}\left[1-\frac{1}{2}\left(\frac{r^2}{R^2}+2\frac{\mathbf{R}\cdot\mathbf{r}}{R^2}\right)+\frac{3}{8}\left(\frac{r^2}{R^2}+2\frac{\mathbf{R}\cdot\mathbf{r}}{R^2}\right)^2\right] \approx \frac{1}{R}\left[1-\frac{1}{2}\left(\frac{r^2}{R^2}+2\frac{\mathbf{R}\cdot\mathbf{r}}{R^2}\right)+\frac{3}{2}\left(\frac{\mathbf{R}\cdot\mathbf{r}}{R^2}\right)^2\right]$$
$$= \frac{1}{R} - \frac{\mathbf{R}\cdot\mathbf{r}}{R^3} + \frac{3\,(\mathbf{R}\cdot\mathbf{r})^2 - r^2R^2}{2R^5}.$$

Applying this to $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$, $\mathbf{r} = -\mathbf{r}_1$ and $\mathbf{r} = \mathbf{r}_2$,

$$\frac{1}{e^2}\mathcal{H}' = \frac{1}{R} + \frac{1}{|\mathbf{R}+\mathbf{r}_2-\mathbf{r}_1|} - \frac{1}{|\mathbf{R}-\mathbf{r}_1|} - \frac{1}{|\mathbf{R}+\mathbf{r}_2|} \approx \cancel{\frac{1}{R}} + \cancel{\frac{1}{R}} - \cancel{\frac{\mathbf{R}\cdot\mathbf{r}_2 - \mathbf{R}\cdot\mathbf{r}_1}{R^3}}$$
$$+\frac{3\,(\mathbf{R}\cdot(\mathbf{r}_2-\mathbf{r}_1))^2 - (\mathbf{r}_2-\mathbf{r}_1)^2R^2}{2R^5} - \cancel{\frac{1}{R}} + \cancel{\frac{\mathbf{R}\cdot\mathbf{r}_1}{R^3}} - \frac{3\,(\mathbf{R}\cdot\mathbf{r}_1)^2 - r_1^2R^2}{2R^5} - \cancel{\frac{1}{R}} + \cancel{\frac{\mathbf{R}\cdot\mathbf{r}_2}{R^3}} - \frac{3\,(\mathbf{R}\cdot\mathbf{r}_2)^2 - r_2^2R^2}{2R^5}$$
$$= \frac{\cancel{3(\mathbf{R}\cdot\mathbf{r}_2)^2} + \cancel{3(\mathbf{R}\cdot\mathbf{r}_1)^2} - 6(\mathbf{R}\cdot\mathbf{r}_2)(\mathbf{R}\cdot\mathbf{r}_1) - \cancel{r_2^2R^2} - \cancel{r_1^2R^2} + 2(\mathbf{r}_1\cdot\mathbf{r}_2)R^2}{2R^5} - \frac{\cancel{3\,(\mathbf{R}\cdot\mathbf{r}_1)^2} - \cancel{r_1^2R^2}}{2R^5}$$
$$- \frac{\cancel{3\,(\mathbf{R}\cdot\mathbf{r}_2)^2} - \cancel{r_2^2R^2}}{2R^5} = \frac{\mathbf{r}_1\cdot\mathbf{r}_2}{R^3} - 3\frac{(\mathbf{R}\cdot\mathbf{r}_2)(\mathbf{R}\cdot\mathbf{r}_1)}{R^5}.$$

b) It is worth recalling that $P_k^2 = 1$, therefore P_k is both hermitian and its own inverse, $P_k^\dagger = P_k^{-1} = P_k$. Since $P_k^{-1}FP_k = (-1)^{\pi_{F_k}}F$, we have

$$\langle\psi_f|F|\psi_i\rangle = (-1)^{\pi_{F_k}}\langle\psi_f|P_k^{-1}FP_k|\psi_i\rangle = (-1)^{\pi_{F_k}}\langle P_k\psi_f \,|\, F \,|\, P_k\psi_i\rangle$$
$$= (-1)^{\pi_{F_k}}(-1)^{\pi_{f_k}}(-1)^{\pi_{i_k}}\langle\psi_f \,|\, F \,|\, \psi_i\rangle = (-1)^{\pi_{F_k}+\pi_{f_k}+\pi_{i_k}}\langle\psi_f \,|\, F \,|\, \psi_i\rangle\,.$$

This equation means that either $(-1)^{\pi_{F_k}+\pi_{f_k}+\pi_{i_k}} = 1$ or $\langle\psi_f|F|\psi_i\rangle = 0$, which is what was to be proven.
It is now straightforward to see that, since

$$\mathcal{H}'(-\mathbf{r}_1,\mathbf{r}_2) = -\mathcal{H}'(\mathbf{r}_1,\mathbf{r}_2), \qquad \text{and} \qquad \psi_0^{(0)}(-\mathbf{r}_1,\mathbf{r}_2) = \psi_0^{(0)}(\mathbf{r}_1,\mathbf{r}_2),$$

the relation we have just proved implies $\langle\psi_0^{(0)}|\mathcal{H}'|\psi_0^{(0)}\rangle = 0$.

c) Let us assume that the electron with $n = 1$ (and consequently $\ell = 0$ and $m = 0$) is the first one. Since $P_1^{-1}\mathcal{H}'P_1 = -\mathcal{H}'$ and $P_1\,|\phi_{100}\rangle = |\phi_{100}\rangle$,

$$\langle\psi_n^{(0)}|\mathcal{H}'|\psi_0^{(0)}\rangle = \langle\phi_{100}\phi_{n\ell m}|\mathcal{H}'|\phi_{100}\phi_{100}\rangle = (-1)\,\langle\phi_{100}\phi_{n\ell m}|P_1^{-1}\mathcal{H}'P_1|\phi_{100}\phi_{100}\rangle$$
$$= (-1)\,\langle P_1\phi_{100}\phi_{n\ell m} \,|\, \mathcal{H}' \,|\, P_1\phi_{100}\phi_{100}\rangle = (-1)\,\langle\phi_{100}\phi_{n\ell m} \,|\, \mathcal{H}' \,|\, \phi_{100}\phi_{100}\rangle$$
$$= -\,\langle\psi_n^{(0)}|\mathcal{H}'|\psi_0^{(0)}\rangle\,.$$

We have arrived at the relation $\langle\psi_n^{(0)}|\mathcal{H}'|\psi_0^{(0)}\rangle = -\langle\psi_n^{(0)}|\mathcal{H}'|\psi_0^{(0)}\rangle$, therefore we conclude that $\langle\psi_n^{(0)}|\mathcal{H}'|\psi_0^{(0)}\rangle = 0$. The case when the electron staying in the same orbital is the second one is proven in the same way.

d) The second order correction in perturbation theory is given by

$$\Delta E_0^{(2)} = \sum_{k\neq 0} \frac{\left|\left\langle \psi_k^{(0)} \middle| \mathcal{H}' \middle| \psi_0^{(0)} \right\rangle\right|^2}{E_0^{(0)} - E_k^{(0)}},$$

where k stands for all quantum numbers of the states $\{n_1, \ell_1, m_1, n_2, \ell_2, m_2\}$. As discussed for the Stark effect, a lower limit for $\Delta E_0^{(2)}$ can be obtained replacing the energy differences $E_0^{(0)} - E_k^{(0)}$ by $E_0^{(0)} - E_1^{(0)}$. However, notice that using the result of part b) it can be seen that the matrix elements for the first excited states, those with $n_1 = 1$, $n_2 = 2$ or $n_1 = 2$, $n_2 = 1$, are zero. Therefore, the lowest energy states contributing to the sum are those with $n_1 = n_2 = 2$ (labelled with the subscript e) for which,

$$E_e^{(0)} = E_0^{(0)} \frac{1}{4},$$

where, since there are two hydrogen atoms at play,

$$E_0^{(0)} = -\frac{e^2}{a_0} = -2 \text{ Ry}.$$

Thus

$$\Delta E_0^{(2)} = \sum_{k\neq 0} \frac{\left|\left\langle \psi_k^{(0)} \middle| \mathcal{H}' \middle| \psi_0^{(0)} \right\rangle\right|^2}{E_0^{(0)} - E_k^{(0)}} > \frac{1}{E_0^{(0)} - E_e^{(0)}} \sum_{k\neq 0} \left|\left\langle \psi_k^{(0)} \middle| \mathcal{H}' \middle| \psi_0^{(0)} \right\rangle\right|^2$$
$$= \frac{4}{3E_0^{(0)}} \sum_{k\neq 0} \left|\left\langle \psi_k^{(0)} \middle| \mathcal{H}' \middle| \psi_0^{(0)} \right\rangle\right|^2.$$

To perform the summation we take into account that $\left\langle \psi_0^{(0)} \middle| \mathcal{H}' \middle| \psi_0^{(0)} \right\rangle = 0$ and use the completeness relation of the eigenstates of the system to obtain,

$$\sum_{k\neq 0} \left|\left\langle \psi_k^{(0)} \middle| \mathcal{H}' \middle| \psi_0^{(0)} \right\rangle\right|^2 = \sum_{k} \left|\left\langle \psi_k^{(0)} \middle| \mathcal{H}' \middle| \psi_0^{(0)} \right\rangle\right|^2$$
$$= \sum_k \left\langle \psi_0^{(0)} \middle| \mathcal{H}' \middle| \psi_k^{(0)} \right\rangle \left\langle \psi_k^{(0)} \middle| \mathcal{H}' \middle| \psi_0^{(0)} \right\rangle = \left\langle \psi_0^{(0)} \middle| (\mathcal{H}')^2 \middle| \psi_0^{(0)} \right\rangle.$$

Seeing that

$$(\mathcal{H}')^2 \approx \left(\frac{e^2}{R^3}\right)^2 \left((\mathbf{r}_1 \cdot \mathbf{r}_2)^2 + 9 \left(\mathbf{r}_1 \cdot \hat{\mathbf{R}}\right)^2 \left(\mathbf{r}_2 \cdot \hat{\mathbf{R}}\right)^2 - 6 (\mathbf{r}_1 \cdot \mathbf{r}_2) \left(\mathbf{r}_1 \cdot \hat{\mathbf{R}}\right) \left(\mathbf{r}_2 \cdot \hat{\mathbf{R}}\right) \right),$$

let us compute

$$\langle\psi_0^{(0)}| (\mathbf{r}_1 \cdot \mathbf{r}_2)^2 |\psi_0^{(0)}\rangle = \langle\phi_{1s}\phi_{1s}|x_1^i x_2^i x_1^j x_2^j|\phi_{1s}\phi_{1s}\rangle$$
$$= \langle\phi_{1s}|x_1^i x_1^j|\phi_{1s}\rangle \langle\phi_{1s}|x_2^i x_2^j|\phi_{1s}\rangle = \frac{\langle r^2\rangle}{3}\delta_{ij}\frac{\langle r^2\rangle}{3}\delta_{ij} = \frac{\left(\langle r^2\rangle\right)^2}{9}\delta_{ii} = \frac{\left(\langle r^2\rangle\right)^2}{3},$$

$$\left\langle \psi_0^{(0)} \middle| \left(\mathbf{r}_1 \cdot \hat{\mathbf{R}}\right)^2 \left(\mathbf{r}_2 \cdot \hat{\mathbf{R}}\right)^2 \middle| \psi_0^{(0)} \right\rangle = \langle\psi_0^{(0)}|x_1^i \hat{X}^i x_1^j \hat{X}^j x_2^k \hat{X}^k x_2^m \hat{X}^m|\psi_0^{(0)}\rangle$$
$$= \hat{X}^i \hat{X}^j \hat{X}^k \hat{X}^m \langle\phi_{1s}|x^i x^j|\phi_{1s}\rangle \langle\phi_{1s}|x^k x^m|\phi_{1s}\rangle$$
$$= \hat{X}^i \hat{X}^j \hat{X}^k \hat{X}^m \frac{\langle r^2\rangle}{3}\delta_{ij}\frac{\langle r^2\rangle}{3}\delta_{km} = \left(\hat{\mathbf{R}}\right)^2 \left(\hat{\mathbf{R}}\right)^2 \frac{\left(\langle r^2\rangle\right)^2}{9} = \frac{\left(\langle r^2\rangle\right)^2}{9},$$

$$\left\langle \psi_0^{(0)} \left| (\mathbf{r}_1 \cdot \mathbf{r}_2) \left(\mathbf{r}_1 \cdot \hat{\mathbf{R}} \right) \left(\mathbf{r}_2 \cdot \hat{\mathbf{R}} \right) \right| \psi_0^{(0)} \right\rangle = \langle \psi_0^{(0)} | x_1^i x_2^i x_1^j \hat{X}^j x_2^k \hat{X}^k | \psi_0^{(0)} \rangle =$$

$$\hat{X}^j \hat{X}^k \langle \phi_{1s} | x^i x^j | \phi_{1s} \rangle \langle \phi_{1s} | x^i x^k | \phi_{1s} \rangle = \hat{X}^j \hat{X}^k \frac{\langle r^2 \rangle}{3} \delta_{ij} \frac{\langle r^2 \rangle}{3} \delta_{ik} = \frac{\left(\langle r^2 \rangle \right)^2}{9}.$$

Therefore

$$\left\langle \psi_0^{(0)} \left| (\mathcal{H}')^2 \right| \psi_0^{(0)} \right\rangle = \frac{2}{3} \left(\frac{e^2}{R^3} \right)^2 \left(\langle r^2 \rangle \right)^2 = 6 \left(\frac{e^2 a_0^2}{R^3} \right)^2,$$

where we have used $\langle r^2 \rangle = 3a_0^2$, which can be obtained using formulae quoted in chapter 1. We can now see that (remember $E_0^{(0)} = -e^2/a_0$),

$$\frac{4}{3E_0^{(0)}} \sum_{k \neq 0} \left| \left\langle \psi_k^{(0)} \left| \mathcal{H}' \right| \psi_0^0 \right\rangle \right|^2 = \frac{4}{3E_0^{(0)}} 6 \left(\frac{e^2 a_0^2}{R^3} \right)^2 = 8 \left(\frac{a_0}{R} \right)^6 E_0^{(0)} < \Delta E_0^{(2)}.$$

2. The lowest excited states of helium have principal quantum numbers $n_1 = 1$, $n_2 = 2$ or $n_1 = 2$, $n_2 = 1$, for electron 1 and 2, respectively. We will treat the electrostatic repulsion between the electrons, $\mathcal{H}' = e^2/r_{12}$, as a perturbation to the unperturbed energies given by

$$E_{n_1,n_2} = -4 \left[\frac{1}{2} \alpha^2 mc^2 \left(\frac{1}{n_1^2} + \frac{1}{n_2^2} \right) \right] = -5 \left(\frac{1}{2} \alpha^2 mc^2 \right) = -68.03\,\text{eV}.$$

Since the $n = 2$ levels of a hydrogenlike atom are fourfold degenerate (we don't take into account the spin degree of freedom in this problem), the first excited unperturbed energy level in this simplified treatment of helium is eightfold degenerate. The eight unperturbed wave functions can be written as

$$\begin{aligned}
\phi_1^{(0)} &= 1s(1)\,2s(2) & \phi_5^{(0)} &= 1s(1)\,2p_y(2) \\
\phi_2^{(0)} &= 2s(1)\,1s(2) & \phi_6^{(0)} &= 2p_y(1)\,1s(2) \\
\phi_3^{(0)} &= 1s(1)\,2p_x(2) & \phi_7^{(0)} &= 1s(1)\,2p_z(2) \\
\phi_4^{(0)} &= 2p_x(1)\,1s(2) & \phi_8^{(0)} &= 2p_z(1)\,1s(2)
\end{aligned}$$

where $1s(1)\,2s(2)$ means the product of a hydrogenlike $1s$ function for electron 1 and a hydrogenlike $2s$ function for electron 2, etc. Note that we are using real hydrogenlike p orbitals, defined as $p_x = \frac{1}{\sqrt{2}}(p_{+1} + p_{-1})$, $p_y = \frac{1}{i\sqrt{2}}(p_{+1} - p_{-1})$ and $p_z = p_0$. The secular determinant,

$$\left| \mathcal{H}'_{ij} - E\delta_{ij} \right| = 0, \qquad \text{where} \qquad \mathcal{H}'_{ij} = \left\langle \phi_i^{(0)} \left| \mathcal{H}' \right| \phi_j^{(0)} \right\rangle$$

contains 64 elements.

(a) Using the hermiticity of $\mathcal{H}'$, the real condition of the orbitals and parity considerations, prove that the only non-zero $\mathcal{H}'_{ij}$ matrix elements are the diagonal ones and $\mathcal{H}'_{12} = \mathcal{H}'_{21}$, $\mathcal{H}'_{34} = \mathcal{H}'_{43}$, $\mathcal{H}'_{56} = \mathcal{H}'_{65}$, $\mathcal{H}'_{78} = \mathcal{H}'_{87}$. Therefore the secular determinant is block-diagonal and factorizes into four determinants of second order.

(b) The first diagonal matrix element of $\mathcal{H}'$ is an example of Coulomb integral, describing the electrostatic repulsion between an electron with probability density function $[1s]^2$ and an electron with probability density function $[2s]^2$,

$$\mathcal{H}'_{11} = \int 1s(1)\,2s(2)\,\frac{e^2}{r_{12}}\,1s(1)\,2s(2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 \equiv J_{1s2s}.$$

The integral corresponding to $\mathcal{H}'_{12}$, denoted as K_{1s2s}, is an example of exchange integral,

$$\mathcal{H}'_{12} = \int 1s(1)\,2s(2)\,\frac{e^2}{r_{12}}\,2s(1)\,1s(2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 \equiv K_{1s2s}.$$

Compute J_{1s2s} and K_{1s2s}. To do the integrals use the Laplace expansion of $1/r_{12}$

$$\frac{1}{r_{12}} = \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} = 4\pi \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \frac{1}{2\ell+1} \frac{r_<^\ell}{r_>^{\ell+1}} Y_\ell^{m*}(\theta_1,\varphi_1) Y_\ell^m(\theta_2,\varphi_2),$$

where $r_<$ means the smaller of r_1 and r_2, and $r_>$ the larger. Use also the orthogonality relations of the spherical harmonics (hint: multiply and divide inside the angular integrals by $Y_0^0 Y_0^{0*} = 1/4\pi$). You should obtain the following results

$$J_{1s2s} = \frac{17}{81}\frac{Ze^2}{a_0}, \qquad K_{1s2s} = \frac{16}{729}\frac{Ze^2}{a_0}.$$

(c) Find the eigenvalues corresponding to the 2×2 determinant of the first block and use them to compute, in eV, the energy of the first two levels up to first-order corrections, i.e., $E_1 = E_1^{(0)} + E_1^{(1)}$ and $E_2 = E_2^{(0)} + E_2^{(1)}$, where $E_1^{(0)} = E_2^{(0)} = -68.03\,\mathrm{eV}$. Find also the corresponding eigenfunctions.

Solution:

(a) From the hermiticity condition of $\mathcal{H}'$ it follows that $\mathcal{H}'^*_{ij} = \mathcal{H}'_{ji}$. This, together with the fact that $\mathcal{H}'$ is real and that we are using real orbitals, implies that $\mathcal{H}'_{ij} = \mathcal{H}'_{ji}$, therefore the secular determinant is symmetric about the principal diagonal.

To study in more detail the matrix elements we should first stablish the parity transformation properties of the different orbitals and operators under consideration. First, all s hydrogenlike orbitals are even under $\mathbf{r} \to -\mathbf{r}$, as well as the operator e^2/r_{12}. Furthermore, the hydrogenlike $2p_x$, $2p_y$ and $2p_z$ orbitals are given by

$$\phi_{2p_x} = \frac{1}{4\sqrt{2\pi}}\left(\frac{Z}{a_0}\right)^{5/2} x\, e^{-Zr/2a_0}, \qquad \phi_{2p_y} = \frac{1}{4\sqrt{2\pi}}\left(\frac{Z}{a_0}\right)^{5/2} y\, e^{-Zr/2a_0},$$

$$\phi_{2p_z} = \frac{1}{\sqrt{\pi}}\left(\frac{Z}{a_0}\right)^{5/2} z\, e^{-Zr/2a_0}.$$

Therefore they are odd under parity transformations. Having this in mind, let's consider for instance the $\mathcal{H}'_{13}$ matrix element,

$$\mathcal{H}'_{13} = \int \phi_1^{(0)}\,\frac{e^2}{r_{12}}\,\phi_3^{(0)})\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2 = \int 1s(1)\,2s(2)\,\frac{e^2}{r_{12}}\,2s(1)\,2p_x(2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2.$$

Since one term in the integrand is odd under parity transformations (the $2p_x$ orbital), and the remaining ones are even, we conclude that the integrand itself is odd and $\mathcal{H}'_{13} = 0$. The same reasoning gives $\mathcal{H}'_{14} = \mathcal{H}'_{15} = \mathcal{H}'_{16} = \mathcal{H}'_{17} = \mathcal{H}'_{18} = 0$ and $\mathcal{H}'_{23} = \mathcal{H}'_{24} = \mathcal{H}'_{25} = \mathcal{H}'_{26} = \mathcal{H}'_{27} = \mathcal{H}'_{28} = 0$.

Now consider $\mathcal{H}'_{35}$,

$$\mathcal{H}'_{35} = \int 1s(1)\,2p_x(2)\,\frac{e^2}{r_{12}}\,1s(1)\,2p_y(2)\,\mathrm{d}\tau_1\,\mathrm{d}\tau_2.$$

If we invert the x coordinates $x_1 \to -x_1$ and $x_2 \to -x_2$, the $1s(1)$ and $2p_y(2)$ orbitals are unaffected, as well as e^2/r_{12}, but the orbital $2p_x(2)$ changes sign, therefore the integrand is odd and $\mathcal{H}'_{35} = 0$. Likewise $\mathcal{H}'_{36} = \mathcal{H}'_{37} = \mathcal{H}'_{38} = 0$ and $\mathcal{H}'_{45} = \mathcal{H}'_{46} = \mathcal{H}'_{47} = \mathcal{H}'_{48} = 0$. Finally, by considering the transformations $y_1 \to -y_1$ and $y_2 \to -y_2$, we see that $\mathcal{H}'_{57} = \mathcal{H}'_{58} = \mathcal{H}'_{67} = \mathcal{H}'_{68} = 0$. Note that, for an integral to be zero, one does not need to require the integrand to

be an odd function of *all variables*. For example, let f be a function of m variables such that it is odd only on a number k of them,

$$f(-x_1, -x_2, \ldots, -x_k, x_{k+1}, \ldots, x_m) = -f(x_1, x_2, \ldots, x_k, x_{k+1}, \ldots, x_m).$$

Then we have

$$\int \mathrm{d}x_1 \ldots \mathrm{d}x_m \, f(x_1, \ldots x_m) = \int \underbrace{\left[\int \mathrm{d}x_1 \ldots \mathrm{d}x_k \, f(x_1, \ldots q_k) \right]}_{=0} \mathrm{d}x_{k+1} \ldots \mathrm{d}x_m = 0.$$

We can write the secular equation as

$$\begin{vmatrix} b_{11} & \mathcal{H}'_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathcal{H}'_{12} & b_{22} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & b_{33} & \mathcal{H}'_{34} & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathcal{H}'_{34} & b_{44} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & b_{55} & \mathcal{H}'_{56} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathcal{H}'_{56} & b_{66} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & b_{77} & \mathcal{H}'_{78} \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathcal{H}'_{78} & b_{88} \end{vmatrix} = 0,$$

with $b_{ii} = \mathcal{H}'_{ii} - E^{(1)}$, $i = 1, 2, \ldots 8$, and $E^{(1)}$ the first-order correction to the energy.

(b) We have (setting $a_0 \equiv a$)

$$J_{1s2s} = \int 1s(1)\, 2s(2)\, \frac{e^2}{r_{12}}\, 1s(1)\, 2s(2)\, \mathrm{d}\tau_1\, \mathrm{d}\tau_2 = \int [1s(1)]^2\, [2s(2)]^2\, \frac{e^2}{r_{12}}\, \mathrm{d}\tau_1\, \mathrm{d}\tau_2.$$

$$[1s(1)]^2 = \left[\left(\frac{Z}{a} \right)^{3/2} \frac{1}{\sqrt{\pi}}\, e^{-Zr_1/a} \right]^2 = \frac{Z^3}{\pi a^3} e^{-2Zr_1/a}.$$

$$[2s(2)]^2 = \left[\left(\frac{Z}{a} \right)^{3/2} \frac{1}{\sqrt{32\pi}} \left(2 - \frac{Zr_2}{a} \right) e^{-Zr_2/2a} \right]^2 = \frac{Z^3}{32\pi a^3} \left(2 - \frac{Zr_2}{a} \right)^2 e^{-Zr_2/a}.$$

$$J_{1s2s} = \frac{Z^6}{32\pi^2 a^6} \int \mathrm{d}\tau_2\, \mathrm{d}\tau_1\, \frac{e^2}{r_{12}} \left(2 - \frac{Zr_2}{a} \right)^2 e^{-Zr_2/a}\, e^{-2Zr_1/a} =$$

$$\frac{e^2 Z^6}{32\pi^2 a^6} \int \mathrm{d}\tau_2\, \mathrm{d}\tau_1 \left(2 - \frac{Zr_2}{a} \right)^2 e^{-Zr_2/a}\, e^{-2Zr_1/a} \sum_{\ell,m} \frac{4\pi}{2\ell+1} \frac{r_<^\ell}{r_>^{\ell+1}} Y_\ell^{m*}(\theta_1, \varphi_1) Y_\ell^m(\theta_2, \varphi_2).$$

To use the orthonormal property of the spherical harmonics,

$$\int \mathrm{d}\theta\, \mathrm{d}\varphi\, \sin\theta\, Y_\ell^{m*}(\theta, \varphi)\, Y_{\ell'}^{m'}(\theta, \varphi) = \delta_{\ell\ell'} \delta_{mm'},$$

we multiply and divide the integrand of J_{1s2s} by $Y_0^0 Y_0^{0*} = 1/4\pi$. Thus

$$J_{1s2s} = \frac{e^2 Z^6}{32\pi^2 a^6} \frac{1}{Y_0^0 Y_0^{0*}} \sum_{\ell,m} \int \mathrm{d}r_2\, \mathrm{d}r_1\, r_2^2 r_1^2 \left(2 - \frac{Zr_2}{a} \right)^2 e^{-Zr_2/a}\, e^{-2Zr_1/a} \frac{r_<^\ell}{r_>^{\ell+1}}$$

$$\times \frac{4\pi}{2\ell+1} \underbrace{\int Y_\ell^{m*}(\theta_1, \varphi_1)\, Y_0^0 \mathrm{d}\Omega_1}_{\delta_{\ell 0}\delta_{m0}} \underbrace{\int Y_0^{0*}\, Y_\ell^m(\theta_2, \varphi_2)\, \mathrm{d}\Omega_2}_{\delta_{\ell 0}\delta_{m0}}$$

$$= \frac{e^2 Z^6}{2a^6} \int \mathrm{d}r_2\, \mathrm{d}r_1\, r_2^2 r_1^2\, \frac{1}{r_>} \left(2 - \frac{Zr_2}{a} \right)^2 e^{-Zr_2/a}\, e^{-2Zr_1/a}.$$

We can integrate first over r_1, taking care that in the range $0 \le r_1 \le r_2$ we should set $r_> = r_2$, and in the range $r_2 \le r_1 \le \infty$ we have $r_> = r_1$. Thus

$$J_{1s2s} = \frac{e^2 Z^6}{2a^6} \int_0^\infty \mathrm{d}r_2\, r_2^2 \left(2 - \frac{Zr_2}{a} \right)^2 e^{-Zr_2/a} \left[\int_0^{r_2} \mathrm{d}r_1\, \frac{r_1^2}{r_2}\, e^{-2Zr_1/a} + \int_{r_2}^\infty \mathrm{d}r_1\, \frac{r_1^2}{r_1}\, e^{-2Zr_1/a} \right]$$

The integrals can now be done in a straightforward way using the following primitives,

$$\int \mathrm{d}x\, x e^{bx} = \left(\frac{x}{b} - \frac{1}{b^2}\right) e^{bx}, \qquad \int \mathrm{d}x\, x^2 e^{bx} = \left(\frac{x^2}{b} - \frac{2x}{b^2} + \frac{2}{b^3}\right) e^{bx}.$$

The result is

$$J_{1s2s} = \frac{17}{81}\frac{Ze^2}{a} = 11.42\,\mathrm{eV}.$$

The K_{1s2s} integral is done in a completely similar way and the result is

$$K_{1s2s} = \frac{16}{729}\frac{Ze^2}{a} = 1.19\mathrm{eV}.$$

(c) The 2×2 first block of the secular equation gives (note that $J_{1s2s} = J_{2s1s}$)

$$\begin{vmatrix} b_{11} & \mathcal{H}'_{12} \\ \mathcal{H}'_{12} & b_{22} \end{vmatrix} = \begin{vmatrix} J_{1s2s} - E & K_{1s2s} \\ K_{1s2s} & J_{2s1s} - E \end{vmatrix} = (J_{1s2s} - E)^2 - K_{1s2s}^2 = 0.$$

The first-order corrections to levels E_1 and E_2 are then

$$E_1^{(1)} = J_{1s2s} - K_{1s2s}, \qquad E_2^{(1)} = J_{1s2s} + K_{1s2s}.$$

The eigenfunctions corresponding to these eigenvalues are found doing (calling $J_{1s2s} \equiv J$ and $K_{1s2s} \equiv K$ to shorten the notation)

$$\begin{pmatrix} J & K \\ K & J \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = (J - K) \begin{pmatrix} a \\ b \end{pmatrix} \implies a = -b.$$

$$\begin{pmatrix} J & K \\ K & J \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = (J + K) \begin{pmatrix} a \\ b \end{pmatrix} \implies a = b.$$

Therefore the normalized zero-order eigenfunctions for which the perturbation is diagonal are

$$\psi_1^{(0)} = \frac{1}{\sqrt{2}}\left(\phi_1^{(0)} - \phi_2^{(0)}\right), \qquad E_1 = E_1^{(0)} + E_1^{(1)} = E_1^{(0)} + J_{1s2s} - K_{1s2s},$$

$$\psi_2^{(0)} = \frac{1}{\sqrt{2}}\left(\phi_1^{(0)} + \phi_2^{(0)}\right), \qquad E_2 = E_2^{(0)} + E_2^{(1)} = E_1^{(0)} + J_{1s2s} + K_{1s2s}.$$

A.4 Chapter 4.

1. The photoelectric effect is the emission of electrons from a material caused by electromagnetic radiation. Electrons emitted in this manner are called photoelectrons. In this assignment we shall obtain the photoelectric effect cross section for a $1s$ hydrogen-like state.

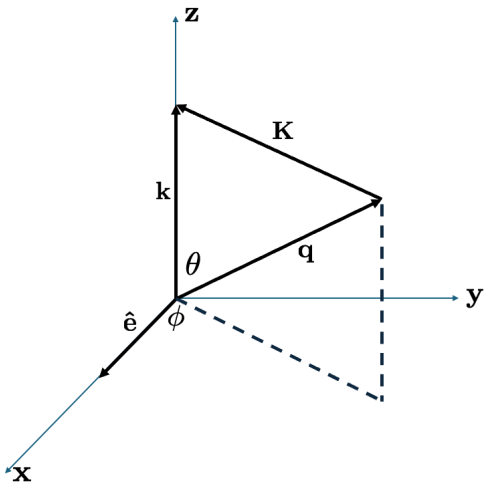
   a) Consider a hydrogen-like atom in a $1s$ initial state: $|i\rangle = \frac{Z^{3/2}}{\sqrt{\pi a_0^3}} e^{-Zr/a_0}$. Consider a final state of a free electron: $|f\rangle = \frac{1}{\sqrt{V}} e^{i\mathbf{q}\cdot\mathbf{r}}$, where the electron momentum is $\mathbf{p} = \hbar\mathbf{q}$. Show that the matrix element involved in photon absorption is

   $$\langle f|e^{i\mathbf{k}\cdot\mathbf{r}}\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p}|i\rangle = \frac{8\sqrt{\pi}(a_0 Z)^{3/2} Z\hbar}{\sqrt{V}(Z^2 + a_0^2 K^2)^2}\,(\hat{\mathbf{e}}_{\mathbf{k}}\cdot\mathbf{q})\,,$$

   where $\mathbf{K} = \mathbf{k} - \mathbf{q}$.

   b) Use the reference frame shown in this figure to prove the relations

   $$\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{q} = q\sin\theta\cos\phi,$$
   $$K^2 = k^2 + q^2 - 2kq\cos\theta.$$



   c) If the photon energy is large compared with the binding energy of the electron in the atom, but small compared to the rest mass of the electron, the final state is that of a free electron and the necessity of a relativistic treatment is not required. Both conditions are satisfied if $ka_0 \approx Z$. Under these assumptions show that the energy of the absorbed photon is

   $$E = Z\alpha mc^2,$$

   where α is the fine-structure constant. This shows that the relation $E \ll mc^2$ holds, provided Z is not very high.

   d) Check also that $Z \approx ka_0 \ll qa_0$. To do this you need to assume that the ionization energy is negligible compared to the energy of the photon, therefore the kinetic energy of the free electron is equal to the photon energy: $\hbar ck = mv^2/2$.

   e) Use the result from d) to demonstrate that

   $$Z^2 + a_0^2 K^2 \approx q^2 a_0^2(1 - \beta\cos\theta), \qquad (\beta = v/c),$$

   and conclude that the result of a) can be approximated by:

   $$\langle f|e^{i\mathbf{k}\cdot\mathbf{r}}\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p}|i\rangle = \frac{8\sqrt{\pi}(a_0 Z)^{3/2} Z\hbar q\sin\theta\cos\varphi}{\sqrt{V}\left[q^2 a_0^2(1 - \beta\cos\theta)\right]^2}.$$

f) The complete interaction matrix element is

$$T_{if} = \frac{e}{m}\sqrt{\frac{2\pi\hbar n}{\omega V}}\,\langle f|e^{i\mathbf{k}\cdot\mathbf{r}}\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p}|i\rangle\,,$$

and the density of initial photons is

$$\frac{\mathrm{d}N}{\mathrm{d}E} = \frac{V}{(2\pi)^3}\frac{mq}{\hbar^2}\,\mathrm{d}\Omega,$$

where Ω is the solid angle of the absorbed photon and n is the number of photons of energy $\hbar ck$.

According to Fermi's golden rule, the probability per unit time for photoelectric absorption (W) and the differential absorption cross section are:

$$\mathrm{d}W = \frac{2\pi}{\hbar}|T_{if}|^2\frac{\mathrm{d}N}{\mathrm{d}E}, \qquad \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{V}{nc}\frac{\mathrm{d}W}{\mathrm{d}\Omega}.$$

Show that

$$\mathrm{d}W = \frac{n}{V}\frac{e^2}{mc}\frac{32Z^5\sin^2\theta\cos^2\varphi}{q^5a_0^5k(1-\beta\cos\theta)^4}\,\mathrm{d}\Omega, \qquad \frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{e^2}{mc^2}\frac{32Z^5}{q^5a_0^5k}\frac{\sin^2\theta\cos^2\varphi}{(1-\beta\cos\theta)^4}.$$

Solution:

a) We need to compute

$$\langle f|e^{i\mathbf{k}\cdot\mathbf{r}}\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p}|i\rangle = \int\frac{1}{\sqrt{V}}e^{-i\mathbf{q}\cdot\mathbf{r}}e^{i\mathbf{k}\cdot\mathbf{r}}\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p}\frac{Z^{3/2}}{\sqrt{\pi a_0^3}}e^{-Zr/a_0}\mathrm{d}\mathbf{r}$$

$$= \frac{Z^{3/2}}{\sqrt{V\pi a_0^3}}(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p})\int e^{i\mathbf{K}\cdot\mathbf{r}}e^{-Zr/a_0}\mathrm{d}\mathbf{r} = \frac{Z^{3/2}}{\sqrt{V\pi a_0^3}}(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p})\int e^{iKr\cos\theta}e^{-Zr/a_0}r^2\sin\theta\mathrm{d}\varphi\mathrm{d}\theta\mathrm{d}r.$$

The integral in φ is 2π whereas

$$\int_0^\pi e^{iKr\cos\theta}\sin\theta\mathrm{d}\theta = \int_{-1}^1 e^{iKrx}\mathrm{d}x = \left.\frac{e^{iKrx}}{iKr}\right]_{-1}^1 = 2\frac{\sin(Kr)}{Kr}.$$

Therefore,

$$\langle f|e^{i\mathbf{k}\cdot\mathbf{r}}\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p}|i\rangle = \frac{4\pi}{K}\frac{Z^{3/2}}{\sqrt{V\pi a_0^3}}(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p})\int_0^\infty\sin(Kr)e^{-Zr/a_0}r\,\mathrm{d}r = \frac{4\pi}{K}\frac{Z^{3/2}}{\sqrt{V\pi a_0^3}}(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p})$$

$$\left[-\frac{a_0\mathrm{e}^{-\frac{Zr}{a_0}}\left(\left(Z\left(K^2a_0^2+Z^2\right)r-K^2a_0^3+Z^2a_0\right)\sin\left(Kr\right)+Ka_0\left(K^2a_0^2r+Z^2r+2Za_0\right)\cos\left(Kr\right)\right)}{\left(K^2a_0^2+Z^2\right)^2}\right]_0^\infty$$

$$= \frac{8\sqrt{\pi}(a_0Z)^{3/2}Z}{\sqrt{V}\left(K^2a_0^2+Z^2\right)^2}(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p}) = \frac{8\sqrt{\pi}(a_0Z)^{3/2}Z\hbar}{\sqrt{V}\left(K^2a_0^2+Z^2\right)^2}(\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{q}).$$

b) With the suggested reference frame:

$$\hat{\mathbf{e}}_{\mathbf{k}\lambda} = \hat{\imath},$$
$$\mathbf{q} = q\sin\theta\cos\varphi\hat{\imath} + q\sin\theta\sin\varphi\hat{\jmath} + q\cos\theta\hat{\mathbf{k}}.$$

Therefore,

$$\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{q} = q\sin\theta\cos\varphi.$$

Also

$$\mathbf{K} = \mathbf{k}-\mathbf{q} \implies K^2 = (\mathbf{k}-\mathbf{q})^2 = k^2+q^2-2\mathbf{k}\cdot\mathbf{q}.$$

Since $\mathbf{k}$ is aligned with the z-axis $\mathbf{k}\cdot\mathbf{q} = kq\cos\theta$.

c) We have that $Z \approx ka_0$. Since the energy of the photon is

$$E = \hbar ck \approx \hbar c \frac{Z}{a_0} = \hbar c \frac{Z}{\hbar/(mc\alpha)} = mc^2 Z\alpha.$$

d) Since $E = \hbar ck = \frac{mv^2}{2}$ and $mv = p = \hbar q$, we have that

$$\hbar ck = E = \frac{mv^2}{2} = \frac{pv}{2} = \hbar q \frac{v}{2} \implies \frac{k}{q} = \frac{v}{2c} \ll 1.$$

Also,

$$k \approx Z/a_0 \implies \frac{Z}{qa_0} \approx \frac{k}{q} \ll 1 \implies Z \ll qa_0.$$

e) We see that

$$\begin{aligned} Z^2 + a_0^2 K^2 &= Z^2 + a_0^2(k^2 + q^2 - 2kq\cos\theta) \approx a_0^2(q^2 - 2kq\cos\theta) = q^2 a_0^2 \left(1 - 2\frac{k}{q}\cos\theta\right) \\ &= q^2 a_0^2 \left(1 - 2\frac{mv^2/(2\hbar c)}{mv/\hbar}\cos\theta\right) = q^2 a_0^2 \left(1 - \frac{v}{c}\cos\theta\right) = q^2 a_0^2 \left(1 - \beta\cos\theta\right). \end{aligned}$$

With the result of b) and this last result, the result of a) can be worked out as

$$\langle f|e^{i\mathbf{k}\cdot\mathbf{r}}\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p}|i\rangle = \frac{8\sqrt{\pi}(a_0 Z)^{3/2} Z\hbar}{\sqrt{V}(Z^2 + a_0^2 K^2)^2}(\hat{\mathbf{e}}_{\mathbf{k}}\cdot\mathbf{q}) = \frac{8\sqrt{\pi}(a_0 Z)^{3/2} Z\hbar q \sin\theta\cos\varphi}{\sqrt{V}[q^2 a_0^2(1 - \beta\cos\theta)]^2}.$$

f) The only thing to do is to introduce the matrix element and the density of states in the probability of absorption:

$$\begin{aligned} \mathrm{d}W = \frac{2\pi}{\hbar}|T_{if}|^2 \frac{\mathrm{d}N}{\mathrm{d}E} = \frac{2\pi}{\hbar}\left|\frac{e}{m}\sqrt{\frac{2\pi\hbar n}{\omega V}}\langle f|e^{i\mathbf{k}\cdot\mathbf{r}}\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p}|i\rangle\right|^2 \frac{V}{(2\pi)^3}\frac{mq}{\hbar^2}\mathrm{d}\Omega \\ = \frac{e^2 nq}{2\pi m\omega\hbar^2}\left|\langle f|e^{i\mathbf{k}\cdot\mathbf{r}}\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p}|i\rangle\right|^2 \mathrm{d}\Omega. \end{aligned}$$

Now we use the result of e) to get

$$\begin{aligned} \mathrm{d}W = \frac{e^2 nq}{2\pi m\omega\hbar^2}\left|\langle f|e^{i\mathbf{k}\cdot\mathbf{r}}\hat{\mathbf{e}}_{\mathbf{k}\lambda}\cdot\mathbf{p}|i\rangle\right|^2 \mathrm{d}\Omega = \frac{e^2 nq}{2\pi m\omega\hbar^2}\left|\frac{8\sqrt{\pi}(a_0 Z)^{3/2} Z\hbar q\sin\theta\cos\varphi}{\sqrt{V}[q^2 a_0^2(1-\beta\cos\theta)]^2}\right|^2 \mathrm{d}\Omega \\ = \frac{n}{V}\frac{e^2}{mc}\frac{32Z^5\sin^2\theta\cos^2\varphi}{q^5 a_0^5 k(1-\beta\cos\theta)^4}\mathrm{d}\Omega. \end{aligned}$$

In the last step we have used that $\omega = ck$. The differential cross-section is then

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{V}{nc}\frac{\mathrm{d}W}{\mathrm{d}\Omega} = \frac{e^2}{mc^2}\frac{32Z^5}{q^5 a_0^5 k}\frac{\sin^2\theta\cos^2\varphi}{(1-\beta\cos\theta)^4}.$$

2. In this exercise we will develop a model for the interaction of a two-level atom with the radiation field. The atomic energy levels are denoted as $|-\rangle$ and $|+\rangle$, with $E_- < E_+$. Let's assume, additionally, that $|-\rangle$ and $|+\rangle$ are parity eigenstates with opposite parity.

   (a) Compute the matrix elements of the atomic hamiltonian in this two-level basis and show that the hamiltonian can be written as

   $$\mathcal{H}_{\text{atom}} = \bar{E}\mathbb{1} + \frac{1}{2}\hbar\omega_0\sigma_3,$$

   where $\mathbb{1}$ is the 2×2 identity matrix, σ_3 is the third Pauli matrix, $\bar{E} = 0.5(E_- +$

E_+) and $\hbar\omega_0 = E_+ - E_-$.

(b) Show that for this two-level system the matrix expression of $\mathbf{p}$ is

$$\mathbf{p} = -m\omega_0 \langle -|\mathbf{r}|+\rangle \, \sigma_2,$$

and, consequently, in the electric-dipole approximation the matrix of $\mathcal{H}_{\text{int}}$ is

$$\mathcal{H}_{\text{int}} = \hbar \sum_{\mathbf{k}\lambda} C_{\mathbf{k}\lambda} \left(a_{\mathbf{k}\lambda} + a^\dagger_{\mathbf{k}\lambda} \right) \sigma_2,$$

where

$$C_{\mathbf{k}\lambda} = -e\omega_0 \sqrt{\frac{2\pi}{V\hbar\omega_{\mathbf{k}}}} \hat{e}_{\mathbf{k}\lambda} \langle -|\mathbf{r}|+\rangle \, .$$

(c) Show that σ_2 can be factorized as $\sigma_2 = i\left(\sigma - \sigma^\dagger\right)$, where $\sigma|+\rangle = |-\rangle$, $\sigma|-\rangle = 0$, $\sigma^\dagger|+\rangle = 0$ and $\sigma^\dagger|-\rangle = |+\rangle$. The σ and $\sigma^\dagger$ are the ladder operators of the two-level atom.

Therefore, the total hamiltonian of the atom and radiation field is,

$$\mathcal{H} = \bar{E}\mathbb{1} + \frac{1}{2}\hbar\omega_0\sigma_3 + \sum_{k\lambda} \hbar\omega_k \left(a^\dagger_{k\lambda} a_{k\lambda} + \frac{1}{2} \right) + i\hbar \sum_{k\lambda} C_{k\lambda} \left(a_{k\lambda} + a^\dagger_{k\lambda} \right) \left(\sigma - \sigma^\dagger \right).$$

The $\bar{E}\mathbb{1} + \sum_{k\lambda} \hbar\omega_k \frac{1}{2}$ term only adds a global constant (infinite in this case) and is often disregarded. Also, if we consider coupling to a single radiation mode, the hamiltonian becomes:

$$\mathcal{H}_{\text{R}} = \frac{1}{2}\hbar\omega_0\sigma_3 + \hbar\omega_c a^\dagger a + i\hbar C \left(a + a^\dagger \right) \left(\sigma - \sigma^\dagger \right).$$

This is called the Rabi hamiltonian. If $|C| \ll \omega_0, \omega_c$ a further simplification can be made,

$$i\hbar C \left(a + a^\dagger \right) \left(\sigma - \sigma^\dagger \right) \approx i\hbar C \left(a^\dagger\sigma - a\sigma^\dagger \right),$$

leading to the Jaynes–Cummings hamiltonian:

$$\mathcal{H}_{\text{JC}} = \frac{1}{2}\hbar\omega_0\sigma_3 + \hbar\omega_c a^\dagger a + i\hbar C \left(a^\dagger\sigma - a\sigma^\dagger \right),$$

which is very important in quantum optics.

(d) Consider the operator $N = \frac{1}{2}\sigma_3 + a^\dagger a$. Show that $[N, \mathcal{H}_{\text{JC}}] = 0$. What is the physical meaning of N?

(e) Consider the 2-state subspace $H_n = \{|+, n\rangle \, |-, n+1\rangle\}$ and diagonalize $\mathcal{H}_{\text{JC}}$ in that subspace.

Solution:

(a) With the name convention of the text and since $\mathcal{H}_{\rm atom}\left|-\right\rangle = E_-\left|-\right\rangle$ and $\mathcal{H}_{\rm atom}\left|+\right\rangle = E_+\left|+\right\rangle$

$$\mathcal{H}_{\rm atom} = \begin{pmatrix} \langle +|\mathcal{H}_{\rm atom}|+\rangle & \langle +|\mathcal{H}_{\rm atom}|-\rangle \\ \langle -|\mathcal{H}_{\rm atom}|+\rangle & \langle -|\mathcal{H}_{\rm atom}|-\rangle \end{pmatrix} = \begin{pmatrix} E_+ & 0 \\ 0 & E_- \end{pmatrix} = \\ \begin{pmatrix} 0.5(E_+ + E_-) & 0 \\ 0 & 0.5(E_+ + E_-) \end{pmatrix} + \frac{1}{2}\begin{pmatrix} (E_+ - E_-) & 0 \\ 0 & -(E_+ - E_-) \end{pmatrix} = \\ 0.5(E_+ + E_-)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{1}{2}(E_+ - E_-)\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \bar{E}\mathbb{1} + \frac{1}{2}\hbar\omega_0\sigma_3$$

(b) To show that $\mathbf{p} = m\omega_0\left\langle -|\mathbf{r}|+\right\rangle\sigma_2$ we proceed similarly:

$$\mathbf{p} = \begin{pmatrix} \langle +|\mathbf{p}|+\rangle & \langle +|\mathbf{p}|-\rangle \\ \langle -|\mathbf{p}|+\rangle & \langle -|\mathbf{p}|-\rangle \end{pmatrix} = \begin{pmatrix} 0 & \langle +|\mathbf{p}|-\rangle \\ \langle -|\mathbf{p}|+\rangle & 0 \end{pmatrix} = \begin{pmatrix} 0 & im\omega_0\left\langle +|\mathbf{r}|-\right\rangle \\ -im\omega_0\left\langle -|\mathbf{r}|+\right\rangle & 0 \end{pmatrix} \\ = -m\omega_0\left\langle -|\mathbf{r}|+\right\rangle\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = -m\omega_0\left\langle -|\mathbf{r}|+\right\rangle\sigma_2,$$

where we have used the relations

$$\langle +|\mathbf{p}|-\rangle = \frac{im}{\hbar}(E_+ - E_-)\left\langle +|\mathbf{r}|-\right\rangle = im\omega_0\left\langle +|\mathbf{r}|-\right\rangle, \qquad \langle -|\mathbf{p}|+\rangle = \langle +|\mathbf{p}|-\rangle^* = -im\omega_0\left\langle -|\mathbf{r}|+\right\rangle,$$

which are proved in the lecture notes. Note that, since the $\mathbf{r}$ matrix elements in the position representation are real, we have $\langle +|\mathbf{r}|-\rangle = \langle -|\mathbf{r}|+\rangle$. Furthermore, parity conservation demands the diagonal elements to vanish.

From eq. (4.71) of the lecture notes and the definition of $C_{\mathbf{k}\lambda}$ in the text, it can be concluded that

$$\mathcal{H}_{\rm int} = \hbar\sum_{\mathbf{k}\lambda} C_{\mathbf{k}\lambda}\left(a_{\mathbf{k}\lambda} + a^\dagger_{\mathbf{k}\lambda}\right)\sigma_2.$$

(c) This is an easy demonstration

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = i\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} - i\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = i(\sigma - \sigma^\dagger).$$

(d) We need to show that N is a constant of motion:

$$[N, \mathcal{H}_{\rm JC}] = \left[\frac{1}{2}\sigma_3 + a^\dagger a, \frac{1}{2}\hbar\omega_0\sigma_3 + \hbar\omega_c a^\dagger a + i\hbar C\left(a^\dagger\sigma - a\sigma^\dagger\right)\right] = \\ i\hbar C\left(\frac{1}{2}\left[\sigma_3, \left(a^\dagger\sigma - a\sigma^\dagger\right)\right] + \left[a^\dagger a, \left(a^\dagger\sigma - a\sigma^\dagger\right)\right]\right) = \\ i\hbar C\left(\frac{a^\dagger}{2}[\sigma_3, \sigma] - \frac{a}{2}[\sigma_3, \sigma^\dagger] + [a^\dagger a, a^\dagger]\sigma - [a^\dagger a, a]\sigma^\dagger\right) = \\ i\hbar C\left(\frac{a^\dagger}{2}[\sigma_3, \sigma] - \frac{a}{2}[\sigma_3, \sigma^\dagger] + a^\dagger[a, a^\dagger]\sigma - [a^\dagger, a]a\sigma^\dagger\right) = i\hbar C\left(-a^\dagger\sigma - a\sigma^\dagger + a^\dagger\sigma + a\sigma^\dagger\right) = 0,$$

where we have used the easy to prove relations:

$$[\sigma_3, \sigma] = -2\sigma, \qquad [\sigma_3, \sigma^\dagger] = 2\sigma^\dagger.$$

N has the physical meaning of the sum of the number of photons plus the excitation of the atom state (0 in the ground $\left|-\right\rangle$ state and 1 in the excited $\left|+\right\rangle$ state).

(e) The hamiltonian $\mathcal{H}_{\rm JC}$ is diagonal in 2×2 blocks in the H_n subspaces. The matrix of $\mathcal{H}_{\rm JC}$ in H_n is

$$\mathcal{H}_{\rm JC} = \begin{pmatrix} \hbar\left(n\omega_c + \frac{\omega_0}{2}\right) & -i\hbar C\sqrt{n+1} \\ i\hbar C\sqrt{n+1} & \hbar\left((n+1)\omega_c - \frac{\omega_0}{2}\right). \end{pmatrix}$$

After some calculation, solving the secular equation it can be found that the eigenvalues of this matrix are:

$$\lambda_\pm = \hbar\omega_c\left(n + \frac{1}{2}\right) \pm \frac{\hbar}{2}\sqrt{(\omega_c - \omega_0)^2 + 4|C|^2(n+1)}.$$

A.5 Chapter 5.

1. A particle is scattered elastically by a *hard-sphere* type potential,

$$V(r) = \begin{cases} \infty & \text{if } r < a, \\ 0 & \text{if } r > a. \end{cases}$$

The energy of the scattered particle satisfies $ka = 1$.

(a) Prove that the phase shifts satisfy the relation $\tan\delta_\ell = j_\ell(ka)/n_\ell(ka)$ and compute, in radians, the values of δ_ℓ and $\sin\delta_\ell$ for $\ell = 0, 1, 2$.

(b) Calculate the differential cross-section, $\mathrm{d}\sigma/\mathrm{d}\Omega$, for angles $\theta = 0, \pm\frac{\pi}{2}, \pi$, taking into account only the partial waves $\ell = 0$ and $\ell = 1$. Leave the results as a function of the phase shifts.

(c) Calculate the total cross-section, σ, taking into account only the partial waves $\ell = 0$ and $\ell = 1$. Compare your result with the values $4\pi a^2$ and $2\pi a^2$, and discuss if we are in the high-energy or low-energy regime.

(d) Compute the correction to σ that you would obtain by including the partial wave with $\ell = 2$. Do you expect large contributions to σ from partial-waves with $\ell > 2$?

You'll need the following spherical Bessel functions

$$j_0(x) = \frac{\sin x}{x}, \qquad j_1(x) = \frac{\sin x}{x^2} - \frac{\cos x}{x}, \qquad j_2(x) = \left(\frac{3}{x^2} - \frac{1}{x}\right)\sin x - \frac{3\cos x}{x^2},$$

$$n_0(x) = -\frac{\cos x}{x}, \qquad n_1(x) = -\frac{\cos x}{x^2} - \frac{\sin x}{x}, \qquad n_2(x) = -\left(\frac{3}{x^2} - \frac{1}{x}\right)\cos x - \frac{3\sin x}{x^2}.$$

Solution:

(a) This is a central potential and partial-wave analysis is suitable. Since $V(r = a) = \infty$, we must have $\psi(a, \theta) = 0$ for all values of θ. Therefore

$$\sum_{\ell=0}^{\infty} i^\ell (2\ell + 1)\, \left[j_\ell(ka) + ik\, a_\ell\, h_\ell^{(1)}(ka)\right]\, P_\ell(\cos\theta) = 0.$$

Multiply this equation by $P_{\ell'}(\cos\theta)\,\sin\theta\,\mathrm{d}\theta$, integrate from 0 to π and use the orthogonality relation of the Legendre polynomials to get

$$\sum_{\ell=0}^{\infty} i^\ell (2\ell + 1)\, \left[j_\ell(kR) + ik\, a_\ell\, h_\ell^{(1)}(kR)\right] \int_0^\pi P_\ell(\cos\theta)\, P_{\ell'}(\cos\theta)\, \sin\theta\, \mathrm{d}\theta$$

$$= \sum_{\ell=0}^{\infty} i^\ell (2\ell + 1)\, \left[j_\ell(kR) + ik\, a_\ell\, h_\ell^{(1)}(kR)\right] \left(\frac{2}{2\ell + 1}\right) \delta_{\ell\ell'}$$

$$= 2i^\ell \left[j_\ell(kR) + ik\, a_\ell\, h_\ell^{(1)}(kR)\right] = 0.$$

From this it follows that

$$a_\ell = i\, \frac{j_\ell(ka)}{k\, h_\ell^{(1)}(ka)} = \frac{i}{k}\, \frac{j_\ell(ka)}{j_\ell(ka) + i n_\ell(ka)}.$$

The phase shifts are related to the partial-wave amplitudes through

$$a_\ell = \frac{1}{k\cot\delta_\ell - ik}.$$

From the two last equations we finally get

$$\tan\delta_\ell = \frac{j_\ell(ka)}{n_\ell(ka)}.$$

Using the given spherical functions for $ka = 1$ we obtain

$$\delta_0 = -1.00, \qquad \delta_1 = -0.21, \qquad \delta_2 = -0.02.$$

$$\sin\delta_0 = -0.84, \qquad \sin\delta_1 = -0.21, \qquad \sin\delta_2 = -0.02.$$

(b) The differential cross-section is

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = |f(\theta)|^2 = \left|\sum_{\ell=0}^{\infty}(2\ell+1)\,a_\ell\,P_\ell(\cos\theta)\right|^2 = \frac{1}{k^2}\left|\sum_{\ell=0}^{\infty}(2\ell+1)\,e^{i\delta_\ell}\sin\delta_\ell\,P_\ell(\cos\theta)\right|^2.$$

For $\ell = 0, 1$ and $k = 1/a$,

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = a^2\left|e^{i\delta_0}\sin\delta_0 + 3e^{i\delta_1}\sin\delta_1\cos\theta\right|^2$$
$$= a^2\Big(\sin^2\delta_0 + 6\sin\delta_0\sin\delta_1\cos(\delta_0-\delta_1)\cos\theta + 9\sin^2\delta_1\cos^2\theta\Big).$$

Setting now $\theta = 0, \pi, \pm\frac{\pi}{2}$,

$$\left.\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right|_{\theta=0} = a^2\Big(\sin^2\delta_0 + 6\sin\delta_0\sin\delta_1\cos(\delta_0-\delta_1) + 9\sin^2\delta_1\Big),$$

$$\left.\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right|_{\theta=\pi} = a^2\Big(\sin^2\delta_0 - 6\sin\delta_0\sin\delta_1\cos(\delta_0-\delta_1) + 9\sin^2\delta_1\Big).$$

$$\left.\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right|_{\theta=\pm\frac{\pi}{2}} = a^2\sin^2\delta_0.$$

(c) The total cross section is given by

$$\sigma = \frac{4\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell+1)\,\sin^2\delta_\ell.$$

For $\ell = 0, 1$ and $k = 1/a$,

$$\sigma = 4\pi a^2\Big(\sin^2\delta_0 + 3\sin^2\delta_1\Big) \approx 0.844 \times 4\pi a^2.$$

This value is close to what was obtained in the lecture notes under the assumption of low-energy scattering ($4\pi a^2$), and far from the value obtained for high-energy scattering ($2\pi a^2$), so we are here in the low-energy regime.

(d) Including in the cross section calculation the term with $\ell = 2$ we get

$$\sigma \approx (0.844 + 0.001)\,4\pi a^2.$$

We do not expect large contributions from partial waves with $\ell > 2$. In fact, as a general rule, $\sigma_{\ell'} < \sigma_\ell$ for $\ell' > \ell$.

2. Starting from the integral form of the Schrödinger equation, prove that the scattering

amplitude in the second Born approximation is given by

$$f(\theta,\varphi) \;=\; -\frac{m}{2\pi\hbar^2}\int e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}_0}\,V(\mathbf{r}_0)\,\mathrm{d}^3\mathbf{r}_0 + \Big(\frac{m}{2\pi\hbar^2}\Big)^2\int e^{-i\mathbf{k}\cdot\mathbf{r}_0}\,V(\mathbf{r}_0)\,\left[\int\frac{e^{ik|\mathbf{r}_0-\mathbf{r}_1|}}{|\mathbf{r}_0-\mathbf{r}_1|}\,V(\mathbf{r}_1)\,e^{ikz_1}\,\mathrm{d}^3\mathbf{r}_1\right]\mathrm{d}^3\mathbf{r}_0,$$

where $\mathbf{k}\equiv k\hat{\mathbf{r}}$ is the wave vector of the scattered particle, and $\mathbf{k}'=k\hat{\mathbf{z}}$ is the wave vector of the incidente particle (for elastic scattering $|\mathbf{k}|=|\mathbf{k}'|=k$). Also, take $\psi_0(\mathbf{r})=Ae^{ikz}$ to be the free-particle solution of the Schrödinger equation for the incident particle. Apply this formula to the case of low-energy soft-sphere scattering, where the potential is

$$V(\mathbf{r})=\begin{cases}V_0, & \text{if } r\le a,\\ 0, & \text{if } r>a,\end{cases}$$

with V_0 very small, and obtain the corresponding scattering amplitude, $f(\theta)$.

Solution: The integral form of the Schrödinger equation is

$$\psi(\mathbf{r}) \;=\; \psi_0(\mathbf{r})-\frac{m}{2\pi\hbar^2}\int\frac{e^{ik|\mathbf{r}-\mathbf{r}_0|}}{|\mathbf{r}-\mathbf{r}_0|}V(\mathbf{r}_0)\psi(\mathbf{r}_0)\,\mathrm{d}^3\mathbf{r}_0 \;=\; \psi_0(\mathbf{r})+\int g(\mathbf{r}-\mathbf{r}_0)V(\mathbf{r}_0)\psi_0(\mathbf{r}_0)\,\mathrm{d}^3\mathbf{r}_0,$$

where $g(\mathbf{r}-\mathbf{r}_0)$ is the Green's function. The zero-th order solution is $\psi_0(\mathbf{r})$, while the first and second order solutions are

$$\psi_1(\mathbf{r})=\psi_0(\mathbf{r})+\int g(\mathbf{r}-\mathbf{r}_1)V(\mathbf{r}_1)\psi_0(\mathbf{r}_1)\,\mathrm{d}^3\mathbf{r}_1,$$
$$\psi_2(\mathbf{r})=\psi_0(\mathbf{r})+\int g(\mathbf{r}-\mathbf{r}_2)V(\mathbf{r}_2)\psi_1(\mathbf{r}_2)\,\mathrm{d}^3\mathbf{r}_2.$$

Note very carefully that ψ_1, when inserted in the previous expression for ψ_2, must be evaluated at $\mathbf{r}_2$. Thus

$$\psi_1(\mathbf{r}_2)=\psi_0(\mathbf{r}_2)+\int g(\mathbf{r}_2-\mathbf{r}_1)V(\mathbf{r}_1)\psi_0(\mathbf{r}_1)\,\mathrm{d}^3\mathbf{r}_1.$$

The second-order Born approximation is then

$$\psi_2(\mathbf{r}) \;=\; \psi_0(\mathbf{r}) \;+\; \int g(\mathbf{r}-\mathbf{r}_2)V(\mathbf{r}_2)\left[\psi_0(\mathbf{r}_2)+\int g(\mathbf{r}_2-\mathbf{r}_1)V(\mathbf{r}_1)\psi_0(\mathbf{r}_1)\,\mathrm{d}^3\mathbf{r}_1\right]\mathrm{d}^3\mathbf{r}_2$$
$$=\; \psi_0(\mathbf{r})+\int g(\mathbf{r}-\mathbf{r}_2)V(\mathbf{r}_2)\psi_0(\mathbf{r}_2)\,\mathrm{d}^3\mathbf{r}_2+\int g(\mathbf{r}-\mathbf{r}_2)V(\mathbf{r}_2)\left[\int g(\mathbf{r}_2-\mathbf{r}_1)V(\mathbf{r}_1)\psi_0(\mathbf{r}_1)\,\mathrm{d}^3\mathbf{r}_1\right]\mathrm{d}^3\mathbf{r}_2.$$

Making explicit the Green's function, and setting $\psi_0(\mathbf{r})=Ae^{ikz}$, we have

$$\psi_2(\mathbf{r})=Ae^{ikz}-\frac{mA}{2\pi\hbar^2}\int\frac{e^{ik|\mathbf{r}-\mathbf{r}_2|}}{|\mathbf{r}-\mathbf{r}_2|}\,V(\mathbf{r}_2)\,e^{ikz_2}\,\mathrm{d}^3\mathbf{r}_2 + \Big(\frac{m}{2\pi\hbar^2}\Big)^2 A\int\frac{e^{ik|\mathbf{r}-\mathbf{r}_2|}}{|\mathbf{r}-\mathbf{r}_2|}\,V(\mathbf{r}_2)\,\left[\int\frac{e^{ik|\mathbf{r}_2-\mathbf{r}_1|}}{|\mathbf{r}_2-\mathbf{r}_1|}\,V(\mathbf{r}_1)\,e^{ikz_1}\,\mathrm{d}^3\mathbf{r}_1\right]\mathrm{d}^3\mathbf{r}_2.$$

In the asymptotic region $r=|\mathbf{r}|\gg|\mathbf{r}_2|=r_2$, which leads to $\dfrac{e^{ik|\mathbf{r}-\mathbf{r}_2|}}{|\mathbf{r}-\mathbf{r}_2|}\approx\dfrac{e^{ikr}}{r}e^{-i\mathbf{k}\cdot\mathbf{r}_2}$. Note, however, that $|\mathbf{r}_2-\mathbf{r}_1|$ is not necessarily small. Therefore

$$\frac{1}{A}\psi_2(\mathbf{r})=e^{ikz}-\frac{m}{2\pi\hbar^2}\frac{e^{ikr}}{r}\int e^{-i\mathbf{k}\cdot\mathbf{r}_2}\,V(\mathbf{r}_2)\,e^{ikz_2}\,\mathrm{d}^3\mathbf{r}_2 + \Big(\frac{m}{2\pi\hbar^2}\Big)^2\frac{e^{ikr}}{r}\int e^{-i\mathbf{k}\cdot\mathbf{r}_2}\,V(\mathbf{r}_2)\,\left[\int\frac{e^{ik|\mathbf{r}_2-\mathbf{r}_1|}}{|\mathbf{r}_2-\mathbf{r}_1|}\,V(\mathbf{r}_1)\,e^{ikz_1}\,\mathrm{d}^3\mathbf{r}_1\right]\mathrm{d}^3\mathbf{r}_2.$$

This expression tells us that the scattering amplitude is

$$f(\theta,\varphi) \;=\; -\frac{m}{2\pi\hbar^2}\int e^{-i\mathbf{k}\cdot\mathbf{r}_2}\,V(\mathbf{r}_2)\,e^{ikz_2}\,\mathrm{d}^3\mathbf{r}_2 \\ + \Big(\frac{m}{2\pi\hbar^2}\Big)^2\int e^{-i\mathbf{k}\cdot\mathbf{r}_2}\,V(\mathbf{r}_2)\,\left[\int\frac{e^{ik|\mathbf{r}_2-\mathbf{r}_1|}}{|\mathbf{r}_2-\mathbf{r}_1|}\,V(\mathbf{r}_1)\,e^{ikz_1}\,\mathrm{d}^3\mathbf{r}_1\right]\mathrm{d}^3\mathbf{r}_2.$$

Since $e^{ikz_2}=e^{i\mathbf{k}'\cdot\mathbf{r}_2}$, we can also write

$$f(\theta,\varphi) \;=\; -\frac{m}{2\pi\hbar^2}\int e^{i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}_2}\,V(\mathbf{r}_2)\,\mathrm{d}^3\mathbf{r}_2 \\ + \Big(\frac{m}{2\pi\hbar^2}\Big)^2\int e^{-i\mathbf{k}\cdot\mathbf{r}_2}\,V(\mathbf{r}_2)\,\left[\int\frac{e^{ik|\mathbf{r}_2-\mathbf{r}_1|}}{|\mathbf{r}_2-\mathbf{r}_1|}\,V(\mathbf{r}_1)\,e^{ikz_1}\,\mathrm{d}^3\mathbf{r}_1\right]\mathrm{d}^3\mathbf{r}_2.$$

For low-energy scattering the exponentials inside the integrals can be approximated to be 1, leading to

$$f(\theta,\varphi) \;=\; -\frac{m}{2\pi\hbar^2}\int V(\mathbf{r}_2)\,\mathrm{d}^3\mathbf{r}_2+\Big(\frac{m}{2\pi\hbar^2}\Big)^2\int V(\mathbf{r}_2)\,\left[\int\frac{1}{|\mathbf{r}_2-\mathbf{r}_1|}\,V(\mathbf{r}_1)\,\mathrm{d}^3\mathbf{r}_1\right]\mathrm{d}^3\mathbf{r}_2.$$

Let's do now the integrals for the specific case of low-energy soft-sphere scattering. Choose the z_1 axis along $\mathbf{r}_2$, so that $|\mathbf{r}_2-\mathbf{r}_1|^2=r_2^2+r_1^2-2r_2r_1\cos\theta_1$, and

$$\int\frac{1}{|\mathbf{r}_2-\mathbf{r}_1|}\,V(\mathbf{r}_1)\,\mathrm{d}^3\mathbf{r}_1 = V_0\int_0^a\frac{1}{|\mathbf{r}_2-\mathbf{r}_1|}\,r_1^2\,\sin\theta_1\,\mathrm{d}r_1\mathrm{d}\theta_1\mathrm{d}\varphi_1 \\ = 2\pi V_0\int_0^a r_1^2\left[\int_0^\pi\frac{\sin\theta_1\,\mathrm{d}\theta_1}{\sqrt{r_2^2+r_1^2-2r_2r_1\cos\theta_1}}\right]\mathrm{d}r_1.$$

Furthermore,

$$\int_0^\pi\frac{\sin\theta_1\,\mathrm{d}\theta_1}{\sqrt{r_2^2+r_1^2-2r_2r_1\cos\theta_1}} = \frac{1}{r_2r_1}\sqrt{r_2^2+r_1^2-2r_2r_1\cos\theta_1}\Big|_0^\pi \\ = \frac{1}{r_2r_1}\Big[(r_2+r_1)-|r_2-r_1|\Big] \;=\; \begin{cases}2/r_2, & \text{if } r_1\le r_2,\\ 2/r_1, & \text{if } r_1>r_2.\end{cases}$$

Note that we have always $r_1<a$ due to the outer integral, therefore

$$\int\frac{1}{|\mathbf{r}_2-\mathbf{r}_1|}\,V(\mathbf{r}_1)\,\mathrm{d}^3\mathbf{r}_1 = 4\pi V_0\left[\int_0^{r_2}r_1^2\,\frac{1}{r_2}\,\mathrm{d}r_1+\int_{r_2}^a r_1^2\,\frac{1}{r_1}\,\mathrm{d}r_1\right]=2\pi V_0\left(a^2-\frac{1}{3}r_2^2\right).$$

$$\int V(\mathbf{r}_2)\,\left[\int\frac{1}{|\mathbf{r}_2-\mathbf{r}_1|}\,V(\mathbf{r}_1)\,\mathrm{d}^3\mathbf{r}_1\right]\mathrm{d}^3\mathbf{r}_2 \;=\; 2\pi V_0^2 4\pi\int_0^a\left(a^2-\frac{1}{3}r_2^2\right)\,r_2^2\,\mathrm{d}r_2 \;=\; \frac{32\pi^2}{15}V_0^2a^5.$$

Finally, the scattering amplitude in the second Born approximation is

$$f(\theta) \;=\; -\frac{mV_0}{2\pi\hbar^2}\frac{4}{3}\pi a^3+\Big(\frac{m}{2\pi\hbar^2}\Big)^2\,\frac{32\pi^2}{15}V_0^2a^5 \;=\; -\left(\frac{2mV_0a^3}{3\hbar^2}\right)\,\left[1-\frac{4}{5}\left(\frac{mV_0a^2}{\hbar^2}\right)\right].$$

A.6 Exam problems from previous years.

1. Without fine-structure corrections the $2s$ and $2p$ Hydrogen levels have exactly the same energy, $E_{2s} = E_{2p}$. However, the relativistic correction to the electron kinetic energy modifies E_{2s} and E_{2p} in such a way that, after the correction, they are no longer equal, $E'_{2s} \neq E'_{2p}$. Compute the energy difference $(\Delta E)_{\text{RelCorr}} = E'_{2s} - E'_{2p}$, giving the result in eV. Compute the value of the electric field $\mathcal{E}$ (in V/cm) needed to make the Stark splitting in the $n = 2$ Hydrogen levels to be the same as the splitting $(\Delta E)_{\text{RelCorr}}$ due to the relativistic correction.

Solution: The relativistic correction to the kinetic energy is

$$E_{\text{r}} = E_n^{(0)} \frac{(Z\alpha)^2}{n} \left[-\frac{3}{4n} + \frac{1}{(\ell + \frac{1}{2})} \right],$$

where $E_n^{(0)}$ is the uncorrected value of the level n energy. Therefore

$$(\Delta E)_{\text{RelCorr}} = E'_{2s} - E'_{2p} = E_2^{(0)} \frac{\alpha^2}{2} \left[\frac{1}{1/2} - \frac{1}{3/2} \right]$$
$$= \frac{2}{3}\alpha^2 E_2^{(0)} = \frac{2}{3} \left(\frac{1}{137} \right)^2 \frac{13.6\,\text{eV}}{4} = 1.2 \times 10^{-4}\,\text{eV}.$$

The Stark splitting in the $n = 2$ Hydrogen levels is $(\Delta E)_{\text{Stark}} = 3e\mathcal{E}a_0$, so

$$3e\mathcal{E}a_0 = 1.2 \times 10^{-4}\,\text{eV} \implies 3\mathcal{E}a_0 = 1.2 \times 10^{-4}\,\text{V} \implies \mathcal{E} = 0.4 \times 10^{-4}\,\frac{\text{V}}{a_0}.$$

$$\mathcal{E} = \frac{0.4 \times 10^{-4}}{5.29 \times 10^{-9}}\,\frac{\text{V}}{\text{cm}} = 7.5 \times 10^3\,\frac{\text{V}}{\text{cm}}.$$

This is quite a high voltage to put across a centimeter, but it can be done. The breakdown voltage in the air is about four times higher than this.

2. In a scattering experiment on a spherically symmetric potential, the incoming particles have an energy energy such that only the partial waves s and p need to be considered. In these conditions:

   (a) Show that the differential cross-section can be written in the form

   $$\frac{\text{d}\sigma}{\text{d}\Omega} = a + b\cos\theta + c\cos^2\theta.$$

   Write the parameters a, b and c as functions of the phase shifts.

   (b) Compute the total cross-section and show its dependence on the parameters a, b and c. Check if this result verifies the optical theorem.

   Help: $P_0(\cos\theta) = 1$, $P_1(\cos\theta) = \cos\theta$.

Solution:

(a) The scattering amplitude for a spherically symmetrical potential is given by, in terms of the phase shifts, as

$$f(\theta) = \frac{1}{k} \sum_{\ell=0}^{\infty} (2\ell + 1)\, e^{i\delta_\ell} \sin(\delta_\ell)\, P_\ell(\cos\theta).$$

Since $P_0(\cos\theta) = 1$, and $P_1(\cos\theta) = \cos\theta$, for only s and p waves we have

$$f(\theta) = \frac{1}{k} \left(e^{i\delta_0} \sin\delta_0 + 3e^{i\delta_1} \sin\delta_1 \cos\theta \right),$$

$$k^2|f(\theta)|^2 = \left(e^{i\delta_0}\sin\delta_0 + 3e^{i\delta_1}\sin\delta_1\,\cos\theta\right)\left(e^{-i\delta_0}\sin\delta_0 + 3e^{-i\delta_1}\sin\delta_1\,\cos\theta\right)$$
$$= \sin^2\delta_0 + 9\sin^2\delta_1\cos^2\theta + 3\left[e^{i(\delta_0-\delta_1)} + e^{-i(\delta_0-\delta_1)}\right]\sin\delta_0\sin\delta_1\,\cos\theta$$
$$= \sin^2\delta_0 \;+\; 6\cos(\delta_0-\delta_1)\sin\delta_0\sin\delta_1\,\cos\theta \;+\; 9\sin^2\delta_1\,\cos^2\theta.$$

Therefore,

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} \;=\; |f(\theta)|^2 \;=\; \frac{1}{k^2}\sin^2\delta_0 \;+\; \frac{6}{k^2}\cos(\delta_0-\delta_1)\sin\delta_0\sin\delta_1\,\cos\theta \;+\; \frac{9}{k^2}\sin^2\delta_1\,\cos^2\theta$$
$$= a + b\cos\theta + c\cos^2\theta,$$

with

$$a = \frac{1}{k^2}\sin^2\delta_0, \qquad b = \frac{6}{k^2}\cos(\delta_0-\delta_1)\sin\delta_0\sin\delta_1, \qquad c = \frac{9}{k^2}\sin^2\delta_1.$$

(b) The total scattering cross-section is

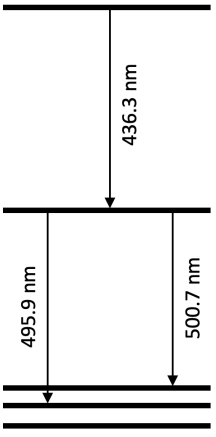
$$\sigma = \int |f(\theta)|^2\,\mathrm{d}\Omega \;=\; 2\pi\int_{-1}^{+1}(a + b\cos\theta + c\cos^2\theta)\,\mathrm{d}(\cos\theta) \;=\; 4\pi a + \frac{4\pi c}{3}$$
$$= \frac{4\pi}{k^2}\Big(\sin^2\delta_0 + 3\sin^2\delta_1\Big).$$

The optical theorem says that

$$\sigma = \frac{4\pi}{k}\,\mathrm{Im}[f(0)] = \frac{4\pi}{k^2}\Big(\sin^2\delta_0 + 3\sin^2\delta_1\Big),$$

giving the same cross-section as before.

3. Consider the five lowest-energy levels of doubly ionized Oxygen ($Z = 8$), O^{2+}, that are shown in this figure.
   a) Determine the electronic configuration of O^{2+} and, under LS coupling, write in the figure the atomic terms with the quantum numbers corresponding to every energy **level** (draw your own figure if you want to take away the exam sheet).
   b) Determine the multipole nature of the 436.3 nm transition.
   c) The green color of the Orion Nebula originates from the 495.9 nm and 500.7 nm transitions. The values of the Einstein's A coefficients are 6.2×10^{-3} and 1.8×10^{-2} s^{-1}, respectively, whose magnitude indicate that they are M1+E2 transitions. Discuss the compatibility of this observation with the selection rules you know.



Solution: **a)** Since O^{2+} has six electrons its configuration, following Madelung's rule, is that of Carbon: $(1s^2)(2s^2)(2p^2)$. The first two shells are filled and, therefore, the terms are described with the total spin and orbital angular momenta of the two equivalent electrons of the p shell ($\ell_1 = 1$ and $\ell_2 = 1$, $s_1 = 1/2$ and $s_2 = 1/2$). With the addition rules of angular momentum we have that

$$L = \ell_1 + \ell_2, ..., |\ell_1 - \ell_2| = 2, 1, 0.$$

And

$$S = s_1 + s_2, ..., |s_1 - s_2| = 1, 0.$$

Since these are two equivalent electrons only $L + S$ even combinations give an antisymmetric wave function according to Pauli's principle. Therefore, the terms for this configuration are: ^{1}D ($L = 2$

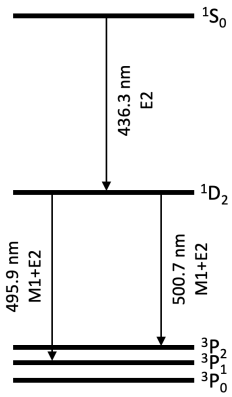
and $S = 0$),^{1}S ($L = 0$ and $S = 0$) and ^{3}P ($L = 1$ and $S = 1$). From these terms we obtain the levels by coupling the orbital and spin angular momenta: ^{1}D$_2$, ^{1}S$_0$, ^{3}P$_2$, ^{3}P$_1$ and ^{3}P$_0$.

Now, since we are under the hypothesis of LS coupling, to order these levels we follow Hund's rules:

(1) The term with the highest total spin, S, has the lowest energy. Therefore, the ^{3}P term is our lowest energy term. Since both the ^{1}D and ^{1}S terms have $S = 0$ this rule does not order them.

(2) For a given spin, S, the term with largest L has the lowest energy. Therefore, the ^{1}D term has lower energy than than the ^{1}S term.

(3) The level with $J = |L - S|$ or $J = L + S$ has the lowest energy depending, respectively, on whether the (sub)shell is less than, or more than, half–full. The only term that splits into more than one level is the ^{3}P term. Since the p shell is less than half full the level with $J = |L - S| = 0$ (^{3}P$_0$) is the least energy level followed by ^{3}P$_1$ and ^{3}P$_2$.



With all this information we deduce the quantum numbers shown on the attached diagram.

b) Now that we have quantum numbers assigned to the different levels we can classify the transitions. The first thing to note is that, since all these states share the same configuration, there is no parity change between the initial and the final level. This excludes E1 transitions.

The 436.3 nm transition corresponds to ^{1}S$_0 \to {}^1$D$_2$. Therefore $\Delta J = 2$, $\Delta S = 0$ and $\Delta L = 2$. There is no change in electron configuration, nor parity change. This means that this is an E2 transition.

c) The 500.7 nm transition corresponds to ^{1}D$_2 \to {}^3$P$_2$. Therefore $\Delta J = 0$, $\Delta S = 1$ and $\Delta L = -1$. There is no change in electron configuration, nor parity change. To be clearly classified as an E2 transition in LS coupling it would require that $\Delta S = 0$. However, this is not a rigorous rule and it can be partially violated. Also, to be classified as a M1 transition in LS coupling it would require $\Delta S = 0$, $\Delta L = 0$ and $\Delta J = \pm 1$. But again, these are only approximate rules that are violated in this case, as indicated by the order of magnitude of the Einstein coefficient.

The 495.9 nm line corresponds to to ^{1}D$_2 \to {}^3$P$_1$. Therefore $\Delta J = -1$, $\Delta S = 1$ and $\Delta L = -1$. As for the 500.7 nm line, the $\Delta S = 0$ selection rule of E2 and M1 transitions is violated. For M1 transitions the $\Delta L = 0$ is also violated. Of course the rigorous selection rules shown in the first three rows of table 4.1 in the lecture notes are satisfied.

Since O^{2+} has six electrons but $Z = 8$ deviation from pure LS coupling is stronger than for C, also with six electrons but $Z = 6$, and it is not surprising that these transitions are observed.

4. a) The bound-state energies of hydrogen-like atoms are given by equation (1.112) of the lecture notes

$$E_n = -\frac{1}{2}\alpha^2 \mu c^2 \frac{Z^2}{n^2},$$

where α is the fine-structure constant, μ is the reduced mass of the electron-nucleus system, c is the speed of light and Z is the number of protons in the nucleus. Compute, in eV, the energy difference between the positronium bound states with $n = 1$ and $n = 2$ (**the positronium is a e^+e^- bound-state system**).

b) The hyperfine interaction in positronium is stronger than the fine interaction, but the denomination *hyperfine structure* is still used in this case, by analogy with hydrogen. Consider the following Hamiltonian, which describes the hyperfine interaction in the positronium states with $n = 1$:

$$\mathcal{H}_{\rm H} = A_{\rm H}\, \mathbf{s}_1 \cdot \mathbf{s}_2,$$

where $A_{\rm H} = \frac{7}{12}\frac{m_e c^2 \alpha^4}{\hbar^2}$, m_e is the electron's mass and $\mathbf{s}_1$, $\mathbf{s}_2$ are the electron and positron spin operators, respectively. Determine, using perturbation theory, the hyperfine correction to the energy of the states $n = 1\,{}^1\mathrm{S}_0$ and $n = 1\,{}^3\mathrm{S}_1$. Give your results in eV.

c) To first order, the positronium's interaction with an external magnetic field does not depend on the orbital angular momenta of the electron and the positron. It only depends on their spin angular momenta and, for a magnetic field aligned with the z axis, is described by the Hamiltonian

$$\mathcal{H}_m = g_e \frac{\mu_B}{\hbar} B\,(s_{1z} - s_{2z}) \approx 2\frac{\mu_B}{\hbar} B\,(s_{1z} - s_{2z})\,.$$

Compute the energy-splittings produced on the $n = 1$ states when $\mathcal{H}_m \sim \mathcal{H}_{\rm H}$ (intermediate field intensity).

d) Draw the energy splittings (obtained in question c) as a function of the magnetic field intensity, paying special attention to the extreme cases of weak-field $\mu_B B \ll \hbar^2 A_{\rm H}$ (similar to the Zeeman effect) and strong-field $\hbar^2 A_{\rm H} \ll \mu_B B$ (similar to the Paschen-Back effect).

Hint: you can get inspiration from Table 2.3 of the lecture notes. You don't need to repeat similar calculations to those you can find in that section of the notes.

Solution:

a) It is kind of easy to see that, since for positronium $\mu = m_e/2$ and $Z = 1$

$$E_n = -\frac{1}{4}\alpha^2 m_e c^2 \frac{Z^2}{n^2} = -\frac{1}{2n^2}\ \mathrm{Ry}.$$

Therefore

$$E_2 - E_1 = \left(-\frac{1}{8} + \frac{1}{2}\right)\ \mathrm{Ry} = \frac{3}{8} \times 13.6\ \mathrm{eV} = 5.1\ \mathrm{eV}.$$

b) We can find the energy shift due to the hyperfine interaction in perturbation theory by using a basis where $\mathcal{H}_{\rm H}$ is diagonal. This is achieved in any basis where $\mathbf{s}_1$ and $\mathbf{s}_2$ are coupled and the states are described by the $S^2 = (\mathbf{s}_1 + \mathbf{s}_2)^2$, s_1^2 and s_2^2 operators and the S, s_1 and s_2 quantum numbers.[2] We need to calculate

$$E_{\rm H} = \langle \alpha S s_1 s_2 \,|\, A_{\rm H} \mathbf{s}_1 \cdot \mathbf{s}_2 \,|\, \alpha S s_1 s_2 \rangle\,,$$

where α indicates other quantum numbers aside from S, s_1 and s_2. Since

$$S^2 = (\mathbf{s_1} + \mathbf{s}_2)^2 = s_1^2 + s_2^2 + 2\mathbf{s}_1 \cdot \mathbf{s}_2 \Rightarrow \mathbf{s}_1 \cdot \mathbf{s}_2 = \frac{1}{2}(S^2 - s_1^2 - s_2^2),$$

we have that

$$E_{\rm H} = \langle \alpha S s_1 s_2 \,|\, A_{\rm H} \mathbf{s}_1 \cdot \mathbf{s}_2 \,|\, \alpha S s_1 s_2 \rangle = \\ = \frac{A_{\rm H}}{2} \langle S s_1 s_2 M_S | S^2 - s_1^2 - s_2^2 | S s_1 s_2 M_S \rangle = \frac{\hbar^2}{2} A_{\rm H} \left(S(S+1) - \frac{3}{2} \right),$$

where it has been already used that $s_1 = s_2 = 1/2$.

[2] These could be, at least, the basis where the observables are $\{L^2, S^2, s_1^2, s_2^2, L_z, S_z\}$ and the basis where the observables are $\{J^2, L^2, S^2, s_1^2, s_2^2, J_z\}$. In positronium, the two spins of the positively and negatively charged particles are coupled first and, when the fine structure is studied, subsequently coupled to the orbital angular momentum, which is noted as $\mathbf{J}$, whereas in hydrogen the electron spin is coupled to the orbital angular momentum first, when studying the fine structure, and then to the proton spin (hyperfine study) and the total angular momentum is noted $\mathbf{F}$.

For the $^3\mathrm{S}_1$ triplet $S=1$ whilst for the $^1\mathrm{S}_0$ singlet $S=0$. Thus

$$E_{\mathrm{H}}(^1\mathrm{S}_0) = -\frac{3\hbar^2}{4}A_{\mathrm{H}} = -\frac{21}{48}m_e c^2\alpha^4 = -\frac{21}{48}0.511\times10^6\,[\mathrm{eV}]\left(\frac{1}{137}\right)^4 = -6.35\times10^{-4}\,\mathrm{eV},$$
$$E_{\mathrm{H}}(^3\mathrm{S}_1) = \quad\frac{\hbar^2}{4}A_{\mathrm{H}} = \quad\frac{7}{48}m_e c^2\alpha^4 = \quad\frac{7}{48}0.511\times10^6\,[\mathrm{eV}]\left(\frac{1}{137}\right)^4 = \quad 2.12\times10^{-4}\,\mathrm{eV}.$$

c) In the intermediate field case both Hamiltonians have to be treated simultaneously. We can consider either the coupled or the uncoupled spin basis. We shall do it in the uncoupled basis where $|\uparrow\uparrow\rangle$ is the state with $m_{s1}=1/2$ and $m_{s2}=1/2$, $|\uparrow\downarrow\rangle$ the state with $m_{s1}=1/2$ and $m_{s2}=-1/2$, etc.

In the uncoupled basis we can compute the matrix elements of $\mathcal{H}_{\mathrm{H}}$ considering that

$$\mathcal{H}_{\mathrm{H}} = A_{\mathrm{H}}\,\mathbf{s}_1\cdot\mathbf{s}_2 = A_{\mathrm{H}}(-s_{1+}s_{2-} + s_{10}s_{20} - s_{1-}s_{2+}),$$

where (s_{i+}, s_{i0}, s_{i-}) are the spherical components of the $i=1,2$ particle. Therefore, the $\mathcal{H}_{\mathrm{H}}$ matrix in the uncoupled spin basis is (it is analogous to the magnetic hyperfine interaction contribution in Table 2.3):

$$\langle\mathcal{H}_{\mathrm{H}}\rangle = \begin{array}{c} \\ \langle\uparrow\uparrow| \\ \langle\uparrow\downarrow| \\ \langle\downarrow\uparrow| \\ \langle\downarrow\downarrow| \end{array}\begin{array}{c} \begin{array}{cccc} |\uparrow\uparrow\rangle & |\uparrow\downarrow\rangle & |\downarrow\uparrow\rangle & |\downarrow\downarrow\rangle \end{array} \\ \begin{pmatrix} \frac{\hbar^2}{4}A_{\mathrm{H}} & 0 & 0 & 0 \\ 0 & -\frac{\hbar^2}{4}A_{\mathrm{H}} & \frac{\hbar^2}{2}A_{\mathrm{H}} & 0 \\ 0 & \frac{\hbar^2}{2}A_{\mathrm{H}} & -\frac{\hbar^2}{4}A_{\mathrm{H}} & 0 \\ 0 & 0 & 0 & \frac{\hbar^2}{4}A_{\mathrm{H}} \end{pmatrix} \end{array}.$$

While the matrix for $\mathcal{H}_m = 2\frac{\mu_B}{\hbar}B\,(s_{1z}-s_{2z})$ is

$$\langle\mathcal{H}_m\rangle = \begin{array}{c} \\ \langle\uparrow\uparrow| \\ \langle\uparrow\downarrow| \\ \langle\downarrow\uparrow| \\ \langle\downarrow\downarrow| \end{array}\begin{array}{c} \begin{array}{cccc} |\uparrow\uparrow\rangle & |\uparrow\downarrow\rangle & |\downarrow\uparrow\rangle & |\downarrow\downarrow\rangle \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2\mu_B B & 0 & 0 \\ 0 & 0 & -2\mu_B B & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{array}.$$

Therefore

$$\langle\mathcal{H}_{\mathrm{H}}+\mathcal{H}_m\rangle = \begin{pmatrix} \frac{\hbar^2}{4}A_{\mathrm{H}} & 0 & 0 & 0 \\ 0 & -\frac{\hbar^2}{4}A_{\mathrm{H}} + 2\mu_B B & \frac{\hbar^2}{2}A_{\mathrm{H}} & 0 \\ 0 & \frac{\hbar^2}{2}A_{\mathrm{H}} & -\frac{\hbar^2}{4}A_{\mathrm{H}} - 2\mu_B B & 0 \\ 0 & 0 & 0 & \frac{\hbar^2}{4}A_{\mathrm{H}} \end{pmatrix}.$$

Diagonalizing this matrix is quite easy if we consider that the first and last elements of the diagonal are already the twice degenerate eigenvalue of the matrix:

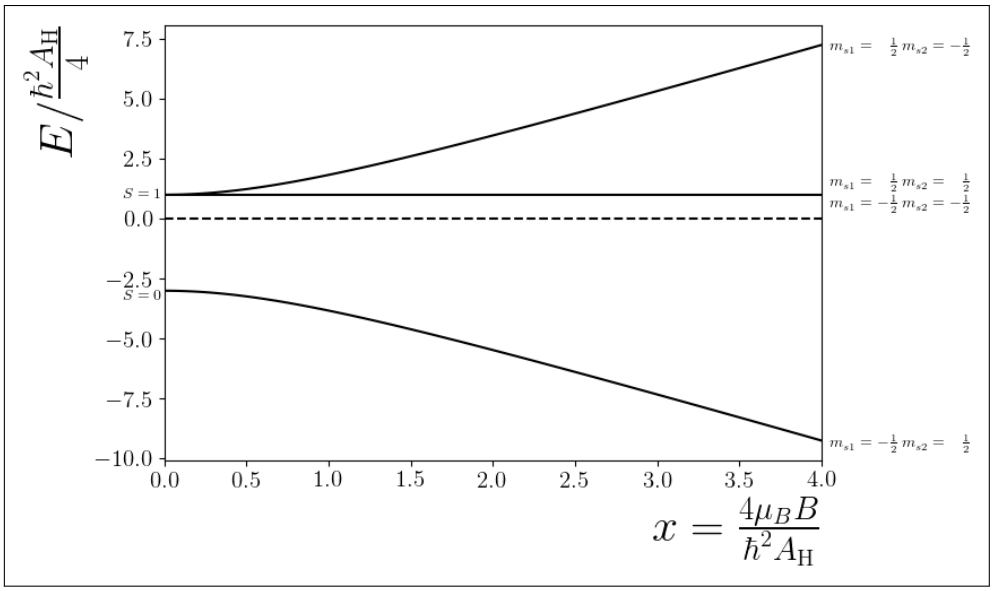
$$E = \frac{\hbar^2}{4}A_{\mathrm{H}}.$$

The non-trivial eigenvalues are the result of diagonalizing the 2×2 central matrix:

$$\begin{vmatrix} -\frac{\hbar^2}{4}A_{\mathrm{H}} + 2\mu_B B - E & \frac{\hbar^2}{2}A_{\mathrm{H}} \\ \frac{\hbar^2}{2}A_{\mathrm{H}} & -\frac{\hbar^2}{4}A_{\mathrm{H}} - 2\mu_B B - E \end{vmatrix} = 0,$$

$$E = -\frac{\hbar^2}{4}A_{\mathrm{H}}\left(1 \pm 2\sqrt{1+\left(\frac{4\mu_B B}{\hbar^2 A_{\mathrm{H}}}\right)^2}\right).$$

d) In the figure the evolution of the energy splitting in units of $\frac{\hbar^2}{4}A_{\mathrm{H}}$ as a function of $x = \frac{4\mu_B B}{\hbar^2 A_{\mathrm{H}}}$ is shown. It can be seen that for $B=0$ the three states of the $S=1$ states are degenerate. Two of them remain unaffected by the external field. For $\hbar^2 A_{\mathrm{H}} << \mu_B B$ the energy dependency with the external magnetic field is linear $E = \pm 2\mu_B B$ for the two states that are sensitive to it.



5. Consider the $n=2$ bound states of hydrogen without fine-structure corrections (uncoupled basis). The system is placed under the simultaneous action of electrostatic and magnetostatic fields aligned with the z axis. Write the 8×8 matrix of the hamiltonian of the perturbation and find its eigenvalues. Some of the eigenvalues are zero and don't give rise to energy shifts, independently of the values of $\mathcal{E}$ and B. In addition, for a particular value of the magnetic to electric field ratio, $B/\mathcal{E}$, the energy shift of some other levels is also zero. What is this value? For an electric field of $\mathcal{E}=10^6$ V/m give the corresponding value of B in Tesla. (You may take the matrix elements you need from the handouts, it is not necessary to repeat the calculations here.)

Solution: In fact the 8×8 matrix reduces to two 4×4 matrices that can be easily computed with table 2.2 and $\mathcal{H}_m$ of equation (2.16),

	$\lvert 0\ 0\ \uparrow\rangle$	$\lvert 1\ 0\ \uparrow\rangle$	$\lvert 1\ 1\ \uparrow\rangle$	$\lvert 1\ \bar{1}\ \uparrow\rangle$
$\langle 0\ 0\ \uparrow\rvert$	$\mu_B B$	$-3e\mathcal{E}a_0/Z$		
$\langle 1\ 0\ \uparrow\rvert$	$-3e\mathcal{E}a_0/Z$	$\mu_B B$		
$\langle 1\ 1\ \uparrow\rvert$			$2\mu_B B$	
$\langle 1\ \bar{1}\ \uparrow\rvert$				

	$\lvert 0\ 0\ \downarrow\rangle$	$\lvert 1\ 0\ \downarrow\rangle$	$\lvert 1\ 1\ \downarrow\rangle$	$\lvert 1\ \bar{1}\ \downarrow\rangle$
$\langle 0\ 0\ \downarrow\rvert$	$-\mu_B B$	$-3e\mathcal{E}a_0/Z$		
$\langle 1\ 0\ \downarrow\rvert$	$-3e\mathcal{E}a_0/Z$	$-\mu_B B$		
$\langle 1\ 1\ \downarrow\rvert$				
$\langle 1\ \bar{1}\ \downarrow\rvert$				$-2\mu_B B$

,

where $\bar{1}=-1$.

This system has 3 trivial eigenvalues: 0, which is twice degenerate, $2\mu_B B$ and $-2\mu_B B$. The other four eigenvalues come from diagonalising the 2×2 matrices

$$\begin{pmatrix} \mu_B B & -3e\mathcal{E}a_0/Z \\ -3e\mathcal{E}a_0/Z & \mu_B B \end{pmatrix}, \qquad \begin{pmatrix} -\mu_B B & -3e\mathcal{E}a_0/Z \\ -3e\mathcal{E}a_0/Z & -\mu_B B \end{pmatrix}.$$

The diagonalisation of these two matrices leads to eigenvalues

$$E_{1,2}=\mu_B B\pm 3e\mathcal{E}a_0/Z, \qquad \text{and,} \qquad E_{3,4}=-\mu_B B\pm 3e\mathcal{E}a_0/Z.$$

Two of these eigenvalues cancel if

$$\mu_B B=3e\mathcal{E}a_0/Z \implies \frac{B}{\mathcal{E}}=\frac{3ea_0}{Z\mu_B}=\frac{6\hbar}{Ze^2}=\frac{6}{Z\alpha c},$$

where the relations $\mu_B=e\hbar/2m$ and $a_0=\hbar^2/me^2=\hbar/m\alpha c$ have been used.

For $\mathcal{E} = 10^6$ V/m (and $Z = 1$), the value of the magnetic field is

$$B = \frac{6 \times 10^6 \times 137}{3 \times 10^8} \frac{\text{V} \cdot \text{s}}{\text{m}^2} = 2.74 \,\text{Tesla}.$$

6. The ground state configuration of Praseodymium ($Z = 59$) is $[\text{Xe}]6\text{s}^2 4\text{f}^3$.
   a) Use Hund's rules to deduce the ground-state term.
   b) What are the levels for this term?
   c) The splitting between the lowest and the highest level of the ground state term is 4381.1 cm^{-1}. Use Landé's interval rule to estimate the energies of the other levels relative to the lowest.

Solution:
a) The 6s shell is full, and so we focus on the three f electrons. Following Hund's rules, to find the ground-state term of the atom we need to maximize S by having three spin-up electrons, so that $\text{S} = 3 \times (1/2) = 3/2$. Next, we have to maximize L with the condition $\text{S} = 3/2$. Since the value of m_s is the same for the three electrons, Pauli's exclusion principle tells us that they must have different values of m_ℓ. The maximum value of the sum of three different values of m_ℓ is $3 + 2 + 1 = 6$, therefore L cannot be larger than 6. We have then $\text{L} = 6$ and $\text{S} = 3/2$ and the ground state term is ^4I.

b) The shell is less than half-full, so the third Hund's rule tells us that this is a normal ordered multiplet with ground state $\text{J} = |\text{L} - \text{S}| = |6 - 3/2| = 9/2$. The ground-state level is thus $^4\text{I}_{9/2}$. The multiplet has the levels

$$^4\text{I}_{\frac{9}{2}}, \;\; ^4\text{I}_{\frac{11}{2}}, \;\; ^4\text{I}_{\frac{13}{2}}, \;\; ^4\text{I}_{\frac{15}{2}}.$$

c) The energy difference between adjacent levels of a multiplet is proportional to J (Landé interval rule),

$$E(J) - E(J-1) = \hbar^2 \zeta J.$$

Let us call E_1, E_2, E_3, and E_4, the energy levels, verifying $E_1 < E_2 < E_3 < E_4$. We can form the differences

$$\Delta_a = E_2 - E_1 = \frac{11}{2}\hbar^2\zeta, \qquad \Delta_b = E_3 - E_2 = \frac{13}{2}\hbar^2\zeta, \qquad \Delta_c = E_4 - E_3 = \frac{15}{2}\hbar^2\zeta.$$

Therefore,

$$\Delta_a + \Delta_b + \Delta_c = E_4 - E_1 = \frac{39}{2}\hbar^2\zeta = 4381.1\,\text{cm}^{-1}.$$

From this we can compute $\hbar^2\zeta = 224.67$ cm^{-1} and

$$E_2 - E_1 = 1235.7\ \text{cm}^{-1}, \qquad E_3 - E_2 = 1460.4\ \text{cm}^{-1}, \qquad E_3 - E_2 = 1685.9\ \text{cm}^{-1}.$$

7. The scattering of particles by a crystal lattice can be described, under certain conditions, by a periodic potential that satisfies the translational invariance condition $V(\mathbf{r}) = V(\mathbf{r} + \mathbf{R})$, where $\mathbf{R}$ is a constant vector. Using the first Born approximation, show that the scattering amplitude (and therefore the cross-section) vanishes unless the condition $\boldsymbol{\kappa} \cdot \mathbf{R} = 2\pi n$, is satisfied, where n is an integer and $\boldsymbol{\kappa} = \mathbf{k}' - \mathbf{k}$, with $\mathbf{k}$ and $\mathbf{k}'$ being the wave vectors of the incoming and outgoing waves, respectively. This gives rise, for example, to Bragg's law for X-ray scattering.
   Hint: Make the variable change $\mathbf{r} = \mathbf{r}_0 + \mathbf{R}$ in the integral that gives the scattering amplitude.

Solution: In the first Born approximation, the scattering amplitude is given by

$$f(\theta, \varphi) = -\frac{m}{2\pi\hbar^2} \int \mathrm{d}^3\mathbf{r}_0 \, e^{i(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{r}_0} \, V(\mathbf{r}_0) = -\frac{m}{2\pi\hbar^2} \int \mathrm{d}^3\mathbf{r}_0 \, e^{i\boldsymbol{\kappa}\mathbf{r}_0} \, V(\mathbf{r}_0),$$

where $\boldsymbol{\kappa} = \mathbf{k}' - \mathbf{k}$. The translation symmetry of the potential, $V(\mathbf{r}) = V(\mathbf{r} + \mathbf{R})$, implies

$$\int \mathrm{d}^3\mathbf{r}_0\, e^{i\boldsymbol{\kappa}\mathbf{r}_0}\, V(\mathbf{r}_0) = \int \mathrm{d}^3\mathbf{r}_0\, e^{i\boldsymbol{\kappa}\mathbf{r}_0}\, V(\mathbf{r}_0 + \mathbf{R}).$$

Since $\mathbf{R}$ is a constant vector, we can change the integration variable to $\mathbf{r} = \mathbf{r}_0 + \mathbf{R}$ in the right-hand side of the previous equation,

$$\int \mathrm{d}^3\mathbf{r}_0\, e^{i\boldsymbol{\kappa}\mathbf{r}_0}\, V(\mathbf{r}_0 + \mathbf{R}) = \int \mathrm{d}^3\mathbf{r}\, e^{i\boldsymbol{\kappa}(\mathbf{r}-\mathbf{R})}\, V(\mathbf{r}).$$

Therefore, the symmetry of the potential leads to

$$\int \mathrm{d}^3\mathbf{r}\, e^{i\boldsymbol{\kappa}\mathbf{r}}\, V(\mathbf{r}) = \int \mathrm{d}^3\mathbf{r}\, e^{i\boldsymbol{\kappa}(\mathbf{r}-\mathbf{R})}\, V(\mathbf{r}) = e^{-i\boldsymbol{\kappa}\mathbf{R}} \int \mathrm{d}^3\mathbf{r}\, e^{i\boldsymbol{\kappa}\mathbf{r}}\, V(\mathbf{r}),$$

$$\left(1 - e^{-i\boldsymbol{\kappa}\mathbf{R}}\right) \int \mathrm{d}^3\mathbf{r}\, e^{i\boldsymbol{\kappa}\mathbf{r}} V(\mathbf{r}) = 0.$$

This equation holds when either of the following two conditions is satisfied

$$\int \mathrm{d}^3\mathbf{r}\, e^{i\boldsymbol{\kappa}\mathbf{r}}\, V(\mathbf{r}) = 0, \qquad \left(1 - e^{-i\boldsymbol{\kappa}\mathbf{R}}\right) = 0.$$

The first condition requires the scattering amplitude in the first Born approximation to be zero. Therefore we conclude that the Born amplitude is non-zero only when the second condition is satisfied, namely $e^{-i\boldsymbol{\kappa}\mathbf{R}} = 1$ for arbitrary $\boldsymbol{\kappa}$. This means that $\boldsymbol{\kappa}\mathbf{R} = 2\pi n$, with n any integer.

8. Suppose the electron in the hydrogen atom is subject to a perturbation similar to the spin-orbit interaction but quadratic in **LS** instead of linear, that is

$$\mathcal{H}_x = \xi(r)\, b\, (\mathbf{LS})^2 = \frac{e^2}{2m^2c^2r^3}\, b\, (\mathbf{LS})^2.$$

Deduce the energy shift this would cause in the hydrogen levels. Calculate the numerical value that the parameter b must have so that the energy difference between the doublet levels corresponding to $n = 2$ and $\ell = 1$ is a factor of 10^{-10} smaller than the value caused by the ordinary spin-orbit interaction (linear in **LS**).

Solution: When studying the spin-orbit interaction in hydrogen we found

$$\langle \ell s j m |\, \mathbf{LS}\, | \ell s j m \rangle = \frac{\hbar^2}{2}\Big[j(j+1) - \ell(\ell+1) - s(s+1)\Big],$$

and

$$E_{\mathrm{so}} = \frac{Z^4\alpha^2}{n^3}\, \frac{j(j+1) - \ell(\ell+1) - s(s+1)}{2\ell(\ell+\frac{1}{2})(\ell+1)} \quad \mathrm{Ry}.$$

It is easy to adapt these equations to study the effect of the Hamiltonian $\mathcal{H}_q$ given in this problem. First, the expectation value of $b(\mathbf{LS})^2$ is

$$\langle \ell s j m |\, b(\mathbf{LS})^2\, | \ell s j m \rangle = \frac{b\hbar^4}{4}\Big[j(j+1) - \ell(\ell+1) - s(s+1)\Big]^2.$$

The expectation value of the radial function does not change, therefore

$$E_q = \frac{Z^4\alpha^2}{n^3}\, \frac{b\hbar^2}{2}\, \frac{\big[j(j+1) - \ell(\ell+1) - s(s+1)\big]^2}{2\ell(\ell+\frac{1}{2})(\ell+1)} \quad \mathrm{Ry}.$$

The two energy levels corresponding to a given value of ℓ are obtained by setting $j = \ell \pm \frac{1}{2}$. Thus

$$\big[j(j+1) - \ell(\ell+1) - s(s+1)\big]^2 = \begin{cases} \ell^2 & \text{for } j = \ell + \frac{1}{2} \\ (\ell+1)^2 & \text{for } j = \ell - \frac{1}{2} \end{cases}$$

$$E_{q,(\ell+\frac{1}{2})} = \frac{Z^4\alpha^2 b\hbar^2}{4n^3}\,\frac{\ell^2}{\ell(\ell+\frac{1}{2})(\ell+1)}\ \text{Ry}, \qquad E_{q,(\ell-\frac{1}{2})} = \frac{Z^4\alpha^2 b\hbar^2}{4n^3}\,\frac{(\ell+1)^2}{\ell(\ell+\frac{1}{2})(\ell+1)}\ \text{Ry}$$

$$\Delta E_q = E_{q,(\ell+\frac{1}{2})} - E_{q,(\ell-\frac{1}{2})} = -\frac{Z^4\alpha^2 b\hbar^2}{4n^3}\,\frac{2\ell+1}{\ell(\ell+\frac{1}{2})(\ell+1)}\ \text{Ry}.$$

Note that, with this hypothetical Hamiltonian, the energy of the $j = \ell - 1/2$ level is increased, while the energy of the $j = \ell + 1/2$ level is decreased. This is contrary to the effect produced by the usual spin-orbit interaction. For $Z = 1$, $n = 2$ and $\ell = 1$, we have

$$\Delta E_q = E_{q,(\ell+\frac{1}{2})} - E_{q,(\ell-\frac{1}{2})} = -\frac{\alpha^2 b\hbar^2}{32}\ \text{Ry}.$$

The equivalent expression for the spin-orbit interaction is

$$\Delta E_{\text{so}} = E_{\text{so},(\ell+\frac{1}{2})} - E_{\text{so},(\ell-\frac{1}{2})} = \frac{\alpha^2}{16}\ \text{Ry}.$$

If the magnitude of the ratio is $|\Delta E_q/\Delta E_{\text{so}}| = 10^{-10}$, then the value of the b parameter is

$$b = \frac{2}{\hbar^2}\left|\frac{\Delta E_q}{\Delta E_{\text{so}}}\right| = \frac{2\times 10^{-10}}{\hbar^2} \;=\; 4.6\times 10^{20}\,(\text{eV}\cdot\text{s})^{-2}.$$

9. Suppose that the potential energy of interaction with the nucleus and electronic repulsion in the Helium atom can be described by a potential of the form

$$V(\mathbf{r}_1,\mathbf{r}_2) = -\frac{2e^2}{r_1}\left(1-\frac{\beta}{2r_1}\right) - \frac{2e^2}{r_2}\left(1-\frac{\beta}{2r_2}\right) + \frac{e^2}{|\mathbf{r}_1-\mathbf{r}_2|},$$

where β is a very small parameter.

(a) Using the variational method with the single-parameter wave function seen in class, $\psi(\mathbf{r}_1,\mathbf{r}_2) = \dfrac{\zeta^3}{\pi a_0^3}e^{-\zeta(r_1+r_2)/a_0}$, where a_0 is the Bohr radius, demonstrate that the expected value of the Hamiltonian as a function of the variational parameter ζ is given by

$$\langle\psi\,|\,\mathcal{H}\,|\,\psi\rangle = \left[-2(1+4\kappa)\zeta^2 + \frac{27}{4}\zeta\right]E_1,$$

where $E_1 = -e^2/2a_0$ is the Hydrogen ground-state energy and $\kappa = \beta/a_0$.

(b) Given the previous result, demonstrate that an estimate for the ground-state energy of Helium is

$$E_{\text{He}}^{\text{v}} = \left(\frac{27}{16}\right)^2\frac{2E_1}{1+4\kappa}.$$

(c) Treat the terms of the potential containing the parameter β as a perturbation, and use time-independent perturbation theory to first order to calculate the correction to the Helium energy (relative to the case where $\beta = \kappa = 0$). Using this correction, show that the Helium ground-state energy obtained via perturbation theory would be

$$E_{\text{He}}^{\text{p}} = \left(\frac{27}{16}\right)^2 2E_1(1-4\kappa).$$

Discuss the compatibility of the expressions obtained in sections (b) and (c).

Hints: Reuse/adapt the calculations from the class manual for Helium as needed, without repeating them. You may even reference equations from the manual by their number. Remember that $\langle 1/r^2\rangle = 2Z^2/a_0^2$ for the ground-state of hydrogen-like atoms.

Solution:

(a) The Hamiltonian can be conveniently written as

$$\mathcal{H} = \underbrace{-\frac{\hbar^2}{2m}(\nabla_1^2+\nabla_2^2) - e^2\left(\frac{\zeta}{r_1}+\frac{\zeta}{r_2}\right)}_{\psi \text{ is an eigenfunction of this part}} + e^2\left(\frac{\beta}{r_1^2}+\frac{\beta}{r_2^2}\right)$$
$$+ e^2\left(\frac{(\zeta-2)}{r_1}+\frac{(\zeta-2)}{r_2}+\frac{1}{|\mathbf{r_1}-\mathbf{r_2}|}\right) = \mathcal{H}_0 + e^2\left(\frac{\beta}{r_1^2}+\frac{\beta}{r_2^2}\right) = \mathcal{H}_0 + \mathcal{H}'.$$

The $\mathcal{H}_0$ term of the Hamiltonian $\mathcal{H} = \mathcal{H}_0 + \mathcal{H}'$ is identical to the one studied in the corresponding section of the lecture notes, therefore when computing $\langle\psi\,|\,\mathcal{H}\,|\,\psi\rangle = \langle\mathcal{H}_0\rangle + \langle\mathcal{H}'\rangle$ we can directly take the result

$$\langle\mathcal{H}_0\rangle = \left[2\zeta^2 - 4\zeta(\zeta-2) - \frac{5}{4}\zeta\right]E_1 = \left[-2\zeta^2+\frac{27}{4}\zeta\right]E_1,$$

and add the value of $\langle\mathcal{H}'\rangle$. Using the relation $\langle 1/r^2\rangle = 2Z^2/a_0^2$ for the ground-state hydrogen-like wavefunctions, and replacing $Z \to \zeta$, this is

$$\langle\mathcal{H}'\rangle = e^2\beta\left\langle\psi\left|\left(\frac{1}{r_1^2}+\frac{1}{r_2^2}\right)\right|\psi\right\rangle = 2e^2\beta\left\langle\psi\left|\frac{1}{r^2}\right|\psi\right\rangle = 4e^2\beta\frac{\zeta^2}{a_0^2} = -8\kappa\zeta^2E_1,$$

where the relation $e^2 = -2E_1a_0$ and the notation $\kappa = \beta/a_0$ were used. Getting all the pieces together,

$$\langle\mathcal{H}\rangle = \left[-2\zeta^2+\frac{27}{4}\zeta\right]E_1 - 8\kappa\zeta^2E_1 = \left[-2(1+4\kappa)\zeta^2+\frac{27}{4}\zeta\right]E_1.$$

(b) Minimize $\langle\mathcal{H}\rangle$ with respect to ζ to obtain

$$\frac{\mathrm{d}\langle\mathcal{H}\rangle}{\mathrm{d}\zeta} = 0 \implies \zeta = \frac{27}{16(1+4\kappa)}.$$

Upon inserting this value of ζ in the expression for $\langle\mathcal{H}\rangle$ we get

$$E_{\mathrm{He}}^{\mathrm{v}} = \left(\frac{27}{16}\right)^2\frac{2E_1}{1+4\kappa}.$$

(c) The energy correction at first order in perturbation theory is given by the value of $\langle\mathcal{H}'\rangle$ obtained in part (a), but setting ζ to the value that takes when $\kappa = 0$, which is $\zeta = 27/16$. Therefore

$$\Delta E = \langle\mathcal{H}'\rangle = -8\kappa\zeta^2E_1 = -8\kappa\left(\frac{27}{16}\right)^2E_1.$$

The energy of the ground state when $\kappa = 0$ is $(27/16)^2 2E_1$. This must be added to ΔE to obtain

$$E_{\mathrm{He}}^{\mathrm{p}} = \left(\frac{27}{16}\right)^2 2E_1 - 8\kappa\left(\frac{27}{16}\right)^2E_1 = \left(\frac{27}{16}\right)^2 2E_1(1-4\kappa).$$

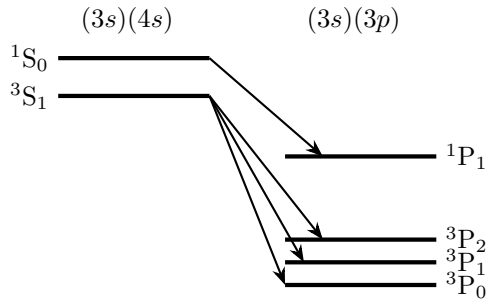
If 4κ is very small, the following approximation holds

$$\frac{1}{1+4\kappa} \approx 1-4\kappa.$$

This proves that, in the $\kappa \ll 1$ regime, the results for $E_{\mathrm{He}}^{\mathrm{v}}$ and $E_{\mathrm{He}}^{\mathrm{p}}$ are compatible.

10. The attached figure shows (not to scale) two excited configurations of magnesium ($Z = 12$), along with their possible terms and levels. Reproduce the diagram from the figure, and using Hund's rules, indicate the quantum numbers L, S, and J for each level (considering normal or regular ordering). Include in the drawing all the allowed (E1) transitions between the levels of the $(3s)(4s)$ configuration and those of the $(3s)(3p)$ configuration. Three of the emissions form a triplet with wavelengths of 516.88, 517.41, and 518.50 nm. Mark those transitions in the diagram and calculate the proportionality factor of Landé's interval rule for this triplet, giving your result in eV. Also, calculate the value of $\zeta(\alpha, L, S)$ in $1/(\mathrm{J}\cdot\mathrm{s}^2)$. Data: $\hbar = 6.58 \times 10^{-16}\,\mathrm{eV}\cdot\mathrm{s}$.

Solution:
The $(3s)(4s)$ configuration has two inequivalent electrons, so all combinations of L and S are possible. We have $L = 0$ and $S = 0, 1$, giving the levels $^1\mathrm{S}_0$ and $^3\mathrm{S}_1$. The $(3s)(3p)$ configuration has also two inequivalent electrons, but now with $L = 1$ and $S = 0, 1$, so the possible terms and levels are $^1\mathrm{P}_1$, $^3\mathrm{P}_0$, $^3\mathrm{P}_1$, and $^3\mathrm{P}_2$. To order energetically the levels in each configuration note that, according to Hund's first rule, the term with the highest total spin, S, has the lowest energy.



From this we see that $E(^1\mathrm{S}_0) > E(^3\mathrm{S}_1)$ in the $(3s)(4s)$ configuration, and $E(^1\mathrm{P}) > E(^3\mathrm{P})$ in the $(3s)(3p)$ configuration. To order the $^3\mathrm{P}_\mathrm{J}$ levels we apply the third rule: since the shell is less than half-full, $E(^3\mathrm{P}_2) > E(^3\mathrm{P}_1) > E(^3\mathrm{P}_0)$. The figure above shows the quantum numbers of every level.

The rules for allowed (E1) transtions in LS coupling are $\Delta S = 0$, $\Delta L = 0, \pm 1$ (except $0 \leftrightarrow 0$), and parity change. The only four transitions verifying these conditions are shown in the figure. The Landé interval rule states that

$$E(J) - E(J-1) = \hbar^2 \zeta(\alpha, L, S)\, J.$$

Since $hc \approx 1240\,\mathrm{eV}\cdot\mathrm{nm}$, emission wavelengths can be converted to eV through the relation $E = hc/\lambda = 1240/\lambda\,\mathrm{eV}$, when λ is expressed in nm. Two energy differences can be computed

$$E(^3\mathrm{P}_2) - E(^3\mathrm{P}_1) = 2\hbar^2\zeta(\alpha,1,1) = \frac{hc}{517.41} - \frac{hc}{518.50} = \frac{1240}{2.46\times10^5}\,\mathrm{eV} = 0.0050\,\mathrm{eV} = 8\times10^{-21}\,\mathrm{J}.$$

$$E(^3\mathrm{P}_1) - E(^3\mathrm{P}_0) = \hbar^2\zeta(\alpha,1,1) = \frac{hc}{516.88} - \frac{hc}{517.41} = 0.0025\,\mathrm{eV} = 4\times10^{-21}\,\mathrm{J}.$$

From the first equation we obtain $\hbar^2\zeta = 0.0025\,\mathrm{eV} = 4\times10^{-21}\,\mathrm{J}$. The second equation gives also the same value within uncertainties (not shown here). The value of ζ can then be computed as

$$\zeta(\alpha,1,1) = \frac{0.0025\,\mathrm{eV}}{\hbar^2} = \frac{0.0025\,\mathrm{eV}}{(6.58\times10^{-16}\,\mathrm{eV}\cdot\mathrm{s})^2} = 5.77\times10^{27}\,\frac{1}{\mathrm{eV}\cdot\mathrm{s}^2}$$
$$= \frac{5.77\times10^{27}}{1.6\times10^{-19}}\,\frac{1}{\mathrm{J}\cdot\mathrm{s}^2} = 3.61\times10^{46}\,\frac{1}{\mathrm{J}\cdot\mathrm{s}^2}.$$

11. Neutrons with a kinetic energy of 50 MeV scatter elastically off a nucleus. The measured phase shifts in this experiment are $\delta_0 = 95°$, $\delta_1 = 72°$, $\delta_2 = 60°$, $\delta_3 = 35°$,

$\delta_4 = 18°$ e $\delta_5 = 5°$. All other phase shifts are negligible, i.e., $\delta_\ell \approx 0$ para $\ell \geq 6$. With these data, calculate the total cross-section, expressing the result in fm^2 and barn. Under the assumption that we are in a sufficiently high-energy case, estimate the radius of the nucleus acting as the scattering center, providing the result in fm. Use non-relativistic kinematics. Data: neutron mass $m_n = 939.57\,\text{MeV}/c^2$, $\hbar c = 197.33$ MeV fm.

Solution: The total cross-section is given by

$$\sigma = \frac{4\pi}{k^2}\sum_{\ell=0}^{5}(2\ell+1)\,\sin^2\delta_\ell$$
$$= \frac{4\pi}{k^2}\Big(\sin^2\delta_0 + 3\sin^2\delta_1 + 5\sin^2\delta_2 + 7\sin^2\delta_3 + 9\sin^2\delta_4 + 11\sin^2\delta_5\Big) = \frac{4\pi}{k^2}\times 10.702.$$

To compute k we use the (non-relativistic) relation

$$k^2 = \frac{2mE}{\hbar^2} = \frac{2mc^2E}{(\hbar c)^2} = \frac{2\times 939.57\times 50}{197.33^2}\,\text{fm}^{-2} = 2.41\,\text{fm}^{-2}.$$

The cross-section is then

$$\sigma = \frac{4\pi}{2.41}\times 10.702\,\text{fm}^2 = 55.80\,\text{fm}^2 = 0.5580\,\text{barn}.$$

To estimate the radius of the nucleus we could try to make the bold assumption that the maximum value of orbital angular momentum contributing to the scattering process is $\ell_{\max} = kR$, where R is the range of the nuclear potential, presumably similar to the nucleus radius. In this experiment the contribution of partial waves with $\ell > 5$ is negligible, therefore

$$R \approx \frac{\ell_{\max}}{k} = \frac{5}{\sqrt{2.41}}\,\text{fm} = 3.2\,\text{fm}.$$

Alternatively one can consider that for a hard sphere in the high energy limit the cross-section is given by:

$$\sigma = 2\pi R^2.$$

Thus

$$R = \sqrt{\frac{\sigma}{2\pi}} = \sqrt{\frac{55.80}{2\pi}}\,\text{fm} \approx 3.0\,\text{fm},$$

which is a very similar result.

12. A Litium atom is in the excited configuration $1s^2 2p$. Suppose the wave function of the $2p$ orbital is given by

$$\Psi_{2p} = Are^{-br/a}Y_{1m}(\theta,\varphi),$$

where a is the Bohr radius, $b = 0.52$, and A is a normalization constant. The $2p$ electron feels a potential $\phi(r) = eZ_{\text{eff}}/r$ with $Z_{\text{eff}}(r) = 1 + 2e^{-4r/a}(1 + 2(r/a))$.

a) Compute the value of the normalization constant A.

b) The spin-orbit interaction is given by $\mathcal{H}_{\text{so}} = \xi(r)\mathbf{LS}$. Find out $\xi(r)$ and prove that $\langle\xi(r)\rangle = 0.0324\,e^2/(m^2c^2a^3)$.

c) Compute the energy splitting due to the spin-orbit interaction between the states $P_{3/2}$ and $P_{1/2}$.

Hint: $\int_0^\infty e^{-ax}x^n\,\text{d}x = n!/a^{n+1}$.

Solution:

a) From the text it is easy to identify $R_{n\ell}(r) = Are^{-br/a}$. The normalization constant can be easily computed using the relation $\int_0^\infty e^{-ax}x^n\,\mathrm{d}x = n!/a^{n+1}$,

$$1 = \int_0^\infty A^2r^2e^{-2br/a}r^2\,\mathrm{d}r = A^2\frac{4!}{(2b/a)^5} \implies A^2 = \frac{4}{3}(b/a)^5.$$

b) The spin-orbit interaction is given by eq. (1.145) of the handouts

$$\mathcal{H}_{\mathrm{so}} = -\frac{e}{2m^2c^2}\frac{1}{r}\frac{d\phi}{dr}\mathbf{LS} = \xi(r)\mathbf{LS}.$$

This hamiltonian is diagonal in the coupled basis and, therefore, the energy shift is given by

$$\langle 2\ell sjm|\mathcal{H}_{\mathrm{so}}|2\ell sjm\rangle = \langle 2\ell sjm|\xi(r)\mathbf{LS}|2\ell sjm\rangle = \langle 2\ell|\xi(r)|2\ell\rangle\,\langle \ell sjm|\mathbf{LS}|\ell sjm\rangle\,.$$

Eq. (1.153) in the handouts gives

$$\langle \ell sjm|\mathbf{LS}|\ell sjm\rangle = \frac{\hbar^2}{2}\left[j(j+1) - \ell(\ell+1) - s(s+1)\right].$$

While

$$\xi(r) = -\frac{e^2}{2m^2c^2}\frac{1}{r}\frac{d}{dr}\left(\frac{1+2e^{-4r/a}(1+2r/a)}{r}\right) = \frac{e^2}{2m^2c^2}\frac{1+2e^{-4r/a}(1+4r/a+8(r/a)^2)}{r^3},$$

and

$$\langle\xi(r)\rangle = \xi_{n\ell} = \langle 2\ell|\xi(r)|2\ell\rangle = \int_0^\infty \xi(r)R_{n\ell}^2r^2\,\mathrm{d}r,$$

where part of eq. (1.154) was used.
Using the result of a) for A we get:

$$\begin{aligned}
\langle\xi(r)\rangle &= \int_0^\infty \xi(r)R_{n\ell}^2r^2\,\mathrm{d}r \\
&= \int_0^\infty \frac{e^2}{2m^2c^2}\frac{1+2e^{-4r/a}(1+4r/a+8(r/a)^2)}{r^3}\frac{4}{3}(b/a)^5r^2e^{-2br/a}r^2\,\mathrm{d}r \\
&= \frac{2e^2b^5}{3m^2c^2a^3}\int_0^\infty \left[r/a + 2e^{-4r/a}(r/a+4(r/a)^2+8(r/a)^3)\right]e^{-2br/a}\,d(r/a) \\
&= \frac{2e^2b^5}{3m^2c^2a^3}\int_0^\infty \left[x + 2e^{-4x}(x+4x^2+8x^3)\right]e^{-2bx}\,dx \\
&= \frac{2e^2b^5}{3m^2c^2a^3}\left[\frac{1}{(2b)^2} + 2\left(\frac{1}{(2b+4)^2} + 4\frac{2}{(2b+4)^3} + 8\frac{3!}{(2b+4)^4}\right)\right] \\
&= \frac{2e^2b^5}{3m^2c^2a^3}\left[\frac{1}{(2b)^2} + \frac{2}{(2b+4)^4}\left((2b+4)^2 + 8(2b+4) + 48\right)\right] \\
&= \frac{e^2b^5}{3m^2c^2a^3}\left[\frac{1}{2b^2} + \frac{b^2+8b+24}{(b+2)^4}\right] = 0.0324\frac{e^2}{m^2c^2a^3}.
\end{aligned}$$

c) We can now get

$$\begin{aligned}
\langle 2\ell sjm|\mathcal{H}_{\mathrm{so}}|2\ell sjm\rangle &= \langle 2\ell|\xi(r)|2\ell\rangle\,\langle \ell sjm|\mathbf{LS}|\ell sjm\rangle \\
&= 0.0324\frac{e^2\hbar^2}{2m^2c^2a^3}\left[j(j+1) - \ell(\ell+1) - s(s+1)\right] \\
&= 0.0324\frac{\hbar^2}{m^2c^2a^2}\left[j(j+1) - \ell(\ell+1) - s(s+1)\right]\frac{e^2}{2a} \\
&= 0.0324\alpha^2\left[j(j+1) - \ell(\ell+1) - s(s+1)\right]\mathrm{Ry},
\end{aligned}$$

where the relations $a = \hbar/mc\alpha$ and $e^2/2a = 1\,\mathrm{Ry}$ were used and also the result for $\langle \ell sjm|\mathbf{LS}|\ell sjm\rangle$ given in a).
The $P_{3/2}$ level corresponds to the quantum numbers $j = 3/2$, $\ell = 1$ and $s = 1/2$, whilst the $P_{1/2}$ level has quantum numbers $j = 1/2$, $\ell = 1$ and $s = 1/2$. Therefore

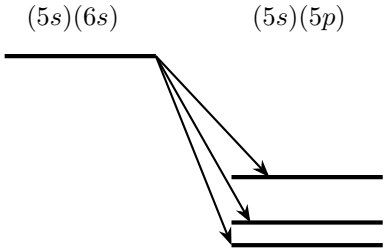
$$\begin{aligned}
\Delta_{\mathrm{so}} &= \langle 21\tfrac{1}{2}\tfrac{3}{2}|\mathcal{H}_{\mathrm{so}}|21\tfrac{1}{2}\tfrac{3}{2}\rangle - \langle 21\tfrac{1}{2}\tfrac{1}{2}|\mathcal{H}_{\mathrm{so}}|21\tfrac{1}{2}\tfrac{1}{2}\rangle = 0.0324\alpha^2\left[\frac{3}{2}\left(\frac{3}{2}+1\right) - \frac{1}{2}\left(\frac{1}{2}+1\right)\right]\mathrm{Ry} \\
&= 0.0324\alpha^2 3\,\mathrm{Ry} = 0.0972\alpha^2\,\mathrm{Ry} = 5.18\times10^{-6}\,\mathrm{Ry} = 7.97\times10^{-5}\,\mathrm{eV} = 0.56\,\mathrm{cm}^{-1}.
\end{aligned}$$

13. The atomic number of Cadmium is $Z = 48$, and the electronic configuration of its fundamental state is $[\mathrm{Kr}]4\mathrm{d}^{10}5\mathrm{s}^2$. The 508.6 nm Cadmium emission line corresponds to the $5\mathrm{s}6\mathrm{s}\,{}^3\mathrm{S}_1 \to 5\mathrm{s}5\mathrm{p}\,{}^3\mathrm{P}_2$ transition and belongs to a triplet, with the remaining two emission lines having wavelengths of 480.0 nm and 467.8 nm.

    a) Work out wich atomic levels are involved in the last two emission lines mentioned above.

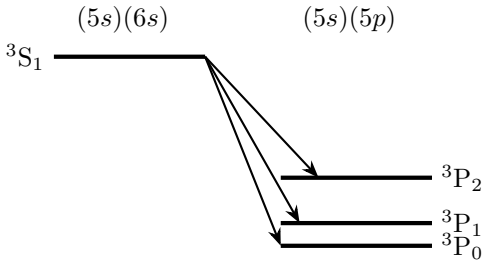
    b) Find the fine-structure splitting of the Cadmin $5\mathrm{s}5\mathrm{p}\,{}^3\mathrm{P}$ term. Is this splitting compatible with Landé interval rule?

    c) Give reasons why the observation of this triplet in an absorption sepctrum would be highly improbable at room temperature.



Solution:

a) We have an initial 5s6s configuration. The possible values of the total spin and orbital angular momentum are $S = 0, 1$ and $L = 0$, respectively. For the final configuration we have $S = 0, 1$ and $L = 1$. Since the initial level is ${}^3\mathrm{S}_1$, only transitions to final states with $S = 1$ are allowed in LS coupling. These would be the ones corresponding to the $S = 1$ and $L = 1$ ${}^3\mathrm{P}$ term. Three $J = 0, 1, 2$ values are possible. If a normal ordering is assumed in the multiplet, the levels involved in the three transitions are those shown in the figure.



b) The longest wavelength corresponds to the transition from the ${}^3\mathrm{S}_1$ level to the ${}^3\mathrm{P}_2$ level:

$$E({}^3\mathrm{S}_1) - E({}^3\mathrm{P}_2) = \frac{hc}{508.6}\,\frac{1}{[\mathrm{nm}]}.$$

Similarly

$$E({}^3\mathrm{S}_1) - E({}^3\mathrm{P}_1) = \frac{hc}{480.0}\,\frac{1}{[\mathrm{nm}]}, \qquad E({}^3\mathrm{S}_1) - E({}^3\mathrm{P}_0) = \frac{hc}{467.8}\,\frac{1}{[\mathrm{nm}]}.$$

Consequently:

$$E({}^3\mathrm{P}_2) - E({}^3\mathrm{P}_1) = hc\left(\frac{1}{480.0} - \frac{1}{508.6}\right)\frac{1}{[\mathrm{nm}]} = 1239.8\left(\frac{1}{480.0} - \frac{1}{508.6}\right)[\mathrm{eV}] = 0.145\,[\mathrm{eV}],$$
$$E({}^3\mathrm{P}_1) - E({}^3\mathrm{P}_0) = hc\left(\frac{1}{467.8} - \frac{1}{480.0}\right)\frac{1}{[\mathrm{nm}]} = 1239.8\left(\frac{1}{467.8} - \frac{1}{480.0}\right)[\mathrm{eV}] = 0.067\,[\mathrm{eV}].$$

The Landé interval rule states that

$$E(J) - E(J-1) = \hbar^2\zeta(\alpha, L, S)\,J.$$

Therefore,

$$E({}^3\mathrm{P}_2) - E({}^3\mathrm{P}_1) = \hbar^2\zeta(\alpha, L, S)\,2,$$
$$E({}^3\mathrm{P}_1) - E({}^3\mathrm{P}_0) = \hbar^2\zeta(\alpha, L, S)\,1.$$

The rule is fulfilled if

$$\frac{E({}^3\mathrm{P}_2) - E({}^3\mathrm{P}_1)}{E({}^3\mathrm{P}_1) - E({}^3\mathrm{P}_0)} \approx 2,$$

which is almost the case, since:

$$\frac{E(^3\mathrm{P}_2) - E(^3\mathrm{P}_1)}{E(^3\mathrm{P}_1) - E(^3\mathrm{P}_0)} = 2.164.$$

c) In absorption spectra atoms are initially in their ground-state level (in this case, $5s^2\,{}^1S_0$) at room temperature, because $k_B T \approx 0.026$ eV at 300 K. This implies that we would only observe transitions from the ground state to some of the excited states with significant probablity. However, none of the quoted triplet lines fall into that category because they are associated to transitions between excited states, hence these lines are not expected to be seen in absorption spectra.

Furthermore, transitions from the ${}^1S_0$ ground-state level to any of the $5s5p\,{}^3P$ excited levels would be highly suppressed because they imply a $\Delta S = 1$ change, which violates the E1 selection rules. Consequently, the triplet initial states in the transition would be in any case very poorly populated for a significant absorption process to take place.

14. Neutrons of 170 MeV kinetic energy scatter elastically off a nucleus with radius $R = 1.05$ fm. Under the hypthesis of high-energy scattering, make an estimation of the maximum angular momentum value $\ell_{\max}$ of the partial wave that significantly contributes to the cross-section. Suppose the phase-shifts measured in this experiment are given by the relation $\delta_\ell = \pi/(\ell+2)^2$. Compute the cross-section due to the $\ell = 0$ partial wave, and call it σ_0. Compute the total cross-section, σ_t, and the ratio σ_0/σ_t. Use non-relativistic kinematics.

Solution: To estimate the value of $\ell_{\max}$ we make the assumption that the maximum value of orbital angular momentum contributing to the scattering process is $\ell_{\max} \approx kR$, where $\hbar k$ is the momentum of the neutron. For neutrons with $E = 170$ MeV the value of k^2 is

$$k^2 = \frac{2mE}{\hbar^2} = \frac{2mc^2E}{(\hbar c)^2} = \frac{2 \times 939.57 \times 170}{197.33^2}\,\mathrm{fm}^{-2} = 8.20\,\mathrm{fm}^{-2}.$$

Therefore

$$\ell_{\max} = kR = \sqrt{8.20} \times 1.05 \approx 3.$$

The total cross-section and the s-wave cross-section are given by

$$\sigma = \frac{4\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell+1)\,\sin^2\delta_\ell, \qquad \sigma_0 = \frac{4\pi}{k^2}\sin^2\frac{\pi}{4} = \frac{2\pi}{k^2.}$$

The s-wave cross-section is then

$$\sigma_0 = \frac{2\pi}{8.20}\,\mathrm{fm}^2 = 0.77\,\mathrm{fm}^2 = 0.0077\,\mathrm{barn}^2.$$

The total cross-section is

$$\begin{aligned}\sigma_t &= \frac{4\pi}{k^2}\sum_{\ell=0}^{3}(2\ell+1)\,\sin^2\left(\frac{\pi}{(\ell+2)^2}\right)\\ &= \frac{4\pi}{k^2}\left[\sin^2\left(\frac{\pi}{4}\right) + 3\sin^2\left(\frac{\pi}{9}\right) + 5\sin^2\left(\frac{\pi}{16}\right) + 7\sin^2\left(\frac{\pi}{25}\right)\right] = \frac{4\pi}{k^2} \times 1.15 = 1.76\,\mathrm{fm}.\end{aligned}$$

The s-wave cross-section is close to half of the total cross-section

$$\frac{\sigma_0}{\sigma_t} = 0.44$$

15. A hydrogen atom, where the fine-structure is neglected, is placed under a perturba-

tion given by

$$V = \frac{V_0}{a_0^2} xy, \qquad V_0 \approx 10^{-2}\,\mathrm{eV}, \qquad a_0 = \frac{\hbar^2}{me^2}.$$

(a) Write V in terms of spherical harmonics and prove that the relations $\Delta\ell \equiv \ell - \ell' = 0, \pm 2$ and $\Delta m \equiv m - m' = \pm 2$ must hold for the matrix elements $\langle n, \ell, m \,|\, V \,|\, n', \ell', m' \rangle$ to vanish.

(b) Compute the matrix elements of the perturbation V for the levels with principal quantum numbers $n = 1, 2$. Find the eigenvalues to first order and the corresponding approximate eigenstates. You may need the following results

$$\left\langle r^2 \right\rangle_{n\ell} = a_0^2 \frac{n^2}{2} \left[5n^2 + 1 - 3\ell(\ell+1) \right],$$

$$\begin{pmatrix} 1 & 2 & 1 \\ -1 & 2 & -1 \end{pmatrix} = \sqrt{\frac{1}{5}}, \qquad \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \sqrt{\frac{2}{15}}.$$

Solution:

(a) In spherical coordinates,

$$V = \frac{V_0}{a_0^2} r^2 \sin^2\theta \, \sin\varphi \cos\varphi = \frac{V_0}{2a_0^2} r^2 \sin^2\theta \, \sin 2\varphi = -i \frac{V_0}{4a_0^2} r^2 \sin^2\theta \, \left(e^{i2\varphi} - e^{-i2\varphi} \right)$$
$$= -i \sqrt{\frac{2\pi}{15}} \frac{V_0}{a_0^2} r^2 \left(Y_2^2 - Y_2^{-2} \right),$$

where $Y_2^{\pm 2} = \sqrt{\frac{15}{32\pi}} \sin^2\theta \, e^{\pm i2\varphi}$ are the spherical harmonics for $\ell = 2$ and $m = \pm 2$. The angular part of the matrix elements can be obtained with Gaunt's formula (B.192):

$$\left\langle \ell, m \,\middle|\, Y_2^{\pm 2} \,\middle|\, \ell', m' \right\rangle \;=\; (-1)^m \sqrt{\frac{5}{4\pi}(2\ell+1)(2\ell'+1)} \begin{pmatrix} \ell & 2 & \ell' \\ -m & \pm 2 & m' \end{pmatrix} \begin{pmatrix} \ell & 2 & \ell' \\ 0 & 0 & 0 \end{pmatrix}.$$

The selection rules derive from 3j properties summarised in eq. (B.182) of the handouts and parity conservation. For the 3j-coefficients not to be null it has to be satisfied that the sum of the elements of the second row cancel. For the first 3j in the previous equation this means that $-m \pm 2 + m' = 0$ and, therefore: $\Delta m \equiv m - m' = \pm 2$. Another thing that 3j coefficients must fulfill to be different from zero is the triangular inequality of the elements of the first row. In this case $\Delta(\ell, 2, \ell')$, which translates into $|\ell - \ell'| \leq 2$. Parity conservation will also make the 3j coefficient cancel if $|\ell - \ell'|$ is odd. From these two last conditions it can be concluded that $\Delta\ell = \ell - \ell' = \pm 2$.

(b) We are asked to obtain the matrix elements for the states with $n = 1, 2$. Since this is a perturbation we consider sets of states with the same energy without the perturbation. So, we separately consider states with $n = 1$ and then later the states with $n = 2$.

For $n = 1$ we have a single state and a single element matrix for V, which gives us the energy shift due to the perturbation. Also $\langle 1, 0, 0 \,|\, V \,|\, 1, 0, 0 \rangle = 0$, because $\Delta m = 0$. This means that V creates no energy shift at first order in perturbation theory in the ground state of hydrogen.

According to our $\Delta m = \pm 2$ selection rule, the only non-zero matrix elements between the $n = 2$ states are $\langle 2, 1, 1 \,|\, V \,|\, 2, 1, -1 \rangle$ and its complex conjugate, $\langle 2, 1, -1 \,|\, V \,|\, 2, 1, 1 \rangle$. Therefore

$$\langle 2, 1, 1 \,|\, V \,|\, 2, 1, -1 \rangle$$
$$= -i \frac{V_0}{a_0^2} \sqrt{\frac{2\pi}{15}} \int r^2 R_{21}^2 \, r^2 \, \mathrm{d}r \, \left[\left\langle 1, 1 \,\middle|\, Y_2^2 \,\middle|\, 1, -1 \right\rangle - \left\langle 1, 1 \,\middle|\, Y_2^{-2} \,\middle|\, 1, -1 \right\rangle \right]$$
$$= -i \frac{V_0}{a_0^2} \sqrt{\frac{2\pi}{15}} \int r^2 R_{21}^2 \, r^2 \, \mathrm{d}r \, \left\langle 1, 1 \,\middle|\, Y_2^2 \,\middle|\, 1, -1 \right\rangle.$$

The last equality uses that $\langle 1,1\,|\,Y_2^{-2}\,|\,1,-1\rangle = 0$. Furthermore,

$$\langle 1,1\,|\,Y_2^2\,|\,1,-1\rangle = -\sqrt{\frac{5}{4\pi}\times 3\times 3}\begin{pmatrix} 1 & 2 & 1\\ -1 & 2 & -1\end{pmatrix}\begin{pmatrix} 1 & 2 & 1\\ 0 & 0 & 0\end{pmatrix}$$
$$= -\sqrt{\frac{5}{4\pi}\times 3\times 3}\sqrt{\frac{1}{5}}\sqrt{\frac{2}{15}} = -\sqrt{\frac{3}{10\pi}},$$

and, given eq. (1.132) of the handouts, for the hydrogen atom $\langle r\rangle_{n\ell} = \frac{a_0^2 n^2}{2}[5n^2+1-3\ell(\ell+1)]$,

$$\int r^2 R_{21}^2\, r^2\,\mathrm{d}r = \langle r^2\rangle_{21} = 30\, a_0^2.$$

Therefore

$$\langle 2,1,1\,|\,V\,|\,2,1,-1\rangle = i\frac{V_0}{a_0^2}\sqrt{\frac{2\pi}{15}}\,30a_0^2\sqrt{\frac{3}{10\pi}} = i6V_0.$$

The V matrix elements for the $|2,0,0\rangle$, $|2,1,0\rangle$, $|2,1,1\rangle$ and $|2,1,-1\rangle$ states, taken in this order, is then

$$\begin{pmatrix} 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & i6V_0\\ 0 & 0 & -i6V_0 & 0\end{pmatrix}$$

The secular equation for the lower-right 2×2 box is

$$\begin{vmatrix} -\lambda & i6V_0\\ -i6V_0 & -\lambda\end{vmatrix} = \lambda^2 - 36V_0^2 = 0 \implies \lambda = \pm 6V_0.$$

The perturbation V does not modify the energy of the states $|2,0,0\rangle$ and $|2,1,0\rangle$, which we may denote as E_0. To find the eigenstante corresponding to the eigenvalue $6V_0$ we do

$$\begin{pmatrix} 0 & i6V_0\\ -i6V_0 & 0\end{pmatrix}\begin{pmatrix} a\\ b\end{pmatrix} = 6V_0\begin{pmatrix} a\\ b\end{pmatrix} \implies b = -ia.$$

So the corresponding normalized eigenstate is

$$\frac{1}{\sqrt{2}}\left[\begin{pmatrix} 1\\ 0\end{pmatrix} - i\begin{pmatrix} 0\\ 1\end{pmatrix}\right] = \frac{1}{\sqrt{2}}\Big(\,|2,1,1\rangle - i\,|2,1,-1\rangle\,\Big).$$

For the eigenvalue $-6V_0$ we have

$$\begin{pmatrix} 0 & i6V_0\\ -i6V_0 & 0\end{pmatrix}\begin{pmatrix} a\\ b\end{pmatrix} = -6V_0\begin{pmatrix} a\\ b\end{pmatrix} \implies b = ia.$$

The corresponding normalized eigenstate is

$$\frac{1}{\sqrt{2}}\left[\begin{pmatrix} 1\\ 0\end{pmatrix} + i\begin{pmatrix} 0\\ 1\end{pmatrix}\right] = \frac{1}{\sqrt{2}}\Big(\,|2,1,1\rangle + i\,|2,1,-1\rangle\,\Big).$$

In summary, after the perturbation the states $|2,0,0\rangle$ and $|2,1,0\rangle$ do not experience an energy shift while the state $\frac{1}{\sqrt{2}}\Big(\,|2,1,1\rangle - i\,|2,1,-1\rangle\,\Big)$ energy is shifted $+6V_0$, and the state $\frac{1}{\sqrt{2}}\Big(\,|2,1,1\rangle + i\,|2,1,-1\rangle\,\Big)$ energy is shifted $-6V_0$ from their unperturbed energy.

16. A hydrogen atom in its first excited state, $2p$, is placed inside a cavity in thermal equilibrium with radiation of energy density $\rho(\omega)$ per unit frequency. At what temperature of the cavity are the transition probabilities for spontaneous and induced emission equal? (the value of Boltzmann's constant is $k_{\mathrm{B}} = 8.62\times 10^{-5}\,\mathrm{eV/K}$).

Solution: The induced and spontaneous emission transition probabilities are, respectively (a sub-index $2p \to 1s$ is omitted in the probabilities, the Einstein coefficients and the photon frequencies):

$$\begin{aligned} W^i =& B \cdot \rho(\omega) = \frac{\pi^2 c^3}{\hbar \omega^3} A \rho(\omega), \\ W^s =& A, \end{aligned}$$

where A and B are the Einstein coefficients, ω is the frequency of the photon emitted in the $2p \to 1s$ transition,

$$\hbar\omega = \left(\frac{1}{1} - \frac{1}{2^2}\right) R_{\mathrm{H}} = \frac{3}{4} R_{\mathrm{H}},$$

and ρ is the density of photons per unit frequency:

$$\rho(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{\exp\left(\frac{\hbar\omega}{k_B T}\right) - 1},$$

where k_B is the Boltzmann constant and T the temperature.

These two transition probabilities are equal if

$$A = \frac{\pi^2 c^3}{\hbar\omega^3} A\rho(\omega) \implies \frac{\pi^2 c^3}{\hbar\omega^3}\rho(\omega) = 1 \implies \exp\left(\frac{\hbar\omega}{k_B T}\right) = 2 \implies T = \frac{\hbar\omega}{k_B \ln 2}.$$

Considering that $k_{\mathrm{B}} = 8.62 \times 10^{-5}$ eV/K and $R_{\mathrm{H}} = 13,6$ eV:

$$T = \frac{\frac{3}{4} \times 13.6}{8.62 \times 10^{-5} \times \ln 2}\,\mathrm{K} = 1.71 \times 10^5\,\mathrm{K}.$$

17. In an experimental setup neutral pions, π^0, with kinetic energy $T = 300$ MeV, are scattered off a heavy nuclei with radius 1.4 fm. It can be assumed that the nucleus behaves as a completely absorbing sphere. Assuming also that we are working in a high-energy regime,

    (a) Find the maximum value of the incident π^0 orbital angular momentum, $\ell_{\max}$, that needs to be taken into account to compute the cross-section (use relativistic kinematics).

    (b) Compute the total elastic cross section and the total inelastic cross section.

    (c) Comute the elastic scattering amplitude and verify the validity of the optical theorem.

    You may find useful the following Legendre polynomials: $P_0 = 1$, $P_1 = x$, $P_2 = \frac{1}{2}(3x^2 - 1)$, $P_3 = \frac{1}{2}(5x^3 - 3x)$, $P_4 = \frac{1}{8}(35x^4 - 30x^2 + 3)$, e que a masa do π^0 é $m_{\pi^0} = 134.98$ MeV$/c^2$.

Solution:

(a) To estimate the value of $\ell_{\max}$ we make the assumption that the maximum value of orbital angular momentum contributing to the scattering process is $\ell_{\max} \approx kR$, where $\hbar k$ is the momentum of the pion. Pions with a kinetic energy of $T = 300$ MeV are relativistic, so their momentum is given by $p^2c^2 = E^2 - m^2c^4 = (T + mc^2)^2 - m^2c^4 = 2mc^2T + T^2$, and the corresponding value of k^2 is

$$k^2 = \frac{2mc^2T}{(\hbar c)^2} + \frac{T^2}{(\hbar c)^2} = \left(\frac{2 \times 135 \times 300}{197.33^2} + \frac{300^2}{197.33^2}\right)\mathrm{fm}^{-2} = 4.39\,\mathrm{fm}^{-2}.$$

Therefore

$$\ell_{\max} \approx kR = \sqrt{4.39} \times 1.4 \approx 3.$$

(b) Since we are assuming the target nuclei to be totally absorbing, incident pions with $\ell \leq 3$ will be absorbed and the inelasticiy must be chosen to be $\eta_\ell = 0$ for $\ell \leq 3$. On the other hand, partial waves with $\ell > 3$ do not contribute to the scattering. The total inelastic (or absorption) cross-section is then

$$\sigma_{\rm tot}^{\rm abs} = \frac{\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell+1)(1-\eta_\ell^2) = \frac{\pi}{k^2}\sum_{\ell=0}^{3}(2\ell+1) = \frac{16\pi}{k^2} = \frac{16\pi}{4.51}\,{\rm fm}^2 = 11.1\,{\rm fm}^2.$$

The total elastic cross-section is

$$\sigma_{\rm tot}^{\rm el} = \frac{\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell+1)\left|1-\eta_\ell\, e^{2i{\rm Re}(\delta_\ell)}\right|^2 = \frac{\pi}{k^2}\sum_{\ell=0}^{3}(2\ell+1) = \sigma_{\rm tot}^{\rm abs} = 11.1\,{\rm fm}^2.$$

The total cross-section is

$$\sigma_{\rm tot} = \sigma_{\rm tot}^{\rm abs} + \sigma_{\rm tot}^{\rm el} = \frac{32\pi}{k^2} = 22.2\,{\rm fm}^2.$$

(c) The elastic scattering amplitude is given by

$$\begin{aligned} f_{\rm el}(k,\theta) &= \frac{1}{2ik}\sum_{\ell=0}^{\infty}(2\ell+1)\,\left[\eta_\ell e^{2i{\rm Re}(\delta_\ell)}-1\right]P_\ell(\cos\theta) = \frac{i}{2k}\sum_{\ell=0}^{3}(2\ell+1)\,P_\ell(\cos\theta) \\ &= \frac{i}{2k}\left[1+3\cos\theta+\frac{5}{2}(3\cos^2\theta-1)+\frac{7}{2}(5\cos^3\theta-3\cos\theta)\right] \end{aligned}$$

The optical theorem states

$$\sigma_{\rm tot} = \frac{4\pi}{k}{\rm Im}\Big[f_{\rm el}(k,\theta=0)\Big].$$

In this case, the forward elastic scattering amplitude is

$$f(0) = \frac{i}{2k}\left[1+3+\frac{5}{2}(3-1)+\frac{7}{2}(5-3)\right] = \frac{8i}{k}.$$

The previous result could be obtained a bit easier using the relation $P_\ell(1) = 1$. Therefore

$$\frac{4\pi}{k}{\rm Im}\Big[f_{\rm el}(k,\theta=0)\Big] = \frac{32\pi}{k^2} = \sigma_{\rm tot},$$

which shows that the optical theorem is verified.

18. A sample of potassium atoms, $^{19}{\rm K} = [{\rm Ar}]4s^1$, is irradiated with electromagnetic radiation with energy between 2 and 3.3 eV. The ionization energy of potassium is 4.34 eV, the energy of the $4s \to 4p$ transition is $13015\,{\rm cm}^{-1}$, and the quantum defect for $\ell = 2$ is $\delta_2 = 0.14$.

    (a) Using the information provided, obtain the energies of the potassium orbitals and show that their ordering is (from lower to higher energy): $4s$, $4p$, $5s$, $3d$, $5p$, $4d$.

    (b) Indicate which orbitals can be excited by the incident radiation, and determine the energies of the corresponding absorption lines.

    (c) Indicate which transitions occur in the decay to the ground state and determine the energies of the corresponding emission lines (neglect spin–orbit interaction). Draw the transitions in a Grotrian diagram.

    (d) When the sample is subjected to a weak constant magnetic field, it is observed that the absorption line corresponding to the transition $4S_{1/2} \to 5P_{1/2}$ splits

into several components. Knowing that the separation between the highest-energy and lowest-energy lines is $\Delta_{\max} = 0.1\,\mathrm{cm}^{-1}$, calculate the strength of the applied magnetic field.

Assume LS coupling and electric-dipole radiation absorption/emission.

Solution:

(a) For alkali atoms the energies of the terms follow approximately the relation

$$E_{nl} = -\frac{R_H}{(n-\delta_\ell)^2}.$$

This formula allows to compute the energies of s, p and d orbitals provided we know the values of δ_0, δ_1 and δ_2. We are explicitly given only $\delta_2 = 0.14$, so we will have to deduce the values of δ_0 and δ_1 using the remaining information. From the potassium ionization energy we know that $E_{4s} = -4.34\,\mathrm{eV}$, therefore δ_s and E_{5s} can be computed as

$$\delta_0 = 4 - \sqrt{\frac{13.61}{4.34}} = 2.23, \qquad E_{5s} = -\frac{13.6}{(5-2.23)^2} = -1,77\,\mathrm{eV}.$$

To find δ_1 use the $4s \to 4p$ transition energy,

$$E_{4p} - E_{4s} = 13015\,\mathrm{cm}^{-1} = 13015 \times 1.240 \times 10^{-4}\,\mathrm{eV} = 1.61\,\mathrm{eV},$$

to compute $E_{4p} = (-4.34 + 1.61)\,\mathrm{eV} = -2.73\,\mathrm{eV}$, and

$$\delta_1 = 4 - \sqrt{\frac{13.61}{2.73}} = 1.77, \qquad E_{5p} = -\frac{13.6}{(5-1.77)^2} = -1,30\,\mathrm{eV}.$$

To get the energies of the $3d$ and $4d$ orbitals we can use directly the given value of $\delta_2 = 0.14$

$$E_{3d} = -\frac{13.6}{(3-0.14)^2} = -1,66\,\mathrm{eV}, \qquad E_{4d} = -\frac{13.6}{(4-0.14)^2} = -0,91\,\mathrm{eV}.$$

The energy ordering of the orbitals is then

$4s$	$4p$	$5s$	$3d$	$5p$	$4d$.
$-4.34\,\mathrm{eV}$	$-2.73\,\mathrm{eV}$	$-1.77\,\mathrm{eV}$	$-1.66\,\mathrm{eV}$	$-1.30\,\mathrm{eV}$	$-0.91\,\mathrm{eV}$.

(b) The ground-state configuration of potassium is $[\mathrm{Ar}]4s^1$. The LS-coupling selection rules for absorption of electric-dipole radiation require $\Delta S = 0$ and $\Delta L = 1$, therefore the transitions must go from the $4s$ orbital to a np orbital: $4p$, $5p$, $6p$ etc. The energy of the incident radiation is in the range 2 to 3.3 eV, while the energy differences between those orbitals and the ground-state are

$$E_{4p} - E_{4s} = 1.61\,\mathrm{eV}, \qquad E_{5p} - E_{4s} = (-1.30 + 4.34)\,\mathrm{eV} = 3.04\,\mathrm{eV},$$

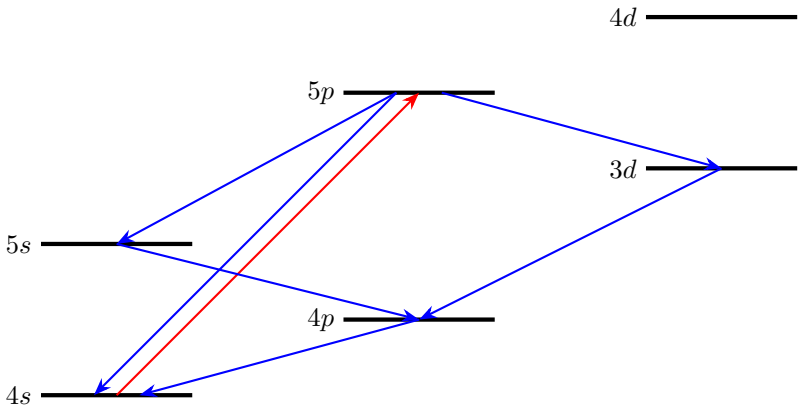
$$E_{6p} - E_{4s} = -\frac{13.6}{\big(6-1.77\big)^2}\,\mathrm{eV} + 4.34\,\mathrm{eV} = 3.58\,\mathrm{eV}.$$

We conclude that only the $5p$ orbital can be excited with the incident radiation.

(c) Once the $5p$ orbital is populated, the following electric-dipole emission transitions, satisfying $\Delta S = 0$ and $\Delta L = 1$, can occur

$5p \to 4s$	$5p \to 5s$	$5p \to 3d$	$5s \to 4p$	$3d \to 4p$	$4p \to 4s$.
$3.04\,\mathrm{eV}$	$0.47\,\mathrm{eV}$	$0.36\,\mathrm{eV}$	$0.96\,\mathrm{eV}$	$1.07\,\mathrm{eV}$	$1.61\,\mathrm{eV}$.

We can draw the transitions in the following Grotrian diagram (not to scale)



(d) The Zeeman energy shifts are given by

$$\Delta E = \mu_{\mathrm{B}} g_{\mathrm{J}} B m_j, \qquad g_{\mathrm{J}} = 1 + \frac{j(j+1) + s(s+1) - \ell(\ell+1)}{2j(j+1)}.$$

For $\mathrm{S}_{1/2}$ we have $g_{\mathrm{J}} = 2$, which leads to $(\Delta E)_{\mathrm{S}_{1/2}} = 2\mu_{\mathrm{B}} B$. For $\mathrm{P}_{1/2}$ we have $g_{\mathrm{J}} = 2/3$ and $(\Delta E)_{\mathrm{P}_{1/2}} = 2\mu_{\mathrm{B}} B/3$. This means that the energy separation between the higher and lower energy lines in the $4\mathrm{S}_{1/2} \to 5\mathrm{P}_{1/2}$ transition is given by

$$\Delta_{\max} = \left(2 + \frac{2}{3}\right) \mu_{\mathrm{B}} B.$$

The value of the applied magnetic field is then

$$B = \frac{3\Delta_{\max}}{8\mu_{\mathrm{B}}} = \frac{3 \times 0.1 \times 1.240 \times 10^{-4}\,\mathrm{eV}}{8 \times 5.788 \times 10^{-5}\mathrm{MeV/T}} = 0.08\,\mathrm{T}.$$

19. Use the variational method with the trial function $\psi(r, \theta, \varphi) = Ae^{-r/\beta}$, where β is the variational parameter, to estimate the binding energy of the ground state of hydrogen. Compare your result with the exact solution obtained by solving the Schrödinger equation.

Solution: First of all, let's normalize the trial function. Using the definite integral $\int_0^\infty r^n e^{-ar}\,\mathrm{d}r = n!/a^{n+1}$, we get

$$1 = \int |\psi|^2 r^2 \sin\theta\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}\varphi = 4\pi|A|^2 \int_0^\infty r^2 e^{-2r/\beta}\,\mathrm{d}r = 4\pi|A|^2 \frac{\beta^3}{4} = \pi|A|^2\beta^3 \implies |A| = \sqrt{\frac{1}{\pi\beta^3}}.$$

The hydrogen energy as a function of the variational parameter is

$$E(\beta) = \langle\psi\,|\,\mathcal{H}\,|\,\psi\rangle = \left\langle \psi \left| \frac{p^2}{2m} - \frac{e^2}{r} \right| \psi \right\rangle = -\left\langle \psi \left| \frac{\hbar^2}{2m}\nabla^2 + \frac{e^2}{r} \right| \psi \right\rangle.$$

Let's compute first the kinetic-energy part. We can take A to be real, so $\psi^* = \psi$ and

$$-\left\langle \psi \left| \frac{\hbar^2}{2m}\nabla^2 \right| \psi \right\rangle = -\frac{\hbar^2}{2m}\int \psi^*\nabla^2\psi\,\mathrm{d}^3\mathbf{r} = -\frac{\hbar^2}{2m}\int \psi\nabla^2\psi\,\mathrm{d}^3\mathbf{r}.$$

Since both ψ and its derivative go to zero at infinity, an integration by parts gives

$$\int \psi\nabla^2\psi\,\mathrm{d}^3\mathbf{r} = -\int (\boldsymbol{\nabla}\psi)(\boldsymbol{\nabla}\psi)\,\mathrm{d}^3\mathbf{r} = -\frac{1}{\beta^2}\int \psi^2\,\mathrm{d}^3\mathbf{r} = -\frac{1}{\beta^2},$$

where he have used the following relations

$$\boldsymbol{\nabla}\psi = \frac{\partial\psi}{\partial r}\hat{\mathbf{r}} = -\frac{1}{\beta}\psi\,\hat{\mathbf{r}}, \qquad (\boldsymbol{\nabla}\psi)(\boldsymbol{\nabla}\psi) = \frac{1}{\beta^2}\psi^2, \qquad \int \psi^2\,\mathrm{d}^3\mathbf{r} = 1.$$

We could also do the integral using the explicit form of the Laplace operator,

$$\nabla^2\psi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) = \frac{1}{\beta}\left(\frac{1}{\beta}-\frac{2}{r}\right)\psi,$$

$$\int \psi\nabla^2\psi\,\mathrm{d}^3\mathbf{r} = \frac{1}{\beta^2}\int\psi^2\,\mathrm{d}^3\mathbf{r} - \frac{2}{\beta}\int\frac{1}{r}\psi^2\,\mathrm{d}^3\mathbf{r}$$
$$= \frac{1}{\beta^2} - \frac{8\pi A^2}{\beta}\int_0^\infty \frac{1}{r}e^{-2r/\beta}r^2\,\mathrm{d}r = \frac{1}{\beta^2} - \frac{8\pi A^2}{\beta}\frac{\beta^2}{4} = -\frac{1}{\beta^2}.$$

Therefore

$$-\left\langle\psi\,\middle|\,\frac{\hbar^2}{2m}\nabla^2\,\middle|\,\psi\right\rangle = \frac{\hbar^2}{2m}\frac{1}{\beta^2}.$$

The Coulomb-energy term is

$$-\left\langle\psi\,\middle|\,\frac{e^2}{r}\,\middle|\,\psi\right\rangle = -4\pi e^2A^2\int_0^\infty r\,e^{-2r/\beta}\,\mathrm{d}r = -\pi e^2A^2\beta^2 = -\frac{e^2}{\beta}.$$

Putting everything together,

$$E(\beta) = \frac{\hbar^2}{2m\beta^2} - \frac{e^2}{\beta}.$$

Minimizing this relation with respect to β gives

$$\frac{\mathrm{d}E}{\mathrm{d}\beta} = -\frac{\hbar^2}{m\beta^3} + \frac{e^2}{\beta^2} = 0 \implies \beta = \frac{\hbar^2}{me^2}.$$

Calling $\beta_0 = \hbar^2/(me^2)$,

$$E(\beta_0) = \frac{\hbar^2}{2m}\frac{m^2e^4}{\hbar^4} - \frac{me^4}{\hbar^2} = -\frac{me^4}{2\hbar^2}.$$

This happens to be equal to the exact result obtained solving the Schrödinger equation because we have started with a trial function that is identical to the hydrogen ground-state wavefunction. The value of the variational parameter that minimizes the energy is also identical to the Bohr radius, $\beta_0 = a_0$.

$$E = -\frac{e^2}{2a_0} = -\frac{me^4}{2\hbar^2}, \qquad a_0 = \frac{\hbar^2}{me^2}.$$

20. Calculate the scattering amplitude, in the first Born approximation, for a potential well of the form $V(r) = -V_0$ for $0 < r < r_0$ and $V(r) = 0$ for $r \geq r_0$. Show that, if the radius of the well, r_0, is much smaller than the wavelength associated with the incident particle, the scattering amplitude is isotropic (i.e., it does not depend on the angle).

Solution: For a spherically symmetric potential, the scattering amplitude in the first Born approximation is

$$f(\theta) = -\frac{2m}{\hbar^2\kappa}\int_0^\infty rV(r)\sin(\kappa r)\,\mathrm{d}r,$$

where $\boldsymbol{\kappa} = \mathbf{k}' - \mathbf{k}$ and $\kappa = |\boldsymbol{\kappa}|$. Therefore, in this case

$$f(\theta) = \frac{2m}{\hbar^2\kappa}\int_0^{r_0} rV_0\sin(\kappa r)\,\mathrm{d}r = \frac{2mV_0}{\hbar^2\kappa}\left(-\left[\frac{r\cos(\kappa r)}{\kappa}\right]_0^{r_0} + \frac{1}{\kappa}\int_0^{r_0}\cos(\kappa r)\,\mathrm{d}r\right)$$
$$= \frac{2mV_0}{\hbar^2\kappa}\left(-\frac{r_0\cos(\kappa r_0)}{\kappa} + \frac{\sin(\kappa r_0)}{\kappa^2}\right) = \frac{2mV_0}{\hbar^2\kappa^3}\Big(\sin(\kappa r_0) - \kappa r_0\cos(\kappa r_0)\Big).$$

Large wavelength compared to r_0 means that $kr_0 \ll 1$ and $\kappa r_0 \ll 1$. In this case we may approximate the sine and cosine functions by their first two power-expansion terms. A quick

way to find the expansions is to remember that $e^{ix} = \sum (ix)^n/n!$ and $e^{ix} = \cos x + i \sin x$, so $\sin(\kappa r_0) \approx \kappa r_0 - \frac{1}{6}(\kappa r_0)^3$ and $\cos(\kappa r_0) \approx 1 - \frac{1}{2}(\kappa r_0)^2$. Therefore, for $\kappa r_0 \ll 1$

$$f(\theta) \approx \frac{2mV_0}{\hbar^2\kappa^3}\left(\kappa r_0 - \frac{1}{6}(\kappa r_0)^3 - \kappa r_0 + \frac{1}{2}(\kappa r_0)^3\right) = \frac{2mV_0 r_0^3}{3\hbar^2}.$$

This shows that that $f(\theta)$ does not have an angular dependence.

Alternatively, one could take $e^{i\boldsymbol{\kappa}\cdot\mathbf{r}} \approx 1$ for $r < r_0$, which is valid in the low-energy limit, in the general first Born approximation expression of the scattering amplitude,

$$f(\theta,\varphi) = -\frac{m}{2\pi\hbar^2}\int e^{i\boldsymbol{\kappa}\cdot\mathbf{r}}\, V(\mathbf{r})\, \mathrm{d}^3\mathbf{r} \approx -\frac{m}{2\pi\hbar^2}\int V(\mathbf{r})\, \mathrm{d}^3\mathbf{r} = \frac{mV_0}{2\pi\hbar^2} 4\pi \int_0^{r_0} r^2\, \mathrm{d}r = \frac{2mV_0 r_0^3}{3\hbar^2}.$$

Note that if you do the integral in the first version of the solution as

$$f(\theta) = \frac{2mV_0}{\hbar^2\kappa}\int_0^{r_0} r\sin(\kappa r)\, \mathrm{d}r = \frac{2mV_0}{\hbar^2\kappa}\frac{1}{2i}\int_0^{r_0} r\left(e^{i\kappa r} - e^{-i\kappa r}\right)\, \mathrm{d}r,$$

you need the first-order term in the approximation of the exponentials to get a non-zero result for the scattering amplitude, $e^{\pm ikr} \approx 1 \pm \frac{i}{2}\kappa r$.

Appendix B

Appendix.

The values of the fundamental constants are taken from [14]. A few of them are quoted with the available experimental accuracy and the uncertainty is indicated by the digits in brackets, but no attempt is made to properly propagate the uncertainties to other derived quantities.

B.1 Systems of Units.

B.1.1 Natural units.

Fundamental constants:

$$c = 3 \times 10^8 \text{ m} \cdot \text{s}^{-1}, \qquad \hbar = 6.6 \times 10^{-22} \text{ MeV} \cdot \text{s}, \qquad [c] = [\text{L}][\text{T}]^{-1}, \qquad [\hbar] = [\text{M}][\text{L}]^2[\text{T}]^{-1}$$

Combined system of natural units and Heaviside-Lorentz units:

$$\hbar = c = \epsilon_0 = \mu_0 = 1$$

Permeability[1] of free space, μ_0. Permittivity of free space, ϵ_0. Now [M], $[\text{L}]^{-1}$ and $[\text{T}]^{-1}$ have equivalent dimensions. The fine-structure constant is a dimensionless parameter with a value of[2],

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} = \frac{e^2}{4\pi} = \frac{1}{137.035999139(31)}, \qquad e > 0$$

After this units transformation all quantities are expressed in powers of GeV.

$$\begin{aligned} \text{Energy} \quad \text{GeV} &\to \text{GeV} & \qquad \text{Time} \quad (\text{GeV}/\hbar)^{-1} &\to \text{GeV}^{-1} \\ \text{Momentum} \quad \text{GeV}/c &\to \text{GeV} & \qquad \text{Length} \quad (\text{GeV}/\hbar c)^{-1} &\to \text{GeV}^{-1} \\ \text{Mass} \quad \text{GeV}/c^2 &\to \text{GeV} & \qquad \text{Area} \quad (\text{GeV}/\hbar c)^{-2} &\to \text{GeV}^{-2} \end{aligned}$$

A useful relation for conversions is $\hbar c = 197.327 \text{ MeV} \cdot \text{fm}$ ($1 \text{ fm} = 10^{-15} \text{ m}$), therefore in natural units we have,

$$1 \text{ fm} \approx \frac{1}{200 \text{ MeV}} = 5 \text{ GeV}^{-1}.$$

In high-energy particle physics, masses, energies and momenta are typically expressed in MeV or GeV. Cross sections are usually given in millibarns (mb) or microbarns (μb), where $1 \text{ mb} = 10^{-31} \text{ m}^2$.

[1] The values for the permeability and permittivity of free space in SI units are:

$$\mu_0 = 12.57 \times 10^{-7} \text{ N} \cdot \text{A}^{-2}, \qquad \epsilon_0 = \frac{1}{\mu_0 c^2} = 8.85 \times 10^{-12} \text{ F/m}.$$

[2] This is the value at $Q^2 = 0$. At higher energy scales the value increases. For example, at $Q^2 \approx m_W^2$ (squared mass of the W gauge boson) $\alpha \sim 1/128$. This *running coupling constant* behaviour is explained in the context of relativistic quantum field theories.

From the value of $\hbar c$ is easy to get the relation $1\ (\mathrm{GeV})^{-2} = 0.38939$ mb. Typical hadronic cross sections correspond to an area of about λ_π^2, where λ_π is the pion Compton wavelength,

$$\lambda_\pi = \frac{\hbar}{m_\pi c} = \frac{\hbar c}{m_\pi c^2} = \frac{197.327\ \mathrm{MeV}\cdot\mathrm{fm}}{139.57\ \mathrm{MeV}} = 1.41\ \mathrm{fm}, \qquad \lambda_\pi^2 = \frac{1}{m_\pi^2} \approx 20\ \mathrm{mb}.$$

Electromagnetic cross sections are an order of magnitude smaller.

B.1.2 Atomic units.

In the formulas below m stands for electron mass. There are two slightly different systems of atomic units.

Hartree units.

- In this system the fundamental constants $\hbar$, e, m_e and $\dfrac{1}{4\pi\epsilon_0}$ are all taken to be unity by definition.
- Unit of length (Bohr radius), $a_0 = \dfrac{\hbar^2}{(e^2/4\pi\epsilon_0)m} = 5.29\times10^{-11}\mathrm{m}$. Sometimes this is written as a_∞ to indicate that we are taking the nucleus to have *infinite* mass compared to the electron.
- The Hartree unit of energy is $E_\mathrm{H} = \dfrac{me^4}{4(4\pi\epsilon_0\hbar)^2} = \dfrac{\alpha^2 mc^2}{4} = \dfrac{e^2}{4\pi\epsilon_0 a_0}$. This is twice the binding energy in the first Bohr orbit of hydrogen.
- Hamiltonian in Hartree units: $H = -\dfrac{1}{2}\nabla^2 - \dfrac{1}{r}$.
- The unit of time is
$$\frac{a_0}{v_0} = \frac{a_0}{\alpha c} = 2.419\times10^{-17}\,\mathrm{s},$$
which is the time required by an electron in the first Bohr orbit to travel one Bohr radius. In this formula v_0 is the magnitude of the electron velocity in the first Bohr orbit, and is taken to be the unit of velocity in atomic units. The kinetic energy of the electron must be equal to the binding energy, therefore assuming that the electron in the Hydrogen atom is very nearly non relativistic we may write $\frac{1}{2}mv_0^2 = \frac{1}{2}m\alpha^2c^2$ and deduce that $v_0 = \alpha c$.
- The unit of velocity and momentum is $v_0 = \alpha c = 2.188\times10^6\ \mathrm{m}\cdot\mathrm{s}^{-1}$ and $p_0 = mv_0$, respectively.
- The unit of electric field strength is
$$\frac{e}{4\pi\varepsilon_0 a_0^2} = \frac{\alpha\hbar c}{ea_0^2} = 5.142\times10^{11}\ \mathrm{V}\cdot\mathrm{m}^{-1}.$$
This is the strength of the Coulomb field experienced by an electron in the first Bohr orbit of atomic Hydrogen (with infinite nuclear mass).
- The unit of magnetic field strength is $\hbar/(ea_0^2) = 2.351\times10^5$ T.

Rydberg units.

- Rydberg constant, $R_\infty = \dfrac{m}{4\pi\hbar^3 c}\left(\dfrac{e^2}{4\pi\epsilon_0}\right)^2$.

- In Ref. [14] the Rydberg constant is used to define the Rydberg energy as

$$E_{\mathrm{R}} = hcR_\infty = \frac{me^4}{2\hbar^2(4\pi\epsilon_0)^2} = \frac{1}{2}mc^2\alpha^2 = 13.605693009(84)\,\mathrm{eV} = 109737.31568\,\mathrm{cm}^{-1}. \tag{B.1}$$

  However, in Ref. [16] the Rydberg constant is defined in such a way that the energy in Rydberg units is directly $E_R = E/R'_\infty$,

$$R'_\infty = \frac{1}{2}mc^2\left(\frac{e^2}{\hbar c}\right)^2 = \frac{1}{2}mc^2\alpha^2 \tag{B.2}$$

  Note that $E_{\mathrm{R}} = E_{\mathrm{H}}/2$.

- The unit of length is again the Bohr radius, $a_0 = \dfrac{4\pi\epsilon_0\hbar^2}{e^2m}$.

- Hamiltonian in Rydberg units: $H = -\nabla^2 - \dfrac{2}{r}$.

- Energy differences between atomic states are usually reported by quoting the inverse wavelength of the electromagnetic radiation emitted in transitions between the states. From the relation between energy and wavelength, $1/\lambda = E/hc$, we have, for $E = 1\,\mathrm{eV}$.

$$\frac{1}{\lambda} = \frac{E}{2\pi\hbar c} = \frac{1\,\mathrm{eV}}{2\pi \times 197.3269788(12) \times 10^6\,\mathrm{eV} \times 10^{-13}\,\mathrm{cm}} = 8065.544005\,\frac{\mathrm{eV}}{\mathrm{cm}}. \tag{B.3}$$

  For 1 Rydberg we have,

$$1\,\mathrm{Ry} \sim 109737.3157\ \mathrm{cm}^{-1}. \tag{B.4}$$

  The equivalence for $1\,\mathrm{cm}^{-1}$ is

$$1\,\mathrm{cm}^{-1} \sim 1.239841974 \times 10^{-4}\,\mathrm{eV} \ \sim\ 9.112670506 \times 10^{-6}\,\mathrm{Ry}. \tag{B.5}$$

$$1\,\mathrm{eV} \sim 8065.544005\ \mathrm{cm}^{-1}.$$

  The spectroscopic usage of cm^{-1} units to report energy differences is just a convention, in the sense that authors choose to report the wavelength of the photon emitted instead of its energy, because it is wavelengths (and not energy) what is really measured by the spectroscopist. This should not be confused with the use of cm^{-1} to measure energy in the natural system of units. In this system the numerical values of $\hbar$ and c are both unity,

$$1 = \hbar c = 197.33\ \mathrm{MeV}\cdot\mathrm{fm} \quad\Longrightarrow\quad 1\,\mathrm{cm}^{-1} = 1.9733 \times 10^{-5}\ \mathrm{eV}. \tag{B.6}$$

  Therefore, 1 cm^{-1} in **natural units** does not correspond to the same amount of energy as 1 cm^{-1} in the spectroscopic convention.

B.1.3 Gaussian units.

We follow Weissbluth in the choice of Gaussian units for most electromagnetic relations in this text. For problems in mechanics, the differences between SI and Gaussian units are just trivial numerical factors. The situation is more difficult when dealing with electromagnetism, where charge, for example, actually has different dimensions for the two sets of units, as it will be shown later in this section. In the following list of formulas, we write on the left the SI version and on the right the Gaussian version. For detailed discussion you may see the appendix on units and dimensions of [10].

SI	**Gaussian**

$$\mathbf{F} = \frac{1}{4\pi\epsilon_0}\frac{q_1 q_2}{r^2}\hat{\mathbf{r}}, \qquad \mathbf{F} = \frac{q_1 q_2}{r^2}\hat{\mathbf{r}}. \tag{B.7}$$

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v}\times\mathbf{B}), \qquad \mathbf{F} = q\left(\mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right). \tag{B.8}$$

$$U = \frac{1}{2}\int\left(\epsilon_0\mathbf{E}^2 + \frac{1}{\mu_0}\mathbf{B}^2\right)\,d\tau, \qquad U = \frac{1}{8\pi}\int\left(\mathbf{E}^2 + \mathbf{B}^2\right)\,d\tau. \tag{B.9}$$

In the SI system the permittivity (ϵ_0) and permeability (μ_0) of vacuum are related through $\epsilon_0\mu_0 = 1/c^2$. For comparison, Maxwell's equations written in both the SI and the Gaussian system of units are shown in the following table

SI	Gaussian
$\boldsymbol{\nabla}\mathbf{E} = \rho/\epsilon_0$	$\boldsymbol{\nabla}\mathbf{E} = 4\pi\rho$
$\boldsymbol{\nabla}\times\mathbf{E} = -\partial\mathbf{B}/\partial t$	$\boldsymbol{\nabla}\times\mathbf{E} = -\frac{1}{c}\partial\mathbf{B}/\partial t$
$\boldsymbol{\nabla}\mathbf{B} = 0$	$\boldsymbol{\nabla}\mathbf{B} = 0$
$\boldsymbol{\nabla}\times\mathbf{B} = \mu_0\mathbf{J} + \mu_0\epsilon_0\partial\mathbf{E}/\partial t$	$\boldsymbol{\nabla}\times\mathbf{B} = \frac{4\pi}{c}\mathbf{J} + \frac{1}{c}\partial\mathbf{E}/\partial t$

The energy stored in the electromagnetic field is given by

$$\text{SI:}\quad E = \frac{1}{2}\int(\varepsilon_0 E^2 + \frac{1}{\mu_0}B^2)\,\mathrm{d}\tau, \qquad \text{Gaussian:}\quad E = \frac{1}{8\pi}\int(E^2 + B^2)\,\mathrm{d}\tau. \tag{B.10}$$

The electric (scalar) potential, ϕ, and the magnetic (vector) potential, $\mathbf{A}$, are combined to form the electromagnetic four-potential A^μ. The next table displays the relation between $\mathbf{E}$, $\mathbf{B}$ and A^μ in the SI and Gaussian systems,

SI	Gaussian
$A^\mu = (\phi/c, \mathbf{A})$	$A^\mu = (\phi, \mathbf{A})$
$\mathbf{E} = -\boldsymbol{\nabla}\phi - \frac{\partial\mathbf{A}}{\partial t}$	$\mathbf{E} = -\boldsymbol{\nabla}\phi - \frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}$
$\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$	$\mathbf{B} = \boldsymbol{\nabla}\times\mathbf{A}$

Although in the previous tables we have used the same symbols (q, $\mathbf{E}$, $\mathbf{B}$, $\mathbf{A}$, etc.) for the corresponding physical quantities in both the Gaussian and the SI systems of units, one must be careful to note that conversions between quantities in Gaussian and SI units are not direct numerical conversions, because the quantities themselves are defined differently in each system and may have different dimensions. For example, in the Gaussian system the electric and magnetic fields have the same dimensions and could be both measured in Gauss (but this is not usually done in practice), whereas this is not the case in the SI system.

We can write in more detail Coulomb's law in both systems of units as

$$\mathbf{F}_{\mathrm{si}} = \frac{1}{4\pi\epsilon_0}\frac{(q_1)_{\mathrm{si}}(q_2)_{\mathrm{si}}}{r^2}\,\hat{\mathbf{r}}, \qquad \mathbf{F}_{\mathrm{g}} = \frac{(q_1)_{\mathrm{g}}(q_2)_{\mathrm{g}}}{r^2}\,\hat{\mathbf{r}}. \tag{B.11}$$

This tells us that the electric charge density, and electric charge, in SI and Gaussian units are related through

$$\rho_{\mathrm{g}} = \frac{\rho_{\mathrm{si}}}{\sqrt{4\pi\epsilon_0}}, \qquad e_{\mathrm{g}} = \frac{e_{\mathrm{si}}}{\sqrt{4\pi\epsilon_0}}. \tag{B.12}$$

The unit of charge in the Gaussian system is the *statcoulomb* (statC).[3] From Coulomb's law (B.11) for $\mathbf{F}_{\mathrm{g}}$ we see that the base dimensions of a statC are purely mechanical

$$[\text{statC}] \;=\; (\text{mass})^{1/2}\cdot(\text{length})^{3/2}\cdot(\text{time})^{-1} \;=\; \mathrm{M}^{1/2}\cdot\mathrm{L}^{3/2}\cdot\mathrm{T}^{-1}. \tag{B.13}$$

[3] Other denominations are the Franklin (Fr) or the electrostatic unit of charge (esu).

However, the SI system takes the unit of electric current as a *base unit*, the Ampere (A), defined as *exactly* $1/(1.602176634 \times 10^{-19})$ times the elementary charge per second (the elementary charge is that of the electron). This means that in the SI system the unit of charge has base dimensions of

$$[\mathrm{C}] \;=\; (\text{current}) \cdot (\text{time}) \;=\; \mathrm{I} \cdot \mathrm{T}. \tag{B.14}$$

By quantifying the ampere in terms of the coulomb, the SI definition directly ties the ampere to an exactly fixed constant of nature (the elementary electric charge) and to the second, which is also exactly defined (in terms of the frequency of energy jumps in the cesium atom). The numerical relation between statC and C is

$$\frac{q_{\mathrm{g}}}{q_{\mathrm{si}}} \;=\; \frac{2997924580\ \mathrm{statC}}{1\ \mathrm{C}} \;\approx\; \frac{3\times 10^{9}\ \mathrm{statC}}{1\ \mathrm{C}} \tag{B.15}$$

The relations between the magnetic flux densities, the electric field, the vector potential and the electric potential are

$$\sqrt{\frac{\mu_0}{4\pi}}\mathbf{B}_{\mathrm{g}} = \mathbf{B}_{\mathrm{si}}, \qquad \sqrt{\frac{1}{4\pi\epsilon_0}}\mathbf{E}_{\mathrm{g}} = \mathbf{E}_{\mathrm{si}}, \qquad \sqrt{\frac{\mu_0}{4\pi}}\mathbf{A}_{\mathrm{g}} = \mathbf{A}_{\mathrm{si}}, \qquad \sqrt{\frac{1}{4\pi\epsilon_0}}\phi_{\mathrm{g}} = \phi_{\mathrm{si}}. \tag{B.16}$$

With all previous relations, it should be easy to convert any expression written in Gaussian units to its equivalent in SI units. For example, in Gaussian units we have

$$\mathbf{E}_{\mathrm{g}} = -\frac{1}{c}\frac{\partial \mathbf{A}_{\mathrm{g}}}{\partial t} - \boldsymbol{\nabla}\phi_{\mathrm{g}}, \tag{B.17}$$

which translates to

$$\sqrt{4\pi\epsilon_0}\,\mathbf{E}_{\mathrm{si}} = -\frac{1}{c}\sqrt{\frac{4\pi}{\mu_0}}\,\frac{\partial \mathbf{A}_{\mathrm{si}}}{\partial t} - \sqrt{4\pi\epsilon_0}\,\boldsymbol{\nabla}\phi_{\mathrm{si}}. \tag{B.18}$$

Using the SI relation $c = 1/\sqrt{\epsilon_0\mu_0}$, this equation reads

$$\mathbf{E}_{\mathrm{si}} = -\frac{\partial \mathbf{A}_{\mathrm{si}}}{\partial t} - \boldsymbol{\nabla}\phi_{\mathrm{si}}. \tag{B.19}$$

Let's study in more detail the conversion between measurements of magnetic field expressed in the Gauss and SI systems of units. From the first relation in (B.16) we see that the ratio $B_{\mathrm{g}}/B_{\mathrm{si}}$ *is not* a dimensionless parameter,

$$\frac{B_{\mathrm{g}}}{B_{\mathrm{si}}} = \sqrt{\frac{4\pi}{\mu_0}}. \tag{B.20}$$

It is useful to do a dimensional analysis of this equation in terms of base units of mass (M), length (L), time (T), and current (I). First, note that[4]

$$\begin{gathered}[\mathrm{G}] = \mathrm{M}^{1/2}\cdot\mathrm{L}^{-1/2}\cdot\mathrm{T}^{-1}, \qquad [\mathrm{Tes}] = \mathrm{M}\cdot\mathrm{T}^{-2}\cdot\mathrm{I}^{-1}, \qquad [\mu_0] = \mathrm{M}\cdot\mathrm{L}\cdot\mathrm{T}^{-2}\cdot\mathrm{I}^{-2}, \\ [\epsilon_0] = \mathrm{M}^{-1}\cdot\mathrm{L}^{-3}\cdot\mathrm{T}^{4}\cdot\mathrm{I}^{2}, \qquad [\mathrm{H}] = \mathrm{M}\cdot\mathrm{L}^{2}\cdot\mathrm{T}^{-2}\cdot\mathrm{I}^{-2}.\end{gathered} \tag{B.21}$$

Therefore

$$\left[\frac{\mathrm{G}}{\mathrm{Tes}}\right] = \frac{\mathrm{M}^{1/2}\cdot\mathrm{L}^{-1/2}\cdot\mathrm{T}^{-1}}{\mathrm{M}\cdot\mathrm{T}^{-2}\cdot\mathrm{I}^{-1}} = \mathrm{M}^{-1/2}\cdot\mathrm{L}^{-1/2}\cdot\mathrm{T}\cdot\mathrm{I} = \left[\sqrt{\frac{4\pi}{\mu_0}}\right]. \tag{B.22}$$

So both sides of Eq. (B.20) have the same dimensions. Let's now insert values in the expression $\sqrt{4\pi/\mu_0}\,B_{\mathrm{si}}$ to obtain the numerical conversion rule between Gauss and Tesla. We must take into account that the base units of mass and length in the Gaussian system (gr and cm), are related to those in the SI system (Kg and m) through

$$\mathrm{M}_{\mathrm{si}} = 10^{3}\,\mathrm{M}_{\mathrm{g}}, \qquad \mathrm{L}_{\mathrm{si}} = 10^{2}\,\mathrm{L}_{\mathrm{g}}. \tag{B.23}$$

[4] To avoid confusion with the base dimension of *time* in the study below, in this section we abbreviate the unit Tesla as Tes, instead of the most common use of T.

The value of permeability of free space in SI units is $\mu_0 = 4\pi \times 10^{-7}\,\mathrm{H\,m^{-1}}$, where H is a Henry, with dimensions $[\mathrm{H}] = \mathrm{M}\cdot\mathrm{L}^2\cdot\mathrm{T}^{-2}\cdot\mathrm{I}^{-2}$. Therefore

$$\left[\sqrt{\frac{4\pi}{\mu_0}}B_{\mathrm{si}}\right] = \sqrt{\frac{4\pi}{4\pi\times10^{-7}\,[\mathrm{H\,m^{-1}}]}}\,[\mathrm{Tes}] = 10^{7/2}\left[\frac{\mathrm{Tes}}{\sqrt{\mathrm{H\,m^{-1}}}}\right] = 10^{7/2}\frac{\mathrm{M_{si}}\cdot\mathrm{T_{si}^{-2}}\cdot\mathrm{I_{si}^{-1}}}{(\mathrm{M_{si}}\cdot\mathrm{L_{si}}\cdot\mathrm{T_{si}^{-2}}\cdot\mathrm{I_{si}^{-2}})^{1/2}}$$
$$= 10^{7/2}\,\mathrm{M_{si}^{1/2}}\cdot\mathrm{L_{si}^{-1/2}}\cdot\mathrm{T_{si}^{-1}} = 10^{7/2}\,(10^3\mathrm{M_g})^{1/2}\cdot(10^2\mathrm{L_{si}})^{-1/2}\cdot\mathrm{T_g^{-1}}$$
$$= 10^4\,\mathrm{M_g^{1/2}}\cdot\mathrm{L_g^{-1/2}}\cdot\mathrm{T_g^{-1}} = 10^4\,[\mathrm{G}]. \quad \text{(B.24)}$$

The conversion rule between the quantity B_{g}, expressed in Gauss, and the quantity B_{g}, expressed in Tesla, is then

$$B_{\mathrm{g}} = \sqrt{\frac{4\pi}{\mu_0}}B_{\mathrm{si}} = \frac{10^4\,\mathrm{G}}{1\,\mathrm{Tes}}\,B_{\mathrm{si}}. \quad \text{(B.25)}$$

Consider now the Bohr magneton written in both systems of units as

$$(\mu_{\mathrm{B}})_{\mathrm{g}} = \frac{e_{\mathrm{g}}\hbar}{2mc}, \qquad (\mu_{\mathrm{B}})_{\mathrm{si}} = \frac{e_{\mathrm{si}}\hbar}{2m}. \quad \text{(B.26)}$$

To get the numerical value of $(\mu_{\mathrm{B}})_{\mathrm{g}}$ we should express e_{g} in statC. It is useful to remember that the fine structure constant in Gaussian units is $\alpha = e_{\mathrm{g}}^2/\hbar c$, so

$$e_{\mathrm{g}} = \sqrt{\alpha\hbar c} = \sqrt{\frac{1}{137}\times(1.055\times10^{-27}\,\mathrm{erg\cdot s})\times(2.9979\times10^{10}\,\mathrm{cm\cdot s^{-1}})}$$
$$= 4.80\times10^{-10}\,\mathrm{statC}. \quad \text{(B.27)}$$

Therefore

$$(\mu_{\mathrm{B}})_{\mathrm{g}} = \frac{e_{\mathrm{g}}\hbar}{2mc} = \frac{(4.8\times10^{-10}\,\mathrm{statC})(1.055\times10^{-27}\,\mathrm{erg\cdot s})}{2\times(9.109\times10^{-28}\,\mathrm{gr})(2.9979\times10^{10}\,\mathrm{cm\cdot s^{-1}})}$$
$$= 9.27\times10^{-21}\,\mathrm{erg\cdot G^{-1}}. \quad \text{(B.28)}$$

The value of $(\mu_{\mathrm{B}})_{\mathrm{si}}$ is

$$(\mu_{\mathrm{B}})_{\mathrm{si}} = \frac{e_{\mathrm{si}}\hbar}{2m} = \frac{(1.6\times10^{-19}\,\mathrm{C})(1.055\times10^{-34}\,\mathrm{J\cdot s})}{2\times(9.109\times10^{-31}\,\mathrm{Kg})} = 9.27\times10^{-24}\,\mathrm{J\cdot Tes^{-1}}. \quad \text{(B.29)}$$

Using the conversion rule $e_{\mathrm{g}} = e_{\mathrm{si}}/\sqrt{4\pi\epsilon_0}$, we see that relation between $(\mu_{\mathrm{B}})_{\mathrm{g}}$ and $(\mu_{\mathrm{B}})_{\mathrm{si}}$ can be expressed as

$$(\mu_{\mathrm{B}})_{\mathrm{g}} = \frac{e_{\mathrm{g}}\hbar}{2mc} = \frac{e_{\mathrm{si}}}{\sqrt{4\pi\epsilon_0}}\,\frac{\hbar}{2m\frac{1}{\sqrt{\mu_0\epsilon_0}}} = \sqrt{\frac{\mu_0}{4\pi}}\frac{e_{\mathrm{si}}\hbar}{2m} = \sqrt{\frac{\mu_0}{4\pi}}(\mu_{\mathrm{B}})_{\mathrm{si}}. \quad \text{(B.30)}$$

We can repeat the exercise of inserting numerical values in $\sqrt{\mu_0/4\pi}\,(\mu_{\mathrm{B}})_{\mathrm{si}}$ and carefully check the base dimensions

$$\left[\sqrt{\frac{\mu_0}{4\pi}}\,(\mu_{\mathrm{B}})_{\mathrm{si}}\right]$$
$$= \sqrt{\frac{4\pi\times10^{-7}\,\mathrm{H\,[m^{-1}]}}{4\pi}}\times9.27\times10^{-24}\,[\mathrm{J\cdot Tes^{-1}}] = 10^{-7/2}\times9.27\times10^{-24}\left[\frac{\mathrm{J}\cdot\sqrt{\mathrm{H\,m^{-1}}}}{\mathrm{Tes}}\right]$$
$$= 10^{7/2}\times9.27\times10^{-24}\,\frac{\mathrm{M_{si}}\cdot\mathrm{L_{si}^2}\cdot\mathrm{T_{si}^{-2}}\cdot(\mathrm{M_{si}}\cdot\mathrm{L_{si}}\cdot\mathrm{T_{si}^{-2}k}\cdot\mathrm{I_{si}^{-2}})^{1/2}}{\mathrm{M_{si}}\cdot\mathrm{T_{si}^{-2}}\cdot\mathrm{I_{si}^{-1}}}$$
$$= 10^{-7/2}\times9.27\times10^{-24}\,\mathrm{M_{si}^{1/2}}\cdot\mathrm{L_{si}^{5/2}}\cdot\mathrm{T_{si}^{-1}}$$
$$= 10^{-7/2}\times9.27\times10^{-24}\,(10^3\mathrm{M_g})^{1/2}\cdot(10^2\mathrm{L_{si}})^{5/2}\cdot\mathrm{T_g^{-1}} = 9.27\times10^{-21}\,\mathrm{M_g^{1/2}}\cdot\mathrm{L_g^{5/2}}\cdot\mathrm{T_g^{-1}}$$
$$= 9.27\times10^{-21}\,\left[\mathrm{erg\cdot G^{-1}}\right]. \quad \text{(B.31)}$$

B.2 Time-independent perturbation theory.

Consider an unperturbed Hamiltonian, $\mathcal{H}_0$, with a set of orthonormal state vectors, ϕ_a, and corresponding eigenvalues E_a,

$$\mathcal{H}_0\phi_a = E_a\phi_a, \qquad \langle\phi_a\,|\,\phi_a\rangle = \delta_{ab}. \tag{B.32}$$

Now we add a small term $\delta\mathcal{H}$ to the Hamiltonian, proportional to some small parameter ε. The energy values will then change to $E_a + \delta E_a$, with corresponding state vectors $\phi_a + \delta_a\phi_a$, where δE_a and $\delta\phi_a$ are presumably given by power series in ε,

$$\delta E_a = \delta_1 E_a + \delta_2 E_a + \ldots, \qquad \delta\phi_a = \delta_1\phi_a + \delta_2\phi_a + \ldots, \tag{B.33}$$

with $\delta_n E_a$ and $\delta_n\phi_a$ proportional to ε^n. The Schrödinger equation can be written as

$$\Big(\mathcal{H}_0 + \delta\mathcal{H}\Big)\Big(\phi_a + \delta\phi_a\Big) = \Big(E_a + \delta E_a\Big)\Big(\phi_a + \delta\phi_a\Big). \tag{B.34}$$

$$\mathcal{H}_0\phi_a + \mathcal{H}_0\delta\phi_a + \delta\mathcal{H}\phi_a + \delta\mathcal{H}\delta\phi_a = E_a\phi_a + E_a\delta\phi_a + \delta E_a\phi_a + \delta E_a\delta\phi_a.$$

$$\mathcal{H}_0\delta\phi_a + \delta\mathcal{H}\phi_a + \delta\mathcal{H}\delta\phi_a = E_a\delta\phi_a + \delta E_a\phi_a + \delta E_a\delta\phi_a. \tag{B.35}$$

B.2.1 First order.

To collect the terms of first order in ε we should drop $\delta\mathcal{H}\delta\phi_a$ and $\delta E_a\delta\phi_a$, whose power series start with terms of order ε^2. To first order in ε we then have

$$\mathcal{H}_0\delta_1\phi_a + \delta\mathcal{H}\phi_a = E_a\delta_1\phi_a + \delta_1 E_a\phi_a. \tag{B.36}$$

To find $\delta_1 E_a$ we take the scalar product of the previous equation with ϕ_a,

$$\langle\phi_a\,|\,\mathcal{H}_0\delta_1\phi_a\rangle + \langle\phi_a\,|\,\delta\mathcal{H}\phi_a\rangle = E_a\,\langle\phi_a\,|\,\delta_1\phi_a\rangle + \delta_1 E_a\,\langle\phi_a\,|\,\phi_a\rangle\,.$$

Since $\mathcal{H}$ is Hermitian, we have $\langle\phi_a\,|\,\mathcal{H}_0\delta_1\phi_a\rangle = \langle\mathcal{H}_0\phi_a\,|\,\delta_1\phi_a\rangle = E_a\,\langle\phi_a\,|\,\delta_1\phi_a\rangle$, which is cancelled by the same term in the right hand side of the previous equation. So we finally arrive to the result that, to first order, the shift in the energy of a bound state is the expectation value of the perturbation $\delta\mathcal{H}$ in the unperturbed state,

$$\delta_1 E_a = \langle\phi_a\,|\,\delta\mathcal{H}\,|\,\phi_a\rangle\,. \tag{B.37}$$

However, this procedure is only applicable to non-degenerate spectra. To see why, let us calculate the change in the state vector produced by the perturbation. To do that we take the scalar product of equation (B.36) with a general unperturbed energy eigenvector ϕ_b,

$$\langle\phi_b\,|\,\mathcal{H}_0\delta_1\phi_a\rangle + \langle\phi_b\,|\,\delta\mathcal{H}\phi_a\rangle = E_a\,\langle\phi_b\,|\,\delta_1\phi_a\rangle + \delta_1 E_a\,\langle\phi_b\,|\,\phi_a\rangle\,.$$

$$\langle\phi_b\,|\,\delta\mathcal{H}\,|\,\phi_a\rangle = (E_a - E_b)\,\langle\phi_b\,|\,\delta_1\phi_a\rangle + \delta_1 E_a\delta_{ab} \tag{B.38}$$

For $a = b$, this is the same as equation (B.37), so the new information is that

$$\langle\phi_b\,|\,\delta\mathcal{H}\,|\,\phi_a\rangle = (E_a - E_b)\,\langle\phi_b\,|\,\delta_1\phi_a\rangle\,, \qquad \text{for } a \neq b. \tag{B.39}$$

The problem is now clear. If there are two states $\phi_b \neq \phi_a$ for which $E_b = E_a$, then the previous equation is inconsistent unless $\langle\phi_b\,|\,\delta\mathcal{H}\,|\,\phi_a\rangle$ vanishes, which need not be the case.

Non-degenerate case.

Let's first assume the case of no degeneracy. In other words, the states whose energies and wave functions we want to calculate do not have the same unperturbed energies as each other or any other states. We are then allowed to write

$$\langle \phi_b \,|\, \delta_1 \phi_a \rangle = \frac{\langle \phi_b \,|\, \delta\mathcal{H} \,|\, \phi_a \rangle}{(E_a - E_b)}, \qquad \text{for } \; a \neq b. \tag{B.40}$$

What about the component of $\delta_1\phi_a$ along ϕ_a, given by $\langle \phi_a \,|\, \delta_1 \phi_a \rangle$? To find it we need to impose the condition that $\phi_a + \delta_1\phi_a$ is properly normalized, that is

$$1 = \langle \phi_a + \delta_1\phi_a \,|\, \phi_a + \delta_1\phi_a \rangle = 1 + \langle \phi_a \,|\, \delta_1\phi_a \rangle + \langle \delta_1\phi_a \,|\, \phi_a \rangle + \mathcal{O}(\varepsilon^2). \tag{B.41}$$

So to order ε we have the condition

$$\mathrm{Re}\Big(\langle \phi_a \,|\, \delta_1 \phi_a \rangle \Big) = 0. \tag{B.42}$$

There is no requirement on the imaginary part of $\langle \phi_a \,|\, \delta_1 \phi_a \rangle$, and we are free to choose its value since this just represents a choice of phase of the whole state vector. In particular, we can choose $\langle \phi_a \,|\, \delta_1 \phi_a \rangle$ to be real, in which case the normalization condition becomes

$$\langle \phi_a \,|\, \delta_1 \phi_a \rangle = 0. \tag{B.43}$$

Now we return to equation (B.40) and use the completeness relation of the $\mathcal{H}_0$ eigenstates to write the first order perturbation to ϕ_a as

$$\delta_1\phi_a = \sum_b |\phi_b\rangle \, \langle \phi_b \,|\, \delta_1 \phi_a \rangle = \sum_{b \neq a} \frac{\langle \phi_b \,|\, \delta\mathcal{H} \,|\, \phi_a \rangle}{(E_a - E_b)} \; |\phi_b\rangle \, . \tag{B.44}$$

The first order perturbation to ϕ_a is then

$$\delta_1\phi_a = \sum_{b \neq a} \frac{\langle \phi_b \,|\, \delta\mathcal{H} \,|\, \phi_a \rangle}{(E_a - E_b)} \; |\phi_b\rangle \, . \tag{B.45}$$

Degenerate case.

Suppose there are a number of states $\phi_{a_1}, \phi_{a_2}, \ldots, \phi_{a_n}$, all with the same energy E_a. Since $\delta\mathcal{H}$ is Hermitian, the quantities $\langle \phi_{a_s} \,|\, \delta\mathcal{H} \,|\, \phi_{a_r} \rangle$ form an Hermitian matrix, so according to a general theorem of matrix algebra the vector space on which this matrix acts is spanned by a set of orthonormal eigenvectors $u_1, u_2, \ldots, u_n$ of this matrix, such that

$$\sum_{r=1}^{n} \langle \phi_{a_s} \,|\, \delta\mathcal{H} \,|\, \phi_{a_r} \rangle \; u_{qr} = \Delta_q u_{qs}. \tag{B.46}$$

Here Δ_q are the eigenvalues and u_{qr} are the elements of the matrix whose columns are the components of the orthonormal eigenvectors $u_1, u_2, \ldots, u_n$.

$$\begin{pmatrix} u_{11} & u_{21} & \ldots & u_{n1} \\ u_{12} & u_{22} & \ldots & u_{n2} \\ \ldots & \ldots & \ldots & \ldots \\ u_{1n} & u_{2n} & \ldots & u_{nn} \end{pmatrix}, \qquad \sum_s u^*_{ks} u_{ms} = \delta_{km}.$$

We can define new eigenstates of $\mathcal{H}_0$ with the same energy E_a,

$$\psi_{a_m} = \sum_r u_{mr} \, \phi_{a_r} \tag{B.47}$$

for which, using the relation (B.46) and the orthonormality condition of the u eigenvectors,

$$\langle \psi_{a_k} \,|\, \delta\mathcal{H} \,|\, \psi_{a_m} \rangle = \sum_{rs} u^*_{ks} u_{mr} \langle \phi_{a_s} \,|\, \delta\mathcal{H} \,|\, \phi_{a_r} \rangle = \sum_{s} u^*_{ks} u_{ms} \Delta_m = \delta_{km} \Delta_m. \tag{B.48}$$

We have therefore proved that it is always possible to perform suitable linear combinations of the degenerate states ϕ_{a_r} to find a new set of sates ψ_{a_r} (also degenerate) for which the off-diagonal matrix elements of the perturbation all vanish. The possible inconsistency problem with equation (B.39) can then be avoided if we start with the ψ states instead of ϕ states. Namely, we can apply

$$\langle \psi_b \,|\, \delta\mathcal{H} \,|\, \psi_a \rangle = (E_a - E_b) \langle \psi_b \,|\, \delta_1 \psi_a \rangle \,, \qquad \text{for all } \; a, b, \tag{B.49}$$

because the matrix $\langle \psi_b \,|\, \delta\mathcal{H} \,|\, \psi_a \rangle$ is diagonal, so $\langle \psi_b \,|\, \delta\mathcal{H} \,|\, \psi_a \rangle = 0$ for $a \neq b$ always, including the case when ψ_a and ψ_b are eigenstates with the same eigenvalue $E_a = E_b$. In terms of the new ψ states, the first-order energy shift and the first-order correction to the unperturbed state are of course still given by

$$\delta_1 E_a = \langle \psi_a \,|\, \delta\mathcal{H} \,|\, \psi_a \rangle \,. \tag{B.50}$$

$$\delta_1 \psi_a = \sum_{b \neq a} \frac{\langle \psi_b \,|\, \delta\mathcal{H} \,|\, \psi_a \rangle}{(E_a - E_b)} \, |\psi_b\rangle \,. \tag{B.51}$$

If we are faced with a set of degenerate unperturbed states ϕ_{a_r} for which the matrix $\langle \phi_{a_s} \,|\, \delta\mathcal{H} \,|\, \phi_{a_r} \rangle$ is not diagonal, all we have to do is make it diagonal by finding its eigenvalues and eigenstates, ψ_{a_r}. The first-order energy correction to a given unperturbed state ψ_{a_r} is its corresponding eigenvalue $\langle \psi_{a_r} \,|\, \delta\mathcal{H} \,|\, \psi_{a_r} \rangle$.

When applying perturbation theory to specific problems it is often the case that one can find an observable $\mathcal{A}$ which commutes with both the unperturbed Hamiltonian $\mathcal{H}_0$ and the perturbation $\delta\mathcal{H}$. This simplifies the problem, as it is easy to show that the perturbation has already a diagonal form if we work with simultaneous eigenstates of $\mathcal{H}_0$ and $\mathcal{A}$. The proof if this handy result is the following:

> Let ψ_{a_1} and ψ_{a_2} be degenerate eigenstates of $\mathcal{H}_0$ which are also eigenstates of $\mathcal{A}$ with *different* eigenvalues,
>
> $$\mathcal{A} \,|\psi_{a_1}\rangle = A_{a_1} \,|\psi_{a_1}\rangle \,, \qquad \mathcal{A} \,|\psi_{a_2}\rangle = A_{a_2} \,|\psi_{a_2}\rangle \,, \qquad \text{with } \; A_{a_1} \neq A_{a_2}. \tag{B.52}$$
>
> Since by assumption $[\mathcal{A}, \delta\mathcal{H}] = 0$, and $\mathcal{A}$ is Hermitian, we have
>
> $$\begin{aligned} 0 &= \langle \psi_{a_1} \,|\, [\mathcal{A}, \delta\mathcal{H}] \,|\, \psi_{a_2} \rangle = \langle \psi_{a_1} \,|\, \mathcal{A}\delta\mathcal{H} \,|\, \psi_{a_2} \rangle - \langle \psi_{a_1} \,|\, \delta\mathcal{H}\mathcal{A} \,|\, \psi_{a_2} \rangle \\ &= A_{a_1} \langle \psi_{a_1} \,|\, \delta\mathcal{H} \,|\, \psi_{a_2} \rangle - A_{a_2} \langle \psi_{a_1} \,|\, \delta\mathcal{H} \,|\, \psi_{a_2} \rangle = (A_{a_1} - A_{a_2}) \langle \psi_{a_1} \,|\, \delta\mathcal{H} \,|\, \psi_{a_2} \rangle \,. \end{aligned} \tag{B.53}$$
>
> Therefore $\langle \psi_{a_1} \,|\, \delta\mathcal{H} \,|\, \psi_{a_2} \rangle = 0$ for $\psi_{a_1} \neq \psi_{a_2}$.

Now we compute the first-order perturbation to the state vectors $\delta_1 \phi_a$ when there is degeneracy. In this case case equation (B.39)

$$\langle \phi_b \,|\, \delta\mathcal{H} \,|\, \phi_a \rangle = (E_a - E_b) \langle \phi_b \,|\, \delta_1 \phi_a \rangle \,, \qquad \text{for } \; a \neq b,$$

does not give any information about the components of $\delta_1 \phi_a$ along the unperturbed state vectors ϕ_b for which $E_b = E_a$, and equation (B.40) only applies for $E_a \neq E_b$. Therefore, the sum in equation (B.44) must now be split into two parts: one for states with different energy than E_a and other for states with the same energy as E_a. Thus

$$\delta_1 \phi_a \;=\; \sum_b |\phi_b\rangle \langle \phi_b \,|\, \delta_1 \phi_a \rangle \;=\; \sum_{c:\, E_c \neq E_a} \frac{\langle \phi_c \,|\, \delta\mathcal{H} \,|\, \phi_a \rangle}{(E_a - E_c)} \, |\phi_c\rangle \;+\; \sum_{b:\, E_b = E_a} |\phi_b\rangle \langle \phi_b \,|\, \delta_1 \phi_a \rangle \,. \tag{B.54}$$

We may impose on the perturbed states the condition that they should be orthonormal,

$$\left\langle \phi_b + \delta_1\phi_b + \mathcal{O}(\varepsilon^2) \,\middle|\, \phi_a + \delta_1\phi_a + \mathcal{O}(\varepsilon^2) \right\rangle = \delta_{ab} \qquad \text{for } E_a = E_b. \tag{B.55}$$

$$\underbrace{\langle \phi_b \,|\, \phi_a \rangle}_{\delta_{ab}} + \langle \phi_b \,|\, \delta_1\phi_a \rangle + \langle \delta_1\phi_b \,|\, \phi_a \rangle + \mathcal{O}(\varepsilon^2) = \delta_{ab} \qquad \text{for } E_a = E_b. \tag{B.56}$$

So we have the condition

$$\langle \phi_b \,|\, \delta_1\phi_a \rangle + \langle \delta_1\phi_b \,|\, \phi_a \rangle = 0 \qquad \text{for } E_a = E_b. \tag{B.57}$$

This means that the matrix $\langle \phi_b \,|\, \delta_1\phi_a \rangle = A_{ba}$ is anti-Hermitian for $E_a = E_b$, that is: $A_{ba} = -A^*_{ab}$. Apart from this condition, the matrix A_{ba} is undetermined at this point (for more details see section 5.1 of [15]).

B.2.2 Second order.

We will compute now the change in energies due to a perturbation $\delta\mathcal{H}$, to second order in whatever small parameter ε appears in the perturbed Hamiltonian. Second-order perturbations are of special importance when the first-order perturbation vanishes, as it happens for the Stark shift of atomic energy levels in an electric field for the $1s_{1/2}$, $2p_{3/2}$, etc., states of hydrogen and almost all states of other atoms.

We could compute second-order corrections due only to a perturbation proportional to ε, which can be denoted as $\delta_1\mathcal{H}$. However, with very little extra effort we can be more general and, besides $\delta_1\mathcal{H}$, include also in the Hamiltonian a possible term $\delta_2\mathcal{H}$ which is itself of second order in ε. Thus $\delta\mathcal{H} = \delta_1\mathcal{H} + \delta_2\mathcal{H}$, and

$$\mathcal{H} = \mathcal{H}_0 + \delta_1\mathcal{H} + \delta_2\mathcal{H}. \tag{B.58}$$

Going back to the Schrödinger equation (B.35),

$$\mathcal{H}_0\delta\phi_a + \delta\mathcal{H}\phi_a + \delta\mathcal{H}\delta\phi_a = E_a\delta\phi_a + \delta E_a\phi_a + \delta E_a\delta\phi_a,$$

and equating the terms of second order in ε on both sides, we get

$$\mathcal{H}_0\delta_2\phi_a + \delta_2\mathcal{H}\phi_a + \delta_1\mathcal{H}\delta_1\phi_a = E_a\delta_2\phi_a + \delta_2E_a\phi_a + \delta_1E_a\delta_1\phi_a. \tag{B.59}$$

Non-degenerate case

Let's first consider the non-degenerate case, where none of the states we are interested in have the same unperturbed energies. To find the second-order energy shift we need to take the scalar product of the previous equation with the state ϕ_a,

$$\langle \phi_a \,|\, \mathcal{H}_0 \,|\, \delta_2\phi_a \rangle + \langle \phi_a \,|\, \delta_2\mathcal{H} \,|\, \phi_a \rangle + \langle \phi_a \,|\, \delta_1\mathcal{H} \,|\, \delta_1\phi_a \rangle = E_a \langle \phi_a \,|\, \delta_2\phi_a \rangle + \delta_2E_a + \delta_1E_a \langle \phi_a \,|\, \delta_1\phi_a \rangle .$$

Since $\mathcal{H}_0$ is Hermitian, the term $\langle \phi_a \,|\, \mathcal{H}_0 \,|\, \delta_2\phi_a \rangle = E_a \langle \phi_a \,|\, \delta_2\phi_a \rangle$ cancels with the corresponding term in the right-hand side and we are left with

$$\langle \phi_a \,|\, \delta_2\mathcal{H} \,|\, \phi_a \rangle + \langle \phi_a \,|\, \delta_1\mathcal{H} \,|\, \delta_1\phi_a \rangle = \delta_1E_a \langle \phi_a \,|\, \delta_1\phi_a \rangle + \delta_2E_a. \tag{B.60}$$

We choose the phase normalization of the perturbed state vector so that $\langle \phi_a \,|\, \delta_1\phi_a \rangle = 0$ as before, so this term can be dropped and the last equation becomes

$$\delta_2E_a = \langle \phi_a \,|\, \delta_2\mathcal{H} \,|\, \phi_a \rangle + \langle \phi_a \,|\, \delta_1\mathcal{H} \,|\, \delta_1\phi_a \rangle . \tag{B.61}$$

The second matrix element of the right-hand side can be computed using the first order result

$$\delta_1\phi_a = \sum_{b\neq a} \frac{\langle \phi_b \,|\, \delta_1\mathcal{H} \,|\, \phi_a \rangle}{(E_a - E_b)} \, |\phi_b\rangle , \tag{B.62}$$

to get

$$\langle \phi_a \,|\, \delta_1\mathcal{H} \,|\, \delta_1\phi_a \rangle = \sum_{b\neq a} \frac{\langle \phi_b \,|\, \delta_1\mathcal{H} \,|\, \phi_a \rangle}{(E_a - E_b)} \, \langle \phi_a \,|\, \delta_1\mathcal{H} \,|\, \phi_b \rangle = \sum_{b\neq a} \frac{|\, \langle \phi_b \,|\, \delta_1\mathcal{H} \,|\, \phi_a \rangle \,|^2}{(E_a - E_b)}. \tag{B.63}$$

Therefore the second-order energy shift is

$$\delta_2 E_a = \sum_{b\neq a} \frac{|\, \langle \phi_b \,|\, \delta_1\mathcal{H} \,|\, \phi_a \rangle \,|^2}{(E_a - E_b)} + \langle \phi_a \,|\, \delta_2\mathcal{H} \,|\, \phi_a \rangle \,. \tag{B.64}$$

When one says that an energy shift is produced by the emission and reabsorption of some virtual particle, as for instance the Lamb shift is produced by the emission and reabsorption of a virtual photon by the electron in the hydrogen atom, what is meant is that $\delta_2 E_a$ (or a higher-order correction) receives an important contribution from a state ϕ_b containing that particle. Note that, if ϕ_a is the state of lowest energy of the system, then (in the absence of $\delta_2\mathcal{H}$) the second-order energy shift of its energy is always negative, because all other states have $E_b > E_a$.

To calculate the second-order shift $\delta_2\phi_a$ in the state vectors we need to take the scalar product of (B.59) with ϕ_b for $b \neq a$, and then impose the condition that $\phi_a + \delta_1\phi_a + \delta_2\phi_a + \ldots$ has unit norm to find the value of $\langle \phi_a \,|\, \delta_2\phi_a \rangle$, which cannot be chosen to be zero now. We do not go into this procedure here since the second-order shift in the state vectors will not be used in this course.

Degenerate case

In the degenerate case some of the states which we are interested in have the same unperturbed energies. The calculation of the second-order energy shift follows the same steps as before until we get again equation (B.60),

$$\langle \phi_a \,|\, \delta_2\mathcal{H} \,|\, \phi_a \rangle + \langle \phi_a \,|\, \delta_1\mathcal{H} \,|\, \delta_1\phi_a \rangle = \delta_1 E_a \, \langle \phi_a \,|\, \delta_1\phi_a \rangle + \delta_2 E_a. \tag{B.60}$$

The orthonormality condition (B.41), $\langle \phi_a + \delta_1\phi_a \,|\, \phi_a + \delta_1\phi_a \rangle = 1$, tells us that $\langle \phi_a \,|\, \delta_1\phi_a \rangle$ is imaginary, so it can again be made to vanish by a suitable choice of phase of $\phi_a + \delta_1\phi_a$. We can use Eq. (B.54)

$$\delta_1\phi_a \;=\; \sum_{c:\, E_c\neq E_a} \frac{\langle \phi_c \,|\, \delta_1\mathcal{H} \,|\, \phi_a \rangle}{(E_a - E_c)} \, |\phi_c\rangle \;+\; \sum_{b:\, E_b = E_a} |\phi_b\rangle \, \langle \phi_b \,|\, \delta_1\phi_a \rangle \,, \tag{B.54}$$

for $\delta_1\phi_a$ in the second term $\langle \phi_a \,|\, \delta_1\mathcal{H} \,|\, \delta_1\phi_a \rangle$ on the left of Eq. (B.60). Since the unperturbed states have been chosen so that $\langle \phi_a \,|\, \delta_1\mathcal{H} \,|\, \phi_b \rangle$ vanishes for $E_b = E_a$ but $b \neq a$, and we have chosen the phase of $\phi_a + \delta_1\phi_a$ so that $\langle \phi_a \,|\, \delta_1\phi_a \rangle$ also vanishes, the second term in Eq. (B.54), which involves unknown matrix elements, does not contribute to the second term on the left of Eq. (B.60),

$$\langle \phi_a \,|\, \delta_1\mathcal{H} \,|\, \delta_1\phi_a \rangle = \sum_{c:\, E_c\neq E_a} \frac{\langle \phi_c \,|\, \delta_1\mathcal{H} \,|\, \phi_a \rangle}{(E_a - E_c)} \, \langle \phi_a \,|\, \delta_1\mathcal{H} \,|\, \phi_c \rangle \;+\; \sum_{b:\, E_b = E_a} \langle \phi_b \,|\, \delta_1\phi_a \rangle \underbrace{\langle \phi_a \,|\, \delta_1\mathcal{H} \,|\, \phi_b \rangle}_{=0}.$$

Therefore we arrive at

$$\delta_2 E_a = \sum_{c:\, E_c\neq E_a} \frac{|\, \langle \phi_a \,|\, \delta_1\mathcal{H} \,|\, \phi_c \rangle \,|^2}{(E_a - E_c)} + \langle \phi_a \,|\, \delta_2\mathcal{H} \,|\, \phi_a \rangle \,. \tag{B.65}$$

This is the same result as in the non-degenerate case (B.64), except that here we have to specify not only that the intermediate states ϕ_c have $c \neq a$, but also that they have $E_c \neq E_a$.

B.3 Time-dependent perturbation theory and Fermi's golden rule.

Let $\phi_k(\mathbf{x},t)$ be the normalised solutions to the Schrödinger equation for the unperturbed time–independent Hamiltonian $\mathcal{H}_0$. We have then

$$\mathcal{H}_0\phi_k = E_k\phi_k, \qquad \langle\phi_j \,|\, \phi_k\rangle = \delta_{jk}. \tag{B.66}$$

Let $V(\mathbf{x},t)$ be an interaction Hamiltonian which will be treated as a small perturbation to $\mathcal{H}_0$. We will see that V is going to induce transitions between the sates. The time–dependent Schrödinger equation is

$$i\hbar\frac{\mathrm{d}\psi}{\mathrm{d}t} = [\mathcal{H}_0 + V(\mathbf{x},t)]\,\psi. \tag{B.67}$$

We assume the ϕ_k states to be also a complete set of the perturbed system, so that a generic state $\psi(\mathbf{x},t)$ can be written as

$$\psi(\mathbf{x},t) = \sum_k c_k(t)\phi_k(\mathbf{x})e^{-iE_kt/\hbar}. \tag{B.68}$$

Note that the coefficients $c_k(t)$ have a time–dependence. This will lead to the possibility of transitions between states to occur. Inserting (B.68) into (B.67) we get a set of differential equations for $c_k(t)$,

$$i\hbar\sum_k\left[\frac{\mathrm{d}c_k}{\mathrm{d}t}\phi_k e^{-iE_kt/\hbar} - \frac{i}{\hbar}E_kc_k\phi_k e^{-iE_kt/\hbar}\right] = \sum_k c_k\mathcal{H}_0\phi_k e^{-iE_kt/\hbar} + \sum_k Vc_k\phi_k e^{-iE_kt/\hbar}, \tag{B.69}$$

$$i\hbar\sum_k\frac{\mathrm{d}c_k}{\mathrm{d}t}\phi_k e^{-iE_kt/\hbar} = \sum_k Vc_k(t)\phi_k e^{-iE_kt/\hbar}. \tag{B.70}$$

Take the system to be at $t = 0$ in the initial state $|i\rangle = \phi_i$. The coefficients at $t = 0$ are then $c_k(0) = \delta_{ik}$. We assume the perturbing Hamiltonian to be sufficiently small so that the coefficients satisfy at all times, *as a first approximation*, the relations $c_i(t) \approx 1$ and $c_{k\neq i}(t) \approx 0$. Therefore equation (B.70) can be written as

$$i\hbar\sum_k\frac{\mathrm{d}c_k}{\mathrm{d}t}\phi_k e^{-iE_kt/\hbar} \;\approx\; V\phi_i e^{-iE_it/\hbar}. \tag{B.71}$$

We are interested in finding the transition probability from the initial state $|i\rangle$ to a particular final state $|f\rangle = \phi_f$. This is related to the coefficient $c_f(t)$. Multiplying both sides of (B.71) on the left by ϕ_f^* and integrating we get

$$i\hbar\sum_k\frac{\mathrm{d}c_k}{\mathrm{d}t}e^{-i(E_k-E_f)t/\hbar}\int\phi_f^*\phi_k\,\mathrm{d}^3\mathbf{x} \;=\; e^{-i(E_i-E_f)t/\hbar}\int\phi_f^*V\phi_i\,\mathrm{d}^3\mathbf{x}, \tag{B.72}$$

$$\frac{\mathrm{d}c_f}{\mathrm{d}t} \;=\; -\frac{i}{\hbar}\langle f\,|\,V\,|\,i\rangle\,e^{i(E_f-E_i)t/\hbar} = -\frac{i}{\hbar}T_{fi}e^{i(E_f-E_i)t/\hbar}. \tag{B.73}$$

where the orthogonality relation $\langle f|k\rangle = \delta_{fk}$ has been used. Note that the transition matrix element

$$T_{fi} = \langle f\,|\,V\,|\,i\rangle = \int\phi_f^*V\phi_i\,\mathrm{d}^3\mathbf{x}, \tag{B.74}$$

has dimensions of energy because both ϕ_i and ϕ_f are normalized by a volume integral. According to the differential equation for c_f (B.73), at time $t = T$ the amplitude for transitions to the state $|f\rangle$ is given by

$$c_f(T) = -\frac{i}{\hbar}\int_0^T T_{fi}\,e^{i(E_f-E_i)t/\hbar}\,\mathrm{d}t. \tag{B.75}$$

To first order, the transition matrix element is $T_{fi} = \langle f \,|\, V \,|\, i \rangle$. It was derived assuming the perturbing Hamiltonian to be sufficiently small so that the coefficients satisfy at all times the relations $c_i(t) \approx 1$ and $c_{k\neq i}(t) \approx 0$. The second-order approximation can be obtained by taking again $c_i(t) \approx 1$ and substituting the expression $c_f(T) = -\frac{i}{\hbar}\int_0^T T_{fi}\, e^{i(E_f-E_i)t/\hbar}\,\mathrm{d}t$ for $c_{k\neq i}(t)$ into (B.70).

$$c_k(t) = -\frac{i}{\hbar}\int_0^t T_{ki}\, e^{i(E_k-E_i)t'/\hbar}\,\mathrm{d}t', \qquad i\hbar\sum_k \frac{\mathrm{d}c_k}{\mathrm{d}t}\phi_k e^{-iE_kt/\hbar} = \sum_k V c_k(t)\phi_k e^{-iE_kt/\hbar}.$$

$$i\hbar\sum_k \frac{\mathrm{d}c_k}{\mathrm{d}t}\phi_k e^{-iE_kt/\hbar} = V c_i(t)\phi_i e^{-iE_it/\hbar} - \frac{i}{\hbar}\sum_{k\neq i}\left[\int_0^t T_{ki}\, e^{i(E_k-E_i)t'/\hbar}\,\mathrm{d}t'\right] V\phi_k e^{-iE_kt/\hbar}.$$

Setting now $c_i(t) \approx 1$ and taking the inner product with a particular final state $\phi_f(\mathbf{x})$ we get

$$i\hbar\sum_k \frac{\mathrm{d}c_k}{\mathrm{d}t}\langle\phi_f\,|\,\phi_k\rangle\, e^{-i(E_k-E_f)t/\hbar} = \langle\phi_f\,|\,V\,|\,\phi_i\rangle\, e^{-i(E_i-E_f)t/\hbar} - \frac{i}{\hbar}\sum_{k\neq i}\left[\int_0^t T_{ki}\, e^{i(E_k-E_i)t'/\hbar}\,\mathrm{d}t'\right]\langle\phi_f\,|\,V\,|\,\phi_k\rangle\, e^{-i(E_k-E_f)t/\hbar}.$$

$$\frac{\mathrm{d}c_f}{\mathrm{d}t} = -\frac{i}{\hbar}\langle\phi_f\,|\,V\,|\,\phi_i\rangle\, e^{i(E_f-E_i)t/\hbar} + \left(-\frac{i}{\hbar}\right)^2\sum_{k\neq i}\langle\phi_f\,|\,V\,|\,\phi_k\rangle\, e^{i(E_f-E_k)t/\hbar}\left[\int_0^t \langle\phi_k\,|\,V\,|\,\phi_i\rangle\; e^{i(E_k-E_i)t'/\hbar}\,\mathrm{d}t'\right].$$

Assuming that the perturbation V is not present at $t = 0$ and that for $t > 0$ is constant, the integral in the previous equation can be written as

$$\int_0^t \langle\phi_k\,|\,V\,|\,\phi_i\rangle\; e^{i(E_k-E_i)t'/\hbar}\,\mathrm{d}t' = \langle\phi_k\,|\,V\,|\,\phi_i\rangle\,\frac{e^{i(E_k-E_i)t/\hbar}}{i(E_k-E_i)/\hbar}.$$

The improved differential equation for the $c_f(t)$ coefficient is then

$$\frac{\mathrm{d}c_f}{\mathrm{d}t} = -\frac{i}{\hbar}\left[\langle\phi_f\,|\,V\,|\,\phi_i\rangle + \sum_{k\neq i}\frac{\langle\phi_f\,|\,V\,|\,\phi_k\rangle\,\langle\phi_k\,|\,V\,|\,\phi_i\rangle}{(E_k-E_i)}\right] e^{i(E_f-E_i)t/\hbar}. \tag{B.76}$$

By comparing with equation (B.73) we can see that, to second order, the transition matrix element is given by

$$T_{fi} = \langle\phi_f\,|\,V\,|\,\phi_i\rangle + \sum_{k\neq i}\frac{\langle\phi_f\,|\,V\,|\,\phi_k\rangle\,\langle\phi_k\,|\,V\,|\,\phi_i\rangle}{(E_k-E_i)}. \tag{B.77}$$

It is important to note that the second order term is obtained by summing the amplitudes of the transitions occurring via all possible intermediate states $|\phi_k\rangle$. The full perturbation expansion can be obtained by successive substitutions.

B.3.1 Constant perturbation.

If the perturbation is nearly constant,

$$V(t) = \begin{cases} 0, & \text{for } t < 0, \\ V, \;\; \text{(independent of } t\text{)} & \text{for } t \geq 0, \end{cases} \tag{B.78}$$

the matrix element $T_{fi} = \langle f \,|\, V \,|\, i \rangle$ can be pulled out of the integral. Then

$$c_f(T) = -\frac{i}{\hbar} T_{fi} \int_0^T e^{i(E_f - E_i)t/\hbar} \, \mathrm{d}t. \tag{B.79}$$

The probability for a transition to the state $|f\rangle$ is

$$P_{fi} = c_f^*(T) c_f(T) = \frac{|T_{fi}|^2}{\hbar^2} \int_0^T \int_0^T e^{-i(E_f - E_i)t/\hbar} e^{+i(E_f - E_i)t'/\hbar} \, \mathrm{d}t \, \mathrm{d}t'. \tag{B.80}$$

The double integral has the straightforward result

$$\int_0^T \int_0^T e^{-i(E_f - E_i)t/\hbar} e^{+i(E_f - E_i)t'/\hbar} \, \mathrm{d}t \, \mathrm{d}t' = \frac{4\hbar^2}{(E_f - E_i)^2} \sin^2 \left[\frac{(E_f - E_i)T}{2\hbar} \right].$$

Therefore

$$P_{fi} = \frac{4|T_{fi}|^2}{(E_f - E_i)^2} \sin^2 \left[\frac{(E_f - E_i)T}{2\hbar} \right]. \tag{B.81}$$

A more meaningful quantity is the *transition rate* or probability per unit time Γ_{fi} from the initial state $|i\rangle$ to the single final state $|f\rangle$, defined as $\Gamma_{fi} = \mathrm{d}P_{fi}/\mathrm{d}T$, so

$$\begin{aligned} \Gamma_{fi} = \frac{4|T_{fi}|^2}{(E_f - E_i)^2} \frac{(E_f - E_i)}{2\hbar} 2 \sin \left[\frac{(E_f - E_i)T}{2\hbar} \right] \cos \left[\frac{(E_f - E_i)T}{2\hbar} \right] \\ = \frac{2|T_{fi}|^2}{\hbar(E_f - E_i)} \sin \left[\frac{(E_f - E_i)T}{\hbar} \right]. \end{aligned}$$

We are interested in what happens at large T. Using the properties of the Dirac delta function we have[5]

$$\begin{aligned} \lim_{T \to \infty} \Gamma_{fi} = \frac{2|T_{fi}|^2}{\hbar^2} \lim_{T \to \infty} \frac{\sin \left[(E_f - E_i)T/\hbar \right]}{(E_f - E_i)/\hbar} = \frac{2|T_{fi}|^2}{\hbar^2} \pi \, \delta\Big((E_f - E_i)/\hbar \Big) \\ = \frac{2|T_{fi}|^2}{\hbar^2} \hbar\pi \, \delta(E_f - E_i) = \frac{2\pi|T_{fi}|^2}{\hbar} \delta(E_f - E_i). \end{aligned}$$

This is the Fermi golden rule for the transition rate from an initial state $|i\rangle$ to a final state $|f\rangle$,

$$\Gamma_{fi} = \frac{2\pi}{\hbar} |T_{fi}|^2 \, \delta(E_f - E_i). \tag{B.82}$$

We can also express the Fermi golden rule in terms of the energy density of final states. In fact, to compute the *total transition rate* we need to sum (integrate) over all possible final states. Denoting by $\rho(E_f)$ the energy density of final states, meaning that $\rho(E_f) \, \mathrm{d}E_f$ is the number of final states with energy between E_f and $E_f + \mathrm{d}E_f$, the total transition rate is given by

$$\Gamma_i = \int \Gamma_{fi} \, \rho(E_f) \, \mathrm{d}E_f = \frac{2\pi}{\hbar} \int |T_{fi}|^2 \, \delta(E_f - E_i) \, \rho(E_f) \, \mathrm{d}E_f. \tag{B.83}$$

Therefore

$$\Gamma_i = \frac{2\pi}{\hbar} |T_{fi}|^2 \, \rho(E_i). \tag{B.84}$$

[5] Remember that $\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ixt} \, \mathrm{d}t$, and $\lim_{t \to \infty} \frac{\sin(xt)}{\pi x} = \delta(x)$. Also, $\delta(ax) = \frac{1}{|a|} \delta(x)$.

B.3.2 Monochromatic perturbation.

If, instead of being nearly constant, we have a weak perturbation that oscillates at a single frequency $\omega/2\pi$ and energy $E_\omega = \hbar\omega$, as could be the case when an atom is subjected to radiation of a given wavelength, we can write

$$V(t) = Ue^{-i\omega t} + U^\dagger e^{+i\omega t} = V(t) = Ue^{-iE_\omega t/\hbar} + U^\dagger e^{+iE_\omega t/\hbar}. \tag{B.85}$$

Now we should return to equation (B.75) and perform the integration

$$\begin{aligned}
c_f(T) &= -\frac{i}{\hbar}\int_0^T T_{fi}\, e^{i(E_f-E_i)t/\hbar}\,\mathrm{d}t = -\frac{i}{\hbar}\int_0^T \langle f\,|\,V\,|\,i\rangle\; e^{i(E_f-E_i)t/\hbar}\,\mathrm{d}t \\
&= -\frac{i}{\hbar}\int_0^T \left[\langle f\,|\,U\,|\,i\rangle\; e^{i(E_f-E_i-E_\omega)t/\hbar} + \left\langle f\,\middle|\,U^\dagger\,\middle|\,i\right\rangle\; e^{i(E_f-E_i+E_\omega)t/\hbar}\right]\mathrm{d}t \\
&= -\frac{i}{\hbar}\int_0^T \left[U_{fi}\, e^{i(E_f-E_i-E_\omega)t/\hbar} + U^\dagger_{fi}\, e^{i(E_f-E_i+E_\omega)t/\hbar}\right]\mathrm{d}t \\
&= -\frac{i}{\hbar}\int_0^T \left[U_{fi}\, e^{i(E_f-E_i-E_\omega)t/\hbar} + U^*_{if}\, e^{i(E_f-E_i+E_\omega)t/\hbar}\right]\mathrm{d}t \\
&= -\frac{i}{\hbar}U_{fi}\int_0^T e^{i(E_f-E_i-E_\omega)t/\hbar}\,\mathrm{d}t \;+\; U^*_{if}\int_0^T e^{i(E_f-E_i+E_\omega)t/\hbar}\,\mathrm{d}t.
\end{aligned}$$

Here U_{fi} denotes the matrix element $\langle f\,|\,U\,|\,i\rangle$ and, from the definition of Hermitian conjugate or adjoint, we have

$$U^\dagger_{fi} = \left\langle f\,\middle|\,U^\dagger\,\middle|\,i\right\rangle = \langle Uf\,|\,U\,|\,i\rangle = \langle i\,|\,Uf\rangle^* = \langle i\,|\,U\,|\,f\rangle^* = U^*_{if}.$$

The integrations are again straightforward. Using the relation $e^{ix} - 1 = 2i\, e^{ix/2}\sin(x/2)$, we have

$$\begin{aligned}
\int_0^T e^{i(E_f-E_i\pm E_\omega)t/\hbar}\,\mathrm{d}t &= \left[\frac{e^{i(E_f-E_i\pm E_\omega)t/\hbar}}{i(E_f-E_i\pm E_\omega)/\hbar}\right]_0^T = \frac{e^{i(E_f-E_i\pm E_\omega)T/\hbar}-1}{i(E_f-E_i\pm E_\omega)/\hbar} \\
&= \frac{\sin\left[(E_f-E_i\pm E_\omega)T/2\hbar\right]}{(E_f-E_i\pm E_\omega)/2\hbar}\,\exp\left[\frac{i(E_f-E_i\pm E_\omega)T}{2\hbar}\right] = \frac{\sin\left(\Omega_\pm T\right)}{\Omega_\pm}\,\exp\left(i\Omega_\pm T\right),
\end{aligned}$$

where

$$\Omega_\pm = \frac{E_f - E_i \pm E_\omega}{2\hbar}.$$

Therefore

$$c_f(T) = -\frac{i}{\hbar}\left[U_{fi}\frac{\sin\left(\Omega_- T\right)}{\Omega_-}\,\exp\left(i\Omega_- T\right) + U^*_{if}\frac{\sin\left(\Omega_+ T\right)}{\Omega_+}\,\exp\left(i\Omega_+ T\right)\right].$$

The function $\sin(\Omega T)/\Omega$ has a prominent maximum at $\Omega = 0$ and a series of decreasing secondary maxima located at points $\Omega \neq 0$. At large times it becomes proportional to the Dirac delta distribution,

$$\lim_{T\to\infty}\frac{\sin\left(\Omega T\right)}{\Omega} = \pi\delta(\Omega). \tag{B.86}$$

Thus, at large times the probability for a transition to the state $|f\rangle$ is given by

$$P_{fi} = c^*_f(T)c_f(T) \;\approx\; \frac{1}{\hbar^2}|U_{fi}|^2\frac{\sin^2\left(\Omega_- T\right)}{\Omega_-^2} \;+\; \frac{1}{\hbar^2}|U_{if}|^2\frac{\sin^2\left(\Omega_+ T\right)}{\Omega_+^2}. \tag{B.87}$$

The approximation made here is to neglect the cross terms in the product $c_f^*(T)c_f(T)$, because at large T the function $\sin(\Omega_- T)/\Omega_-$ vanishes everywhere except at points very near to $E_w = E_f - E_i$, while $\sin(\Omega_+ T)/\Omega_+$ vanishes everywhere except at points very near to $E_w = -(E_f - E_i)$, leading to essentially no overlap. The terms in equation (B.87) may describe, for example, the emission (for $E_f < E_i$) or absorption (for $E_f > E_i$) of photons by the system.

Let us assume we are dealing with the absorption of radiation by the system. Then we take the frequency of the perturbation to be close to $(E_f - E_i)/\hbar$. In other words, $E_\omega \approx E_f - E_i$ and only the first term of equation (B.87) contributes, leading to

$$P_{fi} = \frac{1}{\hbar^2}|U_{fi}|^2 \frac{\sin^2\left(\Omega_- T\right)}{\Omega_-^2} = \frac{1}{\hbar^2}|U_{fi}|^2 \frac{\sin^2\Big[(E_f - E_i - E_\omega)T/2\hbar\Big]}{(E_f - E_i - E_\omega)^2/4\hbar^2}.$$

The *transition rate* is given by $\Gamma_{fi} = \mathrm{d}P_{fi}(T)/\mathrm{d}T$. Denoting now the time by t, instead of T, we have then

$$\Gamma_{fi} = \frac{\mathrm{d}P_{fi}}{\mathrm{d}t} = \frac{1}{\hbar^2}|U_{fi}|^2 \frac{2\sin\left(\Omega_- t\right)\cos\left(\Omega_- t\right)}{\Omega_-} = \frac{1}{\hbar^2}|U_{fi}|^2 \frac{\sin\left(2\Omega_- t\right)}{\Omega_-}$$
$$= \frac{2}{\hbar^2}|U_{fi}|^2 \frac{\sin\Big[(E_f - E_i - E_\omega)t/\hbar\Big]}{(E_f - E_i - E_\omega)/\hbar}.$$

Taking the limit for large t,

$$\lim_{t\to\infty} \frac{\sin\Big[(E_f - E_i - E_\omega)t/\hbar\Big]}{(E_f - E_i - E_\omega)/\hbar} = \pi\,\delta\Big((E_f - E_i - E_\omega)/\hbar\Big) = \pi\hbar\,\delta(E_f - E_i - E_\omega).$$

So finally we get to the Fermi golden rule for the absorption of energy,

$$\Gamma_{fi} = \frac{2\pi}{\hbar}|U_{fi}|^2\,\delta(E_f - E_i - E_\omega). \tag{B.88}$$

In the case of stimulated emission of energy, where $E_\omega = \hbar\omega$ is close to $E_i - E_f$, we have instead

$$\Gamma_{fi} = \frac{2\pi}{\hbar}|U_{if}|^2\,\delta(E_i - E_f - E_\omega). \tag{B.89}$$

Note how the Dirac delta function enforces the conservation of energy.

We have computed the transition rate Γ_{fi} from an initial state $|i\rangle$ to a given final state $|f\rangle$, but we may also compute the total transition rate by summing (integrating) over all possible final states. Denoting by $\rho(E_f)$ the energy density of final states, meaning that $\rho(E_f)\,\mathrm{d}E_f$ is the number of final states with energy between E_f and $E_f + \mathrm{d}E_f$, the total transition rate is given by

$$\Gamma_i = \int \Gamma_{fi}\,\rho(E_f)\,\mathrm{d}E_f = \frac{2\pi}{\hbar}\int |U_{fi}|^2\,\delta(E_f - E_i - E_\omega)\,\rho(E_f)\,\mathrm{d}E_f. \tag{B.90}$$

Therefore

$$\Gamma_i = \frac{2\pi}{\hbar}|U_{fi}|^2\,\rho(E_i + E_\omega). \tag{B.91}$$

B.4 Variational methods.

Let's see how the Schrödinger equation can be derived using a variational principle. To that end, we will find the extrema of the functional $I = \langle\psi|\mathcal{H}|\psi\rangle$ under the condition $\langle\psi|\psi\rangle = 1$. In other words, we want to find the Euler-Lagrange equation which ψ must satisfy for I to have a stationary value ($\delta I = 0$) as ψ is varied arbitrarily. This can be done by requiring that

$$\delta\Big[\,\langle\psi|\mathcal{H}|\psi\rangle - \lambda\,\langle\psi|\psi\rangle\,\Big] = 0, \tag{B.92}$$

where λ is a Lagrange multiplier. We can write this as

$$\delta\Big[\,\langle\psi|\mathcal{H}|\psi\rangle - \lambda\,\langle\psi|\psi\rangle\,\Big] = \delta\Big[\,\langle\psi|(\mathcal{H}-\lambda)|\psi\rangle\,\Big] = \langle\delta\psi|(\mathcal{H}-\lambda)|\psi\rangle + \langle\psi|(\mathcal{H}-\lambda)|\delta\psi\rangle = 0. \quad \text{(B.93)}$$

Although $|\delta\psi\rangle$ and $\langle\delta\psi|$ are not independent, it can be shown that we must require both terms in the previous equation to vanish. Indeed, let's write the variation on ψ as $\delta\psi = \delta u + i\delta v$, where δu and δv are two real independent variations. Then we have

$$\begin{aligned}
&\langle\delta\psi|(\mathcal{H}-\lambda)|\psi\rangle + \langle\psi|(\mathcal{H}-\lambda)|\delta\psi\rangle \\
&\quad= \langle\delta u|(\mathcal{H}-\lambda)|\psi\rangle + \langle i\delta v|(\mathcal{H}-\lambda)|\psi\rangle + \langle\psi|(\mathcal{H}-\lambda)|\delta u\rangle + \langle\psi|(\mathcal{H}-\lambda)|i\delta v\rangle \\
&\quad= \langle\delta u|(\mathcal{H}-\lambda)|\psi\rangle + \langle\psi|(\mathcal{H}-\lambda)|\delta u\rangle^* + \langle i\delta v|(\mathcal{H}-\lambda)|\psi\rangle + \langle i\delta v|(\mathcal{H}-\lambda)|\psi\rangle^* \\
&\quad= \langle\delta u|(\mathcal{H}-\lambda)|\psi+\psi^*\rangle - i\,\langle\delta v|(\mathcal{H}-\lambda)|\psi\rangle + i\,\langle\delta v|(\mathcal{H}-\lambda)|\psi^*\rangle \\
&\quad= \langle\delta u|(\mathcal{H}-\lambda)|\psi+\psi^*\rangle - i\,\langle\delta v|(\mathcal{H}-\lambda)|\psi-\psi^*\rangle = 0, \quad \text{(B.94)}
\end{aligned}$$

for all independent variations δu and δv. Therefore the Euler-Lagrange equations for this variational problem are

$$(\mathcal{H}-\lambda)\,|\psi+\psi^*\rangle = 0, \qquad (\mathcal{H}-\lambda)\,|\psi-\psi^*\rangle = 0, \quad \text{(B.95)}$$

or

$$(\mathcal{H}-\lambda)\,|\psi\rangle + (\mathcal{H}-\lambda)\,|\psi^*\rangle = 0, \qquad (\mathcal{H}-\lambda)\,|\psi\rangle - (\mathcal{H}-\lambda)\,|\psi^*\rangle = 0. \quad \text{(B.96)}$$

Adding and subtracting the two equations we obtain

$$(\mathcal{H}-\lambda)\,|\psi\rangle = 0, \qquad (\mathcal{H}-\lambda)\,|\psi^*\rangle = \langle\psi|\,(\mathcal{H}-\lambda) = 0. \quad \text{(B.97)}$$

These two are equivalent forms of the Schrödinger equation $(\mathcal{H}-E)\,|\psi\rangle = 0$, where the Lagrange multiplier gets to be interpreted as an eigenvalue of the Hamiltonian. Thus any function ψ for which the functional $I = \langle\psi|\mathcal{H}|\psi\rangle$ is stationary is an eigenfunction of $\mathcal{H}$ with eigenvalue E. Conversely, if ψ is an eigenfunction of $\mathcal{H}$ and E the corresponding energy, the functional is stationary.

The variational method is very useful to find approximate solutions to the Schrödinger equation. The method needs to be provided with a first trial function ψ intended to serve as an approximate solution. We define an energy E as

$$E = \frac{\langle\psi|\mathcal{H}|\psi\rangle}{\langle\psi|\psi\rangle}. \quad \text{(B.98)}$$

Let ψ_k be an eigenstate of $\mathcal{H}$ with eigenvalue E_k. Then

$$E - E_k = \frac{\langle\psi|\mathcal{H}|\psi\rangle}{\langle\psi|\psi\rangle} - E_k = \frac{\langle\psi|\mathcal{H}|\psi\rangle - \langle\psi|\psi\rangle\,E_k}{\langle\psi|\psi\rangle} = \frac{\langle\psi|(\mathcal{H}-E_k)|\psi\rangle}{\langle\psi|\psi\rangle}. \quad \text{(B.99)}$$

We will assume that ψ and ψ_k differ by a small variation,

$$\psi = \psi_k + \delta\psi, \qquad \text{with} \qquad \langle\psi_k|\delta\psi\rangle = 0. \quad \text{(B.100)}$$

Taking into account this orthogonality relation and the Hermitian property of $\mathcal{H}$, we have

$$\begin{aligned}
E - E_k &= \frac{\langle\psi_k+\delta\psi|(\mathcal{H}-E_k)|\psi_k+\delta\psi\rangle}{\langle\psi|\psi\rangle} \\
&= \frac{\langle\psi_k|(\mathcal{H}-E_k)|\psi_k\rangle + \langle\psi_k|(\mathcal{H}-E_k)|\delta\psi\rangle + \langle\delta\psi|(\mathcal{H}-E_k)|\psi_k\rangle + \langle\delta\psi|(\mathcal{H}-E_k)|\delta\psi\rangle}{\langle\psi|\psi\rangle} \\
&= \frac{\langle\delta\psi|(\mathcal{H}-E_k)|\delta\psi\rangle}{\langle\psi|\psi\rangle}. \quad \text{(B.101)}
\end{aligned}$$

This tells us that the difference between E and the true eigenvalue E_k when the trial function ψ differs from the true eigenstate ψ_k by $\delta\psi$, varies quadratically with $\delta\psi$. The lowest non-vanishing order of the variation in E is of second order and therefore the first-order variation in E must vanish, $\delta E = 0$. This is the so-called *variational principle*, which says that the best approximation to the eigenvalue E_k is obtained by varying ψ, or the parameters on which it may depend, so that the condition $\delta E = 0$ is satisfied.

Let ψ_k be a complete orthonormal set of eigenstates with eigenvalues E_k. Then we may write

$$\mathcal{H}\psi_k = E_k\psi_k, \qquad \langle\psi_k|\psi_\ell\rangle = \delta_{k\ell}, \qquad \psi = \sum_k a_k\psi_k. \tag{B.102}$$

We will assume ψ is already normalized, so

$$E = \langle\psi|\mathcal{H}|\psi\rangle = \sum_{k\ell} a_k^* a_\ell \langle\psi_k|\mathcal{H}|\psi_\ell\rangle = \sum_{k\ell} a_k^* a_\ell E_\ell \langle\psi_k|\psi_\ell\rangle = \sum_k |a_k|^2 E_k, \tag{B.103}$$

$$\langle\psi|\psi\rangle = \sum_{k\ell} a_k^* a_\ell \langle\psi_k|\psi_\ell\rangle = \sum_{k\ell} a_k^* a_\ell \delta_{k\ell} = \sum_k |a_k|^2 = 1. \tag{B.104}$$

Now let E_0 be the ground state, that is, $E_0 \le E_k$ for all k. Then

$$E = \sum_k |a_k|^2 E_k \ge E_0 \sum_k |a_k|^2 = E_0. \tag{B.105}$$

What we have obtained is the very sensible result that the energy E associated to the trial function ψ can never be below the ground state. The closer ψ is to the ground state, the smaller the difference $E - E_0$ will be. The typical procedure when trying to find the ground state energy of a system using the variational method is to optimize a trial function with respect to one or several parameters. Of course, one needs some physical insight into the problem to choose an appropriate trial function in the first place. It should also be noted that the trial function optimized with respect to the total energy is not necessarily the best function with which to calculate properties other than the energy.

Let's apply the variational method to a specific form of trial function, which we write as a linear combination of some other set of functions which are not themselves necessarily orthogonal,

$$\psi = \sum_i a_i\phi_i, \qquad \text{where} \qquad \langle\phi_i|\phi_j\rangle = S_{ij} \tag{B.106}$$

is known as the *overlap integral*. Now we have

$$E = \frac{\langle\psi|\mathcal{H}|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\langle\sum_i a_i\phi_i|\mathcal{H}|\sum_j a_j\phi_j\rangle}{\langle\sum_i a_i\phi_i|\sum_j a_j\phi_j\rangle} = \frac{\sum_{i,j} a_i^* a_j \langle\phi_i|\mathcal{H}|\phi_j\rangle}{\sum_{i,j} a_i^* a_j S_{ij}} = \frac{\sum_{i,j} a_i^* a_j \mathcal{H}_{i,j}}{\sum_{i,j} a_i^* a_j S_{ij}}, \tag{B.107}$$

where $\mathcal{H}_{i,j} = \langle\phi_i|\mathcal{H}|\phi_j\rangle$. When both sides of the equation

$$E\sum_{i,j} a_i^* a_j S_{ij} = \sum_{i,j} a_i^* a_j \mathcal{H}_{i,j} \tag{B.108}$$

are differentiated with respect to one of the coefficients, for example a_k^*, one obtains

$$\left(\frac{\partial E}{\partial a_k^*}\right)\sum_{i,j} a_i^* a_j S_{ij} + E\sum_j a_j S_{kj} = \sum_j a_j \mathcal{H}_{kj}. \tag{B.109}$$

From this, a stationary value for the energy can be found by setting $\partial E/\partial a_k^* = 0$, which leads to

$$\sum_j a_j(\mathcal{H}_{kj} - ES_{kj}) = 0. \tag{B.110}$$

This procedure has to be repeated for the rest of the coefficients, obtaining additional equations of the same type. We'll have as many conditions as the number of functions ϕ_i, forming a system of homogeneous equations which can be solved when the following determinant (secular equation) vanishes,

$$\begin{vmatrix} \mathcal{H}_{11} - ES_{11} & \mathcal{H}_{12} - ES_{12} & \dots \\ \mathcal{H}_{21} - ES_{21} & \mathcal{H}_{22} - ES_{22} & \dots \\ \vdots & \vdots & \vdots \end{vmatrix} = 0. \tag{B.111}$$

The solution to the secular equation provides a set of energies, the lowest one being an upper bound to the true ground state energy. The remaining energies are often interpreted as approximations to the energies of the excited states, although this has to be done with care and with the necessary attention to the fact that states associated to different energies must be orthogonal.

B.5 Rotations, spin and angular momentum.

The laws of nature should not depend on how we may care to orient our laboratory. It is then expected for our fundamental theories to be rotational invariant. It will be shown that this invariance leads to the existence of a conserved angular momentum three–vector $\mathbf{J}$. A rotation in the ordinary three–dimensional space is a linear transformation $x_i' = \sum_j R_{ij} x_j$ of the Cartesian coordinates x_i that leaves invariant the scalar product $\mathbf{xy} = \sum_i x_i y_i$.

$$\sum_i x_i' x_i' = \sum_i \left(\sum_j R_{ij} x_j \right) \left(\sum_k R_{ik} x_k \right) = \sum_{j,k} \left(\sum_i R_{ij} R_{ik} \right) x_j x_k = \sum_j x_j x_j. \tag{B.112}$$

This implies

$$\sum_i R_{ij} R_{ik} = \delta_{jk}, \qquad \text{or} \qquad R^T R = 1, \tag{B.113}$$

where R^T denotes the transpose of a matrix, $[R^T]_{ji} = [R]_{ij}$, and 1 stands for the unit matrix, $[1]_{jk} = \delta_{jk}$. Since $\det(AB) = \det(A)\det(B)$, and $\det(A^T) = \det(A)$, we see that $[\det(R)]^2 = 1$, so $\det(R) = \pm 1$. Every matrix with non-zero determinant has an inverse, and multiplying the relation $R^T R = 1$ on the right by R^{-1} tells us that $R^T = R^{-1}$. This inverse is also an orthogonal matrix, because $(R^{-1})^T R^{-1} = (R^T)^T R^T = R R^T = 1$. Note also that $(AB)^T = B^T A^T$, which implies that the product of any two orthogonal matrices is also an orthogonal matrix:

$$(AB)^T AB = B^T A^T AB = B^T B = 1. \tag{B.114}$$

The set of all real orthogonal 3×3 matrices forms the group known as $O(3)$.

Space inversions, $\mathbf{x} \to -\mathbf{x}$, satisfy also the relation $R^T R = 1$. To restrict our attention only to rotations we need to demand the determinant of R to be $+1$, $\det(R) = +1$. The rotations form a group of $O(3)$, known as the special orthogonal group in three dimensions $SO(3)$. In N dimensions, the special orthogonal group is denoted as $SO(N)$.

Of course, the operator R acts on the ordinary three-dimensional space, not on the Hilbert space. Nevertheless, a rotation induces a unitary transformation on the Hilbert space of physical states, $\Psi' = U(R)\Psi$. It is clear that the following relation should hold,

$$U(R_2)U(R_1) = U(R_2 R_1). \tag{B.115}$$

On any operator $\mathbf{V}$ representing a vector observable, $U(R)$ must induce the following rotation[6]

$$U^{-1}(R) V_i U(R) = \sum_j R_{ij} V_j. \tag{B.116}$$

[6] This could be taken as the definition of a vector operator. That is, a vector operator is one that transforms under a rotation as shown in equation (B.116). Alternatively, as we will see in a moment, a vector operator can be defined through its commutation relations with the generators of rotations.

An infinitesimal rotation can be written as

$$R_{ij} = \delta_{ij} + \omega_{ij} + O(\omega^2), \qquad R = I + \omega + O(\omega^2). \tag{B.117}$$

The second relation is written in matrix notation. The condition $R^T R = 1$ gives

$$(I + \omega^T + O(\omega^2))(1 + \omega + O(\omega^2)) = 1 + \omega^T + \omega + O(\omega^2), \tag{B.118}$$

so we must have $\omega = -\omega^T$, which means

$$\omega_{ij} = -\omega_{ji}. \tag{B.119}$$

The most general unitary operator $U(R)$ for such infinitesimal rotations is written as a power series in the ω parameters. The first non-trivial term must be a linear combination formed by products of the ω_{ij} parameters and a set of operators J_{ij} on the Hilbert space. Thus[7]

$$U(1+\omega) = \mathbf{I} - \frac{i}{2\hbar}\sum_{ij}\omega_{ij}J_{ij} + O(\omega^2). \tag{B.120}$$

The factor $1/\hbar$ is inserted in order to give J_{ij} the dimensions of $\hbar$, the same as distance times momentum. It can be easily checked that with this definition J_{ij} is a set of Hermitian operators. Indeed, write the relation $U^\dagger U = \mathbf{I}$ as

$$\left(\mathbf{I} + \frac{i}{2\hbar}\sum_{ij}\omega_{ij}J^\dagger_{ij}\right)\left(\mathbf{I} - \frac{i}{2\hbar}\sum_{k\ell}\omega_{k\ell}J_{k\ell}\right) = \mathbf{I} - \frac{i}{2\hbar}\sum_{k\ell}\omega_{k\ell}J_{k\ell} + \frac{i}{2\hbar}\sum_{ij}\omega_{ij}J^\dagger_{ij}$$
$$= \mathbf{I} - \frac{i}{2\hbar}\sum_{k\ell}\omega_{k\ell}\left(J_{k\ell} - J^\dagger_{k\ell}\right) = \mathbf{I} \quad \Longrightarrow \quad J_{k\ell} = J^\dagger_{k\ell}. \tag{B.121}$$

It can be seen that the commutators of the components of a vector operator with the J_{ij} are

$$\frac{i}{\hbar}[V_k, J_{ij}] = \delta_{ik}V_j - \delta_{jk}V_i. \tag{B.122}$$

The commutation rule for the J_{ij} themselves is[8]

$$\frac{i}{\hbar}[J_{ij}, J_{k\ell}] = -\delta_{i\ell}J_{kj} + \delta_{ik}J_{\ell j} - \delta_{j\ell}J_{ik}. \tag{B.123}$$

So far, all this could be applied to rotationally invariant theories in spaces of any dimensionality.[9] In three dimensions it is very convenient to express J_{ij} in terms of a three-component operator $\mathbf{J}$, defined by

$$J_1 \equiv J_{23}, \qquad J_2 \equiv J_{31}, \qquad J_3 \equiv J_{12}. \tag{B.124}$$

[7] Some authors write equation (B.120) with a plus sign, which leads to write the unitary operator for finite rotations as $U(R) = \exp(i\mathbf{J}\boldsymbol{\omega}/\hbar)$. This corresponds to the *passive* point of view for rotations, where the reference frame itself is rotated and the physical system is kept fixed. That is the convention adopted, for example, in Weinberg's book [15]. We are using here the *active* point of view for rotations, where the physical system is rotated and the reference frame is kept fixed. This is the convention adopted in Sakurai-Napolitano's book [13], and leads to $U(R) = \exp(-i\mathbf{J}\boldsymbol{\omega}/\hbar)$, as shown below.

[8] Although this is not difficult to prove, we skip the derivation here. For more details the reader is referred to chapter 4 of [15].

[9] The rotation group in N dimensions, $SO(N)$, formed by the group of $N \times N$ orthogonal matrices with determinant $+1$, has $N(N-1)/2$ generators. It is useful to note that, in N dimensions, a simple rotation is performed in a plane, not around an axis as we are used to think in 3D, so the generators correspond to independent planes of rotations. In more than 3 dimensions there are several different base vectors orthogonal to a given plane defined by any two base vectors, therefore it is impossible to define rotation axes. The number of independent planes you can choose from N dimensions is the number of ways to pick 2 directions out of N:

$$\binom{N}{2} = \frac{N(N-1)}{2}.$$

This gives the number of generators. For $N = 3$ we have 3 generators, for $N = 4$ we have 6 etc.

In a more compact way,

$$J_k \equiv \frac{1}{2}\sum_{ij}\epsilon_{ijk}J_{ij}, \qquad J_{ij} = \sum_k \epsilon_{ijk}J_k, \tag{B.125}$$

where ϵ_{ijk} is the usual totally antisymmetric Levi–Civita tensor, whose only non–vanishing components are $\epsilon_{123} = \epsilon_{231} = \epsilon_{312} = +1$ and $\epsilon_{213} = \epsilon_{321} = \epsilon_{132} = -1$. The unitary operator for infinitesimal rotations takes now the form

$$U(1+\omega) = \mathbf{I} - \frac{i}{\hbar}\boldsymbol{\omega}\cdot\mathbf{J} + \mathbf{O}(\omega^2), \tag{B.126}$$

where

$$w_k \equiv \frac{1}{2}\sum_{ij}\epsilon_{ijk}w_{ij}. \tag{B.127}$$

The rotation here is by an infinitesimal angle $|\boldsymbol{\omega}|$ around an axis in the direction of $\boldsymbol{\omega}$. The three–vector operator $\mathbf{J}$ is the generator of spatial rotations. This can be taken as the *definition* of angular momentum in quantum mechanics. The unitary operator for an arbitrary rotation, not necessarily infinitesimal, is given by

$$U(R) = \exp\left(-\frac{i}{\hbar}\mathbf{J}\cdot\boldsymbol{\omega}\right). \tag{B.128}$$

It can be seen that the commutation relations between the components of the three–vector operator $\mathbf{J}$ and any other three–vector operator $\mathbf{V}$ are

$$[J_i, V_j] = i\hbar\sum_k \epsilon_{ijk}V_k. \tag{B.129}$$

In fact, a vector operator can be defined as one that verifies the previous commutation relations with the generators of spatial rotations. A scalar operator commutes with those generators, $[J_i, S] = 0$. Also, as appropriate for a vector operator, the commutation relations between the components of $\mathbf{J}$ itself are

$$[J_i, J_j] = i\hbar\sum_k \epsilon_{ijk}J_k. \tag{B.130}$$

Using just the these commutation relations it is possible to work out the eigenvalues of $\mathbf{J}^2$ and J_3, and the action of $\mathbf{J}$ on a multiplet of eigenvectors of these operators. It is found that the common eigenstates of $\mathbf{J}^2$ and J_3 can be labelled with two numbers, j and m, with j ***integer or half integer*** and m running by unit steps from $-j$ to $+j$. Denoting the eigenstates by $|jm\rangle$, the eigenvalue equations are

$$J_3\,|jm\rangle = \hbar m\,|jm\rangle\,, \qquad \mathbf{J}^2\,|jm\rangle = \hbar^2 j(j+1)\,|jm\rangle\,, \qquad m = -j, -j+1\ldots+j. \tag{B.131}$$

We have also

$$\left(J_1 \pm iJ_2\right)|jm\rangle = \hbar\sqrt{j(j+1) - m(m\pm 1)}\;|jm\pm 1\rangle\,. \tag{B.132}$$

By carefully studying how the unitary operator $U(R)$ acts on the wave functions in the coordinate representation, $\Psi(\mathbf{X})$, it can be seen that the total angular momentum operator $\mathbf{J}$ acts on $\Psi(\mathbf{X})$ as

$$\mathbf{J}\Psi(\mathbf{X}) = -i\hbar(\mathbf{X}\times\boldsymbol{\nabla})\Psi(\mathbf{X}). \tag{B.133}$$

Defining the orbital angular momentum operator as[10] $\mathbf{L} \equiv \mathbf{X}\times\mathbf{P}$, the previous relation becomes

$$\mathbf{J}\Psi(\mathbf{X}) = \mathbf{L}\Psi(\mathbf{X}). \tag{B.134}$$

[10] Note that Weissbluth uses a slightly different definition: $\mathbf{L} = \frac{1}{\hbar}\mathbf{X}\times\mathbf{P}$.

This says that $\mathbf{J}$ acts on $\Psi(\mathbf{X})$ like the $\mathbf{L}$ operator, but it *does not* say that $\mathbf{J} = \mathbf{L}$. As we will see in a momment, $\mathbf{J}$ may have a term which does not act on space wave functions, like

$$\mathbf{J}\Psi(\mathbf{X}) = (\mathbf{L} + \mathbf{S})\,\Psi(\mathbf{X}) = \mathbf{L}\Psi(\mathbf{X}). \tag{B.135}$$

But we did not define $\mathbf{S}$ yet. By direct computation it is easy to see that $\mathbf{L}$ satisfies the same commutation relations as $\mathbf{J}$,

$$[L_i, L_j] = i\hbar \sum_k \epsilon_{ijk} L_k. \tag{B.136}$$

However, the true generator of rotations is $\mathbf{J}$, not $\mathbf{L}$, and since $\mathbf{L}$ is a vector it must satisfy the same commutation relations with $\mathbf{J}$ as any other vector (equation B.129). Therefore

$$[J_i, L_j] = i\hbar \sum_k \epsilon_{ijk} L_k. \tag{B.137}$$

If we now *define* an operator $\mathbf{S} \equiv \mathbf{J} - \mathbf{L}$, so that

$$\mathbf{J} = \mathbf{L} + \mathbf{S}, \tag{B.138}$$

then by subtracting (B.136) from (B.137) we find

$$[J_i, L_j] - [L_i, L_j] = i\hbar \sum_k \epsilon_{ijk} L_k - i\hbar \sum_k \epsilon_{ijk} L_k = 0.$$

But

$$[J_i, L_j] - [L_i, L_j] = [L_i + S_i, L_j] - [L_i, L_j] = [L_i, L_j] + [S_i, L_j] - [L_i, L_j] = [S_i, L_j].$$

Therefore

$$[S_i, L_j] = 0. \tag{B.139}$$

From equations (B.139), (B.138), (B.136) and (B.130) it is possible to infer that

$$[S_i, S_j] = i\hbar \sum_k \epsilon_{ijk} S_k. \tag{B.140}$$

The operator $\mathbf{S}$ acts as a new kind of angular momentum, and may be thought of as an internal property of the particle, called the *spin*. The spin operator is not constructed from the particle's position and momentum operators, and commutes with them. Direct calculation gives

$$[L_i, X_j] = i\hbar \sum_k \epsilon_{ijk} X_k, \qquad [L_i, P_j] = i\hbar \sum_k \epsilon_{ijk} P_k. \tag{B.141}$$

On the other hand, as every vector operator, $\mathbf{X}$ and $\mathbf{P}$ must satisfy

$$[J_i, X_j] = i\hbar \sum_k \epsilon_{ijk} X_k, \qquad [J_i, P_j] = i\hbar \sum_k \epsilon_{ijk} P_k. \tag{B.142}$$

The difference between equations (B.141) and (B.142) gives

$$[S_i, X_j] = [S_i, P_j] = 0. \tag{B.143}$$

The *ladder operators*, defined as $J_\pm = (J_1 \pm iJ_2)$, satisfy

$$J_\pm \left|jm\right\rangle = \hbar\sqrt{j(j+1) - m(m\pm 1)}\ \left|j, (m\pm 1)\right\rangle, \tag{B.144}$$

$$J_1 = \frac{1}{2}(J_+ + J_-), \qquad J_2 = \frac{1}{2i}(J_+ - J_-). \tag{B.145}$$

It is usual to define, by analogy with the spherical harmonics $Y_1^m(\theta,\varphi)$, the spherical components of a general angular momentum $\mathbf{J}$ as

$$J_{+1} = -\frac{1}{\sqrt{2}}(J_x + iJ_y), \qquad J_0 = J_z, \qquad J_{-1} = \frac{1}{\sqrt{2}}(J_x - iJ_y). \tag{B.146}$$

$$J_x = -\frac{1}{\sqrt{2}}(J_{+1} - J_{-1}), \qquad J_z = J_0, \qquad J_y = \frac{i}{\sqrt{2}}(J_{+1} + J_{-1}). \tag{B.147}$$

The following relations are verified,

$$[J_0, J_{\pm 1}] = \pm J_{\pm 1}, \quad [J_{+1}, J_{-1}] = -J_0, \quad J^2 = J_x^2 + J_y^2 + J_z^2. \tag{B.148}$$

$$\begin{aligned} J^2 = -J_{+1}J_{-1} + J_0^2 - J_{-1}J_{+1} &= \sum_q (-1)^q J_q J_{-q} \\ &= -2J_{+1}J_{-1} + J_0(J_0 - 1) = -2J_{-1}J_{+1} + J_0(J_0 + 1) \end{aligned} \tag{B.149}$$

J^2 commutes with all the rectangular and spherical components of $\mathbf{J}$, $[J^2, J_\mu] = 0$. The action of the $J_{\pm 1}$ operators on the angular momentum eigenstates is

$$J_{\pm 1}\,|jm\rangle = \mp\hbar\sqrt{\tfrac{1}{2}\left[j(j+1) - m(m\pm 1)\right]}\;|j,(m\pm 1)\rangle\,. \tag{B.150}$$

We have written before the decomposition of the angular momentum operator in terms of its spherical components. This can be done for any other vector (or tensor) operator. For example, if $\mathbf{A} = (A_1, A_2, A_3) = (A_x, A_y, A_x)$ are the cartesian components of an arbitrary vector operator, its spherical components are given by

$$A_1^{+1} = -\frac{1}{\sqrt{2}}(A_x + iA_y), \qquad A_1^0 = A_z, \qquad A_1^{-1} = \frac{1}{\sqrt{2}}(A_x - iA_y). \tag{B.151}$$

They are the components of the vector $\mathbf{A}$ when it is expanded in the basis of spherical vectors, which are

$$\xi_1 = -\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \qquad \xi_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \qquad \xi_{-1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}. \tag{B.152}$$

B.5.1 Angular momentum operators for $j = 1/2$, $j = 1$ and $j = 3/2$.

The spin-j representation matrix elements of the angular momentum operators can be computed using the following relations,

$$\left\langle jm' \middle| J_3^{(j)} \middle| jm \right\rangle = \left[J_3^{(j)}\right]_{m'm} = m\hbar\,\delta_{m'm}. \tag{B.153}$$

$$\left[J_1^{(j)}\right]_{m'm} = \frac{1}{2}\left(\left[J_+^{(j)}\right]_{m'm} + \left[J_-^{(j)}\right]_{m'm}\right). \tag{B.154}$$

$$\left[J_2^{(j)}\right]_{m'm} = \frac{1}{2i}\left(\left[J_+^{(j)}\right]_{m'm} - \left[J_-^{(j)}\right]_{m'm}\right). \tag{B.155}$$

$$\left[J_+^{(j)}\right]_{m'm} = \hbar\sqrt{j(j+1) - m(m+1)}\,\delta_{m',m+1}. \tag{B.156}$$

$$\left[J_-^{(j)}\right]_{m'm} = \hbar\sqrt{j(j+1) - m(m-1)}\,\delta_{m',m-1}. \tag{B.157}$$

The spin-j representation of the angular momentum operators for $j = 1/2$ is

$$J_3^{(1/2)} = \frac{\hbar}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad J_1^{(1/2)} = \frac{\hbar}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad J_2^{(1/2)} = \frac{\hbar}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{B.158}$$

$$J_1^{(1/2)} = \frac{\hbar}{2}\sigma_1, \qquad J_2^{(1/2)} = \frac{\hbar}{2}\sigma_2, \qquad J_3^{(1/2)} = \frac{\hbar}{2}\sigma_3, \tag{B.159}$$

where σ_1, σ_2 and σ_3 are the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{B.160}$$

The basis states are

$$|1/2, 1/2\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad |1/2, -1/2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{B.161}$$

The spin-j representation of the angular momentum operators for $j = 1$ is

$$J_3^{(1)} = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \qquad J_1^{(1)} = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad J_2^{(1)} = \frac{i\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}. \tag{B.162}$$

The basis states are

$$|1, 1\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad |1, 0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad |1, -1\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{B.163}$$

The spin-j representation of the angular momentum operators for $j = 3/2$ is

$$J_3^{(3/2)} = \hbar \begin{pmatrix} 3/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & -1/2 & 0 \\ 0 & 0 & 0 & -3/2 \end{pmatrix}, \qquad J_1^{(3/2)} = \frac{\hbar}{2} \begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ \sqrt{3} & 0 & 2 & 0 \\ 0 & 2 & 0 & \sqrt{3} \\ 0 & 0 & \sqrt{3} & 0 \end{pmatrix}, \tag{B.164}$$

$$J_2^{(3/2)} = \frac{\hbar}{2i} \begin{pmatrix} 0 & \sqrt{3} & 0 & 0 \\ -\sqrt{3} & 0 & 2 & 0 \\ 0 & -2 & 0 & \sqrt{3} \\ 0 & 0 & -\sqrt{3} & 0 \end{pmatrix}, \tag{B.165}$$

The basis states are

$$\left|\frac{3}{2}, \frac{3}{2}\right\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \qquad \left|\frac{3}{2}, \frac{1}{2}\right\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \left|\frac{3}{2}, \frac{-1}{2}\right\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \left|\frac{3}{2}, \frac{-3}{2}\right\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \tag{B.166}$$

Note that the previous representation matrices are written in the J_3-eigenstate basis, therefore the J_3 matrix is always diagonal. However, the matrix representations (B.268) and (B.270) for $j = 1$ are written in the Cartesian basis. The relation between the cartesian components and the spherical components of a vector can be seen in equation (B.151).[11]

Rotations of 2π for integer and half-integer spin systems.

From the previous matrix representations it is easy to see that a rotation of a $|j, m\rangle$ state by an angle θ about the z-axis can be written as

$$U_3(\theta)\,|j, m\rangle = \exp\left(-\frac{i}{\hbar}\theta J_3\right)\,|j, m\rangle \;=\; e^{-i\theta m}\,|j, m\rangle\,. \tag{B.167}$$

[11] Using the J_z matrix representation given by expression (B.268), check that the state vectors ξ_1, ξ_0 and ξ_{-1} given in (B.152) are in fact J_z eigenstates.

This tells us that a $\theta = 2\pi$ rotation of a half-integer spin system (m half-integer) does not return the same state, because we get an extra minus sign. Indeed,

$$U_3(2\pi)\,|j,m\rangle \;=\; e^{-i2\pi m}\;|j,m\rangle \;=\; e^{-i\pi}\;|j,m\rangle \;=\; -\,|j,m\rangle\,. \tag{B.168}$$

To recover the exact original state we need to perform a 4π rotation

$$U_3(4\pi)\,|j,m\rangle \;=\; e^{-i4\pi m}\;|j,m\rangle \;=\; e^{-i2\pi}\;|j,m\rangle \;=\; +\,|j,m\rangle\,. \tag{B.169}$$

However, for integer spin systems (m integer) a 2π rotation does return the original state,

$$U_3(2\pi)\,|j,m\rangle \;=\; e^{-i2\pi m}\;|j,m\rangle \;=\; e^{-i2\pi}\;|j,m\rangle \;=\; +\,|j,m\rangle\,. \tag{B.170}$$

A mathematical object that picks up a minus sign after a 2π rotation wouldn't be admissible in classical mechanics. However, remember that observable magnitudes in quantum mechanics are quadratic in the wavefunctions, therefore the minus sign is completely acceptable in this case. It turns out that Nature fully exploits the mathematics of rotations and there are plenty of half-integer spin systems. In fact, as far as we know, all elementary matter particles (quarks and leptons) are spin-1/2 systems.

B.5.2 Representations of the rotation operator: rotation matrices.

Consider a system described by a set of vector states $|j,m\rangle$ with definite value of angular momentum j and 3-component m. We want to find the matrix elements of the unitary operator $U(R)$ that implements spatial rotations on this system. Remember that the operator $U(R)$ is given by

$$U(R) = \exp\left(-\frac{i}{\hbar}\mathbf{J}\cdot\boldsymbol{\omega}\right). \tag{B.171}$$

The rotation is performed around the direction of the vector $\boldsymbol{\omega}$ by an angle of $|\boldsymbol{\omega}|$ radians. $[J^2, J_i] = 0$, so $[U(R), J^2] = 0$ and the matrix elements of $U(R)$ must be diagonal in j,

$$\left\langle j', m' \,\middle|\, U(R) \,\middle|\, j, m\right\rangle \propto \delta_{j'j}. \tag{B.172}$$

For a given j, the $U(R)$ operator can be represented by a $(2j+1)\times(2j+1)$ matrix, denoted by $D^{(j)}_{m'm}(R)$, whose elements are

$$D^{(j)}_{m'm}(R) \equiv \left\langle j, m' \,\middle|\, U(R) \,\middle|\, j, m\right\rangle. \tag{B.173}$$

These are called rotation matrices or Wigner matrices. A rotated vector state can be written as

$$|j,m\rangle_R = U(R)\,|j,m\rangle = \sum_{m'} |j,m'\rangle\,\langle j,m'|U(R)|j,m\rangle\,. \tag{B.174}$$

This tells us directly how to obtain $|j,m\rangle_R$ in terms of the rotation matrix elements

$$|j,m\rangle_R = \sum_{m'} D^{(j)}_{m'm}(R)\,|j,m'\rangle\,. \tag{B.175}$$

As an example let's quote the rotation matrix for a system of $j = 1/2$,

$$D^{(1/2)}(\theta,\varphi) = \begin{pmatrix} e^{-i\varphi/2}\cos\frac{\theta}{2} & -e^{-i\varphi/2}\sin\frac{\theta}{2} \\ e^{i\varphi/2}\sin\frac{\theta}{2} & e^{i\varphi/2}\cos\frac{\theta}{2} \end{pmatrix}, \tag{B.176}$$

where θ and φ are the usual polar and azimuth angular variables of the spherical coordinates.

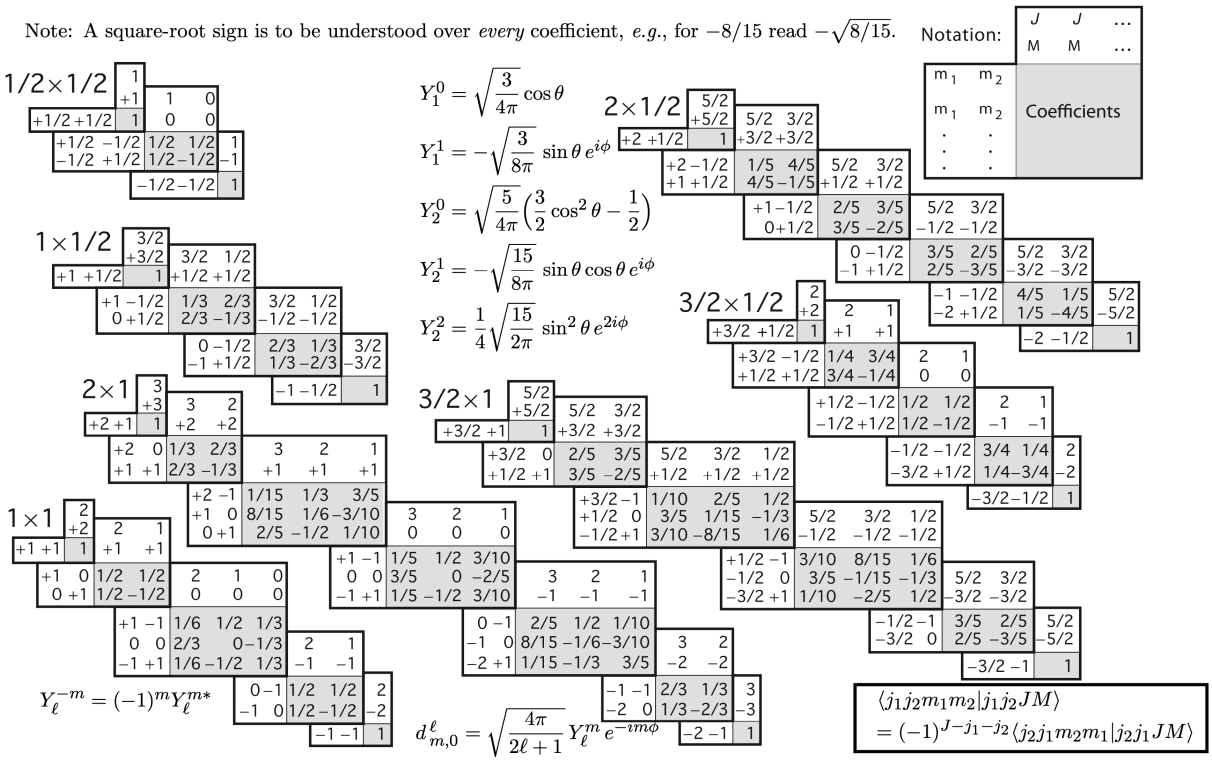


Figure B.1: Clebsch-Gordan coefficients for a few angular momentum couplings. Taken from the Review of Particle Physics [14].

B.5.3 Clebsch-Gordan coefficients, $3j$-symbols and $6j$-symbols.

Two systems with angular momentum $\mathbf{J}_1$ and $\mathbf{J}_2$ can be considered together as a global system with total angular momentum $\mathbf{J}_3 = \mathbf{J}_1 + \mathbf{J}_2$. Two different basis states of this third system can be written as $|j_1 j_2 j_3 m_3\rangle$ and $|j_1 j_2 m_1 m_2\rangle$. Using the completeness relation of the states $|j_1 j_2 m_1 m_2\rangle$ we may write

$$|j_1 j_2 j_3 m_3\rangle = \sum_{m_1,m_2} |j_1 j_2 m_1 m_2\rangle \underbrace{\langle j_1 j_2 m_1 m_2 \,|\, j_1 j_2 j_3 m_3\rangle}_{\text{CG}} . \tag{B.177}$$

The special matrix elements $\langle j_1 j_2 m_1 m_2 \,|\, j_1 j_2 j_3 m_3\rangle$ are called Clebsch-Gordan coefficients. They can be all chosen to be real, so we can also write

$$|j_1 j_2 j_3 m_3\rangle = \sum_{m_1,m_2} \langle j_1 j_2 j_3 m_3 | j_1 j_2 m_1 m_2\rangle \, |j_1 j_2 m_1 m_2\rangle \,, \tag{B.178}$$

with the inverse relations

$$|j_1 j_2 m_1 m_2\rangle = \sum_{j_3,m_3} \langle j_1 j_2 j_3 m_3 | j_1 j_2 m_1 m_2\rangle \, |j_1 j_2 j_3 m_3\rangle \,. \tag{B.179}$$

An alternative notation is,

$$\begin{aligned} \Psi^{m_3}_{j_1 j_2 j_3} &= \sum_{m_1,m_2} C_{j_1 j_2}(j_3 m_3; m_1 m_2)\, \Psi^{m_1 m_2}_{j_1 j_2}, \\ \Psi^{m_1 m_2}_{j_1 j_2} &= \sum_{j_3,m_3} C_{j_1 j_2}(j_3 m_3; m_1 m_2)\, \Psi^{m_3}_{j_1 j_2 j_3}. \end{aligned} \tag{B.180}$$

Figure (B.1) shows CG coefficients for a number of angular momentum configurations. As an example of how to use it, consider the table corresponding to the $1 \otimes 1$ coupling. Since in

that table j_1 and j_2 are fixed (to 1), we can shorten the notation and write $|m_1, m_2\rangle$ instead of $|j_1, j_2, m_1, m_2\rangle$, and $|j_3, m_3\rangle \equiv |J, M\rangle$ instead of $|j_1 j_2 j_3 m_3\rangle$. Thus

$$|m_1, m_2\rangle = \sum_{J,M} \langle m_1, m_2 | J, M\rangle \; |J, M\rangle \, . \tag{B.181}$$

From the middle portion of the $1 \otimes 1$ table is now very easy to read the following relations

$$\begin{array}{lcllll} |+1,-1\rangle & = & \sqrt{\frac{1}{6}}\,|2,0\rangle & + \sqrt{\frac{1}{2}}\,|1,0\rangle & + \sqrt{\frac{1}{3}}\,|0,0\rangle \\ |0,0\rangle & = & \sqrt{\frac{2}{3}}\,|2,0\rangle & & - \sqrt{\frac{1}{3}}\,|0,0\rangle \\ |-1,+1\rangle & = & \sqrt{\frac{1}{6}}\,|2,0\rangle & - \sqrt{\frac{1}{2}}\,|1,0\rangle & + \sqrt{\frac{1}{3}}\,|0,0\rangle \end{array} \tag{B.182}$$

Written in full detail, $|j_1 j_2 m_1 m_2\rangle = \sum_{j_3,m_3} \langle j_1 j_2 m_1 m_2 | j_1 j_2 j_3 m_3\rangle \, |j_1 j_2 j_3 m_3\rangle$, they are

$$\begin{array}{lcllll} |1,1,+1,-1\rangle & = & \sqrt{\frac{1}{6}}\,|1,1,2,0\rangle & + \sqrt{\frac{1}{2}}\,|1,1,1,0\rangle & + \sqrt{\frac{1}{3}}\,|1,1,0,0\rangle \\ |1,1,0,0\rangle & = & \sqrt{\frac{2}{3}}\,|1,1,2,0\rangle & & - \sqrt{\frac{1}{3}}\,|1,1,0,0\rangle \\ |1,1,-1,+1\rangle & = & \sqrt{\frac{1}{6}}\,|1,1,2,0\rangle & - \sqrt{\frac{1}{2}}\,|1,1,1,0\rangle & + \sqrt{\frac{1}{3}}\,|1,1,0,0\rangle \end{array} \tag{B.183}$$

Instead of using the Clebsch-Gordan coefficients to construct states of total angular momentum $\mathbf{J_3}$ from states with individual angular momenta $\mathbf{J_1}$ and $\mathbf{J_2}$, it is possible to use these coefficients to construct a state Ψ of total angular momentum zero from states $\Psi^{m_1 m_2 m_3}_{j_1 j_2 j_3}$ with *three* individual angular momenta:

$$\Psi = \sum_{m_1,m_2,m_3} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \Psi^{m_1 m_2 m_3}_{j_1 j_2 j_3} . \tag{B.184}$$

The coefficients multiplying the states $\Psi^{m_3 m_1 m_2}_{j_3 j_1 j_2}$ are called $3j$ symbols, and are related to the CG coefficients by

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \frac{(-1)^{j_1 - j_2 - m_3}}{\sqrt{2j_3 + 1}} \, \langle j_1 j_2 m_1 m_2 \, | \, j_1 j_2 j_3 - m_3 \rangle \tag{B.185}$$

Because of the symmetric way in which the three angular momenta appear in the coupling, the $3j$ symbols are symmetric under even permutations of columns,

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{pmatrix} j_2 & j_3 & j_1 \\ m_2 & m_3 & m_1 \end{pmatrix} = \begin{pmatrix} j_3 & j_1 & j_2 \\ m_3 & m_1 & m_2 \end{pmatrix} . \tag{B.186}$$

Under odd permutations of columns they are related by a factor $(-1)^{j_1+j_2+j_3}$, thus

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = (-1)^{j_1+j_2+j_3} \begin{pmatrix} j_2 & j_1 & j_3 \\ m_2 & m_1 & m_3 \end{pmatrix} . \tag{B.187}$$

Furthermore,

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = \begin{pmatrix} j_2 & j_3 & j_1 \\ m_3 & m_1 & m_2 \end{pmatrix} = \begin{pmatrix} j_3 & j_1 & j_2 \\ m_2 & m_3 & m_1 \end{pmatrix} . \tag{B.188}$$

$$\begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} = 0, \quad \text{unless} \quad m_1 + m_2 + m_3 = 0 \;\text{ and } \left.\begin{array}{c} j_1 + j_2 - j_3 \\ j_1 - j_2 + j_3 \\ -j_1 + j_2 + j_3 \end{array}\right\} \geq 0. \tag{B.189}$$

The last condition is usually denoted as $\triangle(j_1 j_2 j_3)$. A useful recursion relation for the CG coefficients is

$$\begin{aligned}\sqrt{j(j+1)-m(m-1)}\,\langle j_1 j_2 m_1 m_2|j_1 j_2 j(m-1)\rangle \\ = \sqrt{j_1(j_1+1)-m_1(m_1+1)}\,\langle j_1 j_2(m_1+1)m_2|j_1 j_2 j m\rangle \\ + \sqrt{j_2(j_2+1)-m_2(m_2+1)}\,\langle j_1 j_2 m_1(m_2+1)|j_1 j_2 j m\rangle\end{aligned} \tag{B.190}$$

Note that all three coefficients in this formula vanish unless $m_1 + m_2 = m - 1$.

When coupling three angular momenta it is common practice to introduce the so called $6j$ symbols. They are denoted as

$$\left\{ \begin{array}{ccc} j_1 & j_2 & j_3 \\ \ell_1 & \ell_2 & \ell_3 \end{array} \right\}$$

and are defined for integers and half-integers $j_1, j_2, j_3, \ell_1, \ell_2, \ell_3$ whose triads (j_1, j_2, j_3), (j_1, ℓ_2, ℓ_3), (ℓ_1, j_2, ℓ_3), and (ℓ_1, ℓ_2, j_3) satisfy the following conditions:

1. Each triad satisfies the triangular inequalities.[12]

2. The sum of the elements of each triad is an integer. Therefore, the members of each triad are either all integers or contain two half-integers and one integer.

If these conditions are not satisfied the $6j$ is zero. The $6j$ symbols are invariant under permutation of their columns, e.g.,

$$\left\{ \begin{array}{ccc} j_1 & j_2 & j_3 \\ \ell_1 & \ell_2 & \ell_3 \end{array} \right\} = \left\{ \begin{array}{ccc} j_2 & j_1 & j_3 \\ \ell_2 & \ell_1 & \ell_3 \end{array} \right\}, \tag{B.191}$$

and under exchange of two corresponding elements between rows, e.g.,

$$\left\{ \begin{array}{ccc} j_1 & j_2 & j_3 \\ \ell_1 & \ell_2 & \ell_3 \end{array} \right\} = \left\{ \begin{array}{ccc} \ell_1 & \ell_2 & j_3 \\ j_1 & j_2 & \ell_3 \end{array} \right\}. \tag{B.192}$$

The $6j$ symbols can be computed using the following relation,

$$\begin{aligned}\left\{ \begin{array}{ccc} j_1 & j_2 & j_3 \\ \ell_1 & \ell_2 & \ell_3 \end{array} \right\} = (-1)^{j_1+j_2+\ell_1+\ell_2}\,\Delta(j_1 j_2 j_3)\,\Delta(\ell_1 \ell_2 j_3)\,\Delta(\ell_1 j_2 \ell_3)\,\Delta(j_1 \ell_2 \ell_3) \\ \times \sum_k \frac{(-1)^k (j_1+j_2+\ell_1+\ell_2+1-k)!}{k!(j_1+j_2-j_3-k)!(\ell_1+\ell_2-j_3-k)!(j_1+\ell_2-\ell_3-k)!(\ell_1+j_2-\ell_3-k)!\;\alpha\;\beta}\end{aligned} \tag{B.193}$$

where the sum must be taken over all integer values of k so that the arguments of the factorials are not negative, the last two coefficients in the denominator are given by

$$\alpha = (-j_1 - \ell_1 + j_3 + \ell_3 + k)!, \qquad \beta = (-j_2 - \ell_2 + j_3 + \ell_3 + k)! \tag{B.194}$$

and the triangle relations are defined by

$$\Delta(abc) = \sqrt{\frac{(a+b-c)!(a-b+c)!(-a+b+c)!}{(a+b+c+1)!}} \tag{B.195}$$

[12] The triangular inequality for the (j_1, j_2, j_3) triad is $|j_1 - j_2| \le j_3 \le j_1 + j_2$.

B.5.4 Spherical harmonics.

The spherical harmomics, $Y_\ell^m(\theta,\varphi)$, are eigenfunctions of the orbital angular momentum and satisfy the following differential equation,

$$\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right]Y_\ell^m+\ell(\ell+1)Y_\ell^m=0. \tag{B.196}$$

They are given explicitly by

$$Y_\ell^m(\theta,\varphi)=(-1)^m\left[\frac{2\ell+1}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}\right]^{1/2}P_\ell^m(\cos\theta)e^{im\varphi}, \tag{B.197}$$

and satisfy the orthogonality relation

$$\int_{\varphi=0}^{2\pi}\int_{\theta=0}^{\pi}(Y_\ell^m)^*(\theta,\varphi)Y_{\ell'}^{m'}(\theta,\varphi)\,\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi=\delta_{mm'}\delta_{\ell\ell'}. \tag{B.198}$$

We will need to use the integral of the product of three spherical harmonics, which is given by the so-called *Gaunt formula*

$$\begin{aligned}\left\langle \ell'm'\,\middle|\,Y_L^M\,\middle|\,\ell m\right\rangle &\equiv \int (Y_{\ell'}^{m'})^*(\theta,\varphi)Y_L^M(\theta,\varphi)Y_\ell^m(\theta,\varphi)\,d\Omega\\ &=\int_0^\pi \sin\theta\,d\theta\int_0^{2\pi}d\varphi\,(Y_{\ell'}^{m'})^*(\theta,\varphi)Y_L^M(\theta,\varphi)Y_\ell^m(\theta,\varphi)\\ &=(-1)^{m'}\sqrt{\frac{1}{4\pi}(2\ell'+1)(2L+1)(2\ell+1)}\begin{pmatrix}\ell' & L & \ell\\ -m' & M & m\end{pmatrix}\begin{pmatrix}\ell' & L & \ell\\ 0 & 0 & 0\end{pmatrix}.\end{aligned} \tag{B.199}$$

From the properties of the $3j$-symbols it can be seen that this integral, or matrix element, is zero unless $-m'+M+m=0$, the sum $\ell'+L+\ell$ is an even integer, and the relation $\triangle(\ell' L\ell)$ is satisfied. It is easier to remember these rules by realizing that they follow from the requirement that the angular momentum quantum numbers ℓ', m' must arise from the angular momentum addition rules aplied to L, M and ℓ, m, which are

$$\ell'=|L-\ell|,|L-\ell|+1,\ldots,L+\ell \qquad \text{and} \qquad m'=M+m.$$

This will become clearer after reading the next section about the Wigner-Eckart theorem since, in fact, the Gaunt formula is a special case of the theorem.

B.5.5 The Wigner-Eckart theorem.

The angular momentum eigenstates $|jm\rangle$ can depend on any number of other dynamical variables which are invariant under the action of the symmetry generators J_i. In the following it will be convenient to include possible additional quantum numbers in the labelling of the eigenstates, such as $|\Phi jm\rangle$ or, in an even better notation, $|\Phi_j^m\rangle$. Using this notation we can write

$$\begin{aligned}\left(J_1+iJ_2\right)\left|\Phi_j^m\right\rangle &=\hbar\sqrt{j(j+1)-m(m+1)}\,\left|\Phi_j^{m+1}\right\rangle\\ \left(J_1-iJ_2\right)\left|\Phi_j^m\right\rangle &=\hbar\sqrt{j(j+1)-m(m-1)}\,\left|\Phi_j^{m-1}\right\rangle\\ \left(J_1\mp iJ_2\right)\left|\Phi_j^m\right\rangle &=\hbar\sqrt{j(j+1)-m(m\mp1)}\,\left|\Phi_j^{m\mp1}\right\rangle\\ \left(J_1-iJ_2\right)\left|\Phi_j^{m+1}\right\rangle &=\hbar\sqrt{j(j+1)-(m+1)(m+1-1)}\,\left|\Phi_j^m\right\rangle\\ &=\hbar\sqrt{j(j+1)-m(m+1)}\,\left|\Phi_j^m\right\rangle\end{aligned}$$

$$\begin{aligned}(J_1 + iJ_2)\left|\Phi_j^{m-1}\right\rangle &= \hbar\sqrt{j(j+1)-(m-1)(m-1+1)}\,\left|\Phi_j^m\right\rangle \\ &= \hbar\sqrt{j(j+1)-m(m-1)}\,\left|\Phi_j^m\right\rangle \\ (J_1 \mp iJ_2)\left|\Phi_j^{m\pm1}\right\rangle &= \hbar\sqrt{j(j+1)-(m\pm1)(m\pm1\mp1)}\,\left|\Phi_j^{m\pm1\mp1}\right\rangle \\ &= \hbar\sqrt{j(j+1)-m(m\pm1)}\,\left|\Phi_j^m\right\rangle\end{aligned}$$

As an application of these relations we can obtain a handy result that we will use to prove the Wigner-Eckart theorem. Let $|\Psi jm\rangle \equiv |\Psi_j^m\rangle$ be another multiplet of states which are also eigenstates of angular momentum with the same eigenvalues. Then we have

$$\begin{aligned}\hbar\sqrt{j(j+1)-m(m\pm1)}\,\left\langle\Phi_j^{m\pm1}\,\middle|\,\Psi_j^{m\pm1}\right\rangle &= \left\langle\Phi_j^{m\pm1}\,\middle|\,(J_1\pm iJ_2)\Psi_j^m\right\rangle \\ = \left\langle(J_1\mp iJ_2)\Phi_j^{m\pm1}\,\middle|\,\Psi_j^m\right\rangle &= \hbar\sqrt{j(j+1)-m(m\pm1)}\,\left\langle\Phi_j^m\,\middle|\,\Psi_j^m\right\rangle. \quad \text{(B.200)}\end{aligned}$$

Therefore

$$\left\langle\Phi_j^{m\pm1}\,\middle|\,\Psi_j^{m\pm1}\right\rangle = \left\langle\Phi_j^m\,\middle|\,\Psi_j^m\right\rangle. \quad \text{(B.201)}$$

This shows that the quantity $\left\langle\Phi_j^m\,\middle|\,\Psi_j^m\right\rangle$ is independent of m. We consider only the matrix elements in which both state vectors have equal values of j and m, because both are eigenstates of the Hermitian operators $\mathbf{J}^2$ and J_3, so the matrix element would vanish unless they both had the same eigenvalues (remember that eigenvectors with different eigenvalues are orthogonal),

$$\left\langle\Phi_{j_3}^{m_3}\,\middle|\,\Psi_{j_2}^{m_2}\right\rangle = 0, \quad \text{unless} \quad j_2 = j_3 \quad \text{and} \quad m_2 = m_3. \quad \text{(B.202)}$$

For a general operator, $\mathcal{O}$, we would not expect its matrix elements between angular momentum eigenstates, $\left\langle\Phi_{j_3}^{m_3}\,\middle|\,\mathcal{O}\,\middle|\,\Psi_{j_2}^{m_2}\right\rangle$, to vanish or to be independent of the 3-components m_3 and m_2. However, there is a special type of operators, called *irreducible tensor operators*, whose matrix elements have a very simple dependence on the angular momentum 3-components: this dependence is entirely contained in a Clebsch-Gordan coefficient.

An irreducible tensor operator of rank j is a set of $2j+1$ operators O_j^m (a tensor of $2j+1$ components), with $m = -j, -j+1, \ldots, +j$, such that the commutators of the rotation generators with these operators are

$$[J_3, O_j^m] = \hbar m\, O_j^m. \quad \text{(B.203)}$$

$$[J_1 \pm iJ_2, O_j^m] = \hbar\sqrt{j(j+1)-m(m\pm1)}\, O_j^{m\pm1}. \quad \text{(B.204)}$$

It is also said the the operators O_j^m have spin j. Examples of irreducible tensor operators are the spherical harmonics. Section B.6 contains a short introduction to irreducible tensor operators. The previous commutation relations can also be written as

$$[J_i, O_j^m] = \sum_{m'} \left[J_i^{(j)}\right]_{m'm} O_j^{m'},$$

with $i = 1, 2, 3$. These three relations are written in a compact way as

$$[\mathbf{J}, O_j^m] = \sum_{m'} \left[\mathbf{J}^{(j)}\right]_{m'm} O_j^{m'}, \quad \text{(B.205)}$$

where $\left[\mathbf{J}^{(j)}\right]_{m'm}$ is a matrix element of the spin-j representation of the angular momentum operators,[13]

$$\left[J_3^{(j)}\right]_{m'm} \equiv \hbar m \delta_{m'm}, \quad \text{(B.206)}$$

$$\left[J_1^{(j)}\right]_{m'm} \pm i\left[J_2^{(j)}\right]_{m'm} \equiv \hbar\sqrt{j(j+1)-m(m\pm1)}\,\delta_{m',m\pm1}. \quad \text{(B.207)}$$

[13] In section B.5.1 you can see the explicit representations for $j = 1/2$, $j = 1$ and $j = 3/2$.

A scalar operator, for example, is one that commutes with all components of $\mathbf{J}$, which trivially agrees with the relations (B.203) and (B.204) if we set $j = m = 0$, for which $\mathbf{J}^{(j)}_{m'm} = 0$. These relations are also consistent with our previous definition of a vector operator $\mathbf{V}$ as one that satisfies the commutation relations

$$[J_i, V_j] = i\hbar \sum_k \epsilon_{ijk} V_k. \tag{B.129}$$

Indeed, since the spherical components of this vector are

$$V^{+1} = -\frac{1}{\sqrt{2}}(V_1 + iV_2), \qquad V^{-1} = \frac{1}{\sqrt{2}}(V_1 - iV_2), \qquad V^0 = V_3, \tag{B.208}$$

from the previous commutation relations it is easy to find

$$[J_3, V^m] = m\hbar V^m, \qquad [J_1 \pm iV_2, V^m] = \hbar\sqrt{2 - m(m \pm 1)}\, V^{m\pm 1}. \tag{B.209}$$

This shows that the V^m components form a set of $2j + 1$ operators V^m_1, with $j = 1$.

The Wigner-Eckart theorem states that

$$\boxed{\left\langle \Phi^{m_3}_{j_3} \middle| O^{m_1}_{j_1} \middle| \Psi^{m_2}_{j_2} \right\rangle = \frac{1}{2j_3 + 1} \langle j_1 j_2 m_1 m_2 \,|\, j_1 j_2 j_3 m_3 \rangle \langle \Phi \,\|\, O \,\|\, \Psi \rangle} \tag{B.210}$$

where $\langle j_1 j_2 m_1 m_2 \,|\, j_1 j_2 j_3 m_3 \rangle$ is the Clebsch-Gordan coefficient associated with the coupling of two angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$ to give $\mathbf{J}_3$, and $\langle \Phi \,\|\, O \,\|\, \Psi \rangle$ is the so-called reduced matrix element, which is independent of the angular momentum orientation. It can depend on everything except the 3-components m_1, m_2 and m_3. The Clebsch-Gordan coefficients are chosen to be all real, so we can as well write $\langle j_1 j_2 j_3 m_3 \,|\, j_1 j_2 m_1 m_2 \rangle$ or, $C_{j_1 j_2}(j_3 m_3; m_1 m_2)$, giving these other two equivalent expressions of the theorem,

$$\left\langle \Phi^{m_3}_{j_3} \middle| O^{m_1}_{j_1} \middle| \Psi^{m_2}_{j_2} \right\rangle = \frac{1}{2j_3 + 1} \langle j_1 j_2 j_3 m_3 \,|\, j_1 j_2 m_1 m_2 \rangle \langle \Phi \,\|\, O \,\|\, \Psi \rangle \,. \tag{B.211}$$

$$\left\langle \Phi^{m_3}_{j_3} \middle| O^{m_1}_{j_1} \middle| \Psi^{m_2}_{j_2} \right\rangle = \frac{1}{2j_3 + 1} C_{j_1 j_2}(j_3 m_3; m_1 m_2) \, \langle \Phi \,\|\, O \,\|\, \Psi \rangle \,. \tag{B.212}$$

In some sense, the Wigner–Eckart theorem states that operating with an irreducible tensor operator of rank j_1 on the angular momentum eigenstate $\Psi^{m_2}_{j_2}$ is like adding a state with angular momentum j_1 to it. The matrix element for the spherical tensor operator is proportional to the Clebsch–Gordan coefficient which arises when adding the two angular momenta.

As an example of the usefullness of this theorem suppose we want to compute, in the dipole approximation, the transition probability of an electron from a $4d$ to a $2p$ orbital in a hydrogen atom. This involves calculating matrix elements of the form $\langle 2p, m_3 \,|\, r_i \,|\, 4d, m_2 \rangle$, where r_i is any of the x, y, z components of the position operator, and $m_3 = (-1, 0, 1)$, $m_2 = (-2, -1, 0, 1, 2)$ are the magnetic quantum numbers of the $2p$ and $4d$ orbitals. In principle there are $3 \times 5 \times 3 = 45$ integrals to do, but the Wigner-Eckart theorem allows us to compute just one of them and infer the remaining with the help of a CG table. From a qualitative point of view, this is due to the fact that all these 45 calculations are all related to each other by rotations. If an electron is in one of the three $2p$ orbitals, rotating the system is equivalent to moving the electron to a different $2p$ orbital (in general it will be a quantum superposition of the 3 basis states, $m_3 = -1, 0, +1$). Analogous statements apply for the five $4d$ orbitals and the three r_i operator components. The Wigner-Eckart theorem factorises out the angular dependence, which is encoded in the CG coefficients.

To prove the theorem, let's consider a general operator $O^{m_1}_{j_1}$ of spin j_1 and let it act on $\Psi^{m_2}_{j_2}$ to get

$$O^{m_1}_{j_1} \Psi^{m_2}_{j_2} \equiv \Omega^{m_1 m_2}_{j_1 j_2}. \tag{B.213}$$

When multiplied with the angular momentum generators this state vector becomes

$$\begin{aligned} J_i\,\Omega^{m_1m_2}_{j_1j_2} = J_i\,O^{m_1}_{j_1}\Psi^{m_2}_{j_2} &= J_i\,O^{m_1}_{j_1}\Psi^{m_2}_{j_2} - O^{m_1}_{j_1}\,J_i\,\Psi^{m_2}_{j_2} + O^{m_1}_{j_1}\,J_i\,\Psi^{m_2}_{j_2} \\ &= \left[J_i,\,O^{m_1}_{j_1}\right]\Psi^{m_2}_{j_2} + O^{m_1}_{j_1}\,J_i\,\Psi^{m_2}_{j_2}. \end{aligned} \tag{B.214}$$

The state $J_i\,\Psi^{m_2}_{j_2}$ can be written as

$$J_i\,\Psi^{m_2}_{j_2} = \sum_{m_3}\left|\Psi^{m_3}_{j_2}\right\rangle\left\langle\Psi^{m_3}_{j_2}\,\middle|\,J_i\Psi^{m_2}_{j_2}\right\rangle = \hbar\sum_{m_3}\left[J_i^{(j_2)}\right]_{m_3m_2}\Psi^{m_3}_{j_2}. \tag{B.215}$$

Furthermore, according to (B.205),

$$\left[J_i,\,O^{m_1}_{j_1}\right]\Psi^{m_2}_{j_2} = \hbar\sum_{m_3}\left[J_i^{(j_1)}\right]_{m_3m_1}O^{m_3}_{j_1}\,\Psi^{m_2}_{j_2}. \tag{B.216}$$

Therefore

$$J_i\,\Omega^{m_1m_2}_{j_1j_2} = \hbar\sum_{m_3}\left[J_i^{(j_1)}\right]_{m_3m_1}O^{m_3}_{j_1}\,\Psi^{m_2}_{j_2} \;+\; O^{m_1}_{j_1}\,\hbar\sum_{m_3}\left[J_i^{(j_2)}\right]_{m_3m_2}\Psi^{m_3}_{j_2}. \tag{B.217}$$

$$J_i\,\Omega^{m_1m_2}_{j_1j_2} = \hbar\sum_{m_3}\left[J_i^{(j_1)}\right]_{m_3m_1}\Omega^{m_3m_2}_{j_1j_2} \;+\; \hbar\sum_{m_3}\left[J_i^{(j_2)}\right]_{m_3m_2}\Omega^{m_1m_3}_{j_1j_2}. \tag{B.218}$$

The last expression shows that J_i acts on $\Omega^{m_1m_2}_{j_1j_2}$ just as if this were a state vector for a system consisting of two particles with spins j_1 and j_2, and 3–components m_1 and m_2. Therefore, using the CG coefficients, the state $\Omega^{m_1m_2}_{j_1j_2}$ can be written as the linear combination

$$\begin{aligned} \Omega^{m_1m_2}_{j_1j_2} = O^{m_1}_{j_1}\Psi^{m_2}_{j_2} &= \sum_{j_3m_3}\langle j_1j_2m_1m_2\,|\,j_1j_2j_3m_3\rangle\;\Omega^{m_3}_{j_1j_2j_3}, \\ &= \sum_{j_3m_3}C_{j_1j_2}(j_3m_3;m_1m_2)\,\Omega^{m_3}_{j_1j_2j_3}, \end{aligned} \tag{B.219}$$

where $\Omega^{m_3}_{j_1j_2j_3}$ are state vectors of angular momentum j_3 with 3–component m_3. Thus

$$\begin{aligned} \left\langle\Phi^{m_3}_{j_3}\,\middle|\,O^{m_1}_{j_1}\,\middle|\,\Psi^{m_2}_{j_2}\right\rangle &= \left\langle\Phi^{m_3}_{j_3}\,\middle|\,\sum_{j'm'}C_{j_1j_2}(j'm';m_1m_2)\,\Omega^{m'}_{j_1j_2j'}\right\rangle \\ &= \sum_{j'm'}C_{j_1j_2}(j'm';m_1m_2)\left\langle\Phi^{m_3}_{j_3}\,\middle|\,\Omega^{m'}_{j_1j_2j'}\right\rangle = C_{j_1j_2}(j_3m_3;m_1m_2)\left\langle\Phi^{m_3}_{j_3}\,\middle|\,\Omega^{m_3}_{j_1j_2j_3}\right\rangle. \end{aligned} \tag{B.220}$$

In the last step we have used the fact that $\left\langle\Phi^{m_3}_{j_3}\,\middle|\,\Omega^{m'}_{j_1j_2j'}\right\rangle$ is zero unless $j'=j_3$ and $m'=m_3$. Now, equation (B.201) applied to the state vectors Φ and Ω means that

$$\left\langle\Phi^{m_3\pm1}_{j_3}\,\middle|\,\Omega^{m_3\pm1}_{j_1j_2j_3}\right\rangle = \left\langle\Phi^{m_3}_{j_3}\,\middle|\,\Omega^{m_3}_{j_1j_2j_3}\right\rangle. \tag{B.221}$$

As a consequence, the matrix element $\left\langle\Phi^{m_3}_{j_3}\,\middle|\,\Omega^{m_3}_{j_1j_2j_3}\right\rangle$ does not depend on the 3–component m_3 value and we may write

$$\left\langle\Phi^{m_3}_{j_3}\,\middle|\,O^{m_1}_{j_1}\,\middle|\,\Psi^{m_2}_{j_2}\right\rangle = \frac{1}{2j_3+1}C_{j_1j_2}(j_3m_3;m_1m_2)\;\langle\Phi\,\|\,O\,\|\,\Psi\rangle. \tag{B.222}$$

This completes the proof of the theorem. The $1/(2j_3+1)$ factor is just chosen for convenience, but there are other conventions.

Using this result it is easy to see that the matrix elements of all vector operators for state vectors of definite angular momentum are parallel. Indeed, let $\mathbf{A}$ and $\mathbf{B}$ be any pair of vectors. Applying the Wigner–Eckart theorem we have

$$\left\langle \Phi_{j_3}^{m_3} \middle| A_1^{m_1} \middle| \Psi_{j_2}^{m_2} \right\rangle = \frac{1}{2j_3+1} C_{1j_2}(j_3 m_3; m_1 m_2)\ \langle \Phi \,\|\, \mathbf{A} \,\|\, \Psi \rangle\,. \tag{B.223}$$

$$\left\langle \Phi_{j_3}^{m_3} \middle| B_1^{m_1} \middle| \Psi_{j_2}^{m_2} \right\rangle = \frac{1}{2j_3+1} C_{1j_2}(j_3 m_3; m_1 m_2)\ \langle \Phi \,\|\, \mathbf{B} \,\|\, \Psi \rangle\,. \tag{B.224}$$

Therefore, as long as $\langle \Phi \,\|\, \mathbf{B} \,\|\, \Psi \rangle$ does not vanish,

$$\left\langle \Phi_{j_3}^{m_3} \middle| A_1^{m_1} \middle| \Psi_{j_2}^{m_2} \right\rangle = \frac{\langle \Phi \,\|\, \mathbf{A} \,\|\, \Psi \rangle}{\langle \Phi \,\|\, \mathbf{B} \,\|\, \Psi \rangle} \left\langle \Phi_{j_3}^{m_3} \middle| B_1^{m_1} \middle| \Psi_{j_2}^{m_2} \right\rangle. \tag{B.225}$$

Since this is true for the spherical components of the vectors, it is also true for the Cartesian components of the operators, A_i and B_i,

$$\left\langle \Phi_{j_3}^{m_3} \middle| A_i \middle| \Psi_{j_2}^{m_2} \right\rangle = \frac{\langle \Phi \,\|\, \mathbf{A} \,\|\, \Psi \rangle}{\langle \Phi \,\|\, \mathbf{B} \,\|\, \Psi \rangle} \left\langle \Phi_{j_3}^{m_3} \middle| B_i \middle| \Psi_{j_2}^{m_2} \right\rangle. \tag{B.226}$$

We could as well write the previous equation in vector notation, with the understanding that stands for three equations, one for each component

$$\left\langle \Phi_{j_3}^{m_3} \middle| \mathbf{A} \middle| \Psi_{j_2}^{m_2} \right\rangle = \frac{\langle \Phi \,\|\, \mathbf{A} \,\|\, \Psi \rangle}{\langle \Phi \,\|\, \mathbf{B} \,\|\, \Psi \rangle} \left\langle \Phi_{j_3}^{m_3} \middle| \mathbf{B} \middle| \Psi_{j_2}^{m_2} \right\rangle. \tag{B.227}$$

In particular, since $\mathbf{J}$ is itself a vector operator,

$$\left\langle \Phi_{j_2}^{m_3} \middle| \mathbf{A} \middle| \Psi_{j_2}^{m_2} \right\rangle = \frac{\langle \Phi \,\|\, \mathbf{A} \,\|\, \Psi \rangle}{\langle \Phi \,\|\, \mathbf{J} \,\|\, \Psi \rangle} \left\langle \Phi_{j_2}^{m_3} \middle| \mathbf{J} \middle| \Psi_{j_2}^{m_2} \right\rangle. \tag{B.228}$$

Note that we have written this result only for the case $j_2 = j_3$ because, since $\mathbf{J}$ commutes with $\mathbf{J}^2$, the matrix elements $\left\langle \Phi_{j_3}^{m_3} \middle| \mathbf{J} \middle| \Psi_{j_2}^{m_2} \right\rangle$ would vanish if Φ and Ψ had different angular momenta, $j_2 \neq j_3$. However, the third components m_2 and m_3 don't need to be equal because J_x and J_y, are expressed in terms of the ladder operators, which change the values of the third components. It is important to point out that this is a general rule only for the angular momentum operator. Other operators may have non-vanishing matrix elements between states of different total angular momentum.

The Wigner-Eckart theorem can be expressed using the $3j$-symbols as

$$\left\langle \Phi_{j_3}^{m_3} \middle| O_{j_1}^{m_1} \middle| \Psi_{j_2}^{m_2} \right\rangle = (-1)^{j_3 - m_3} \begin{pmatrix} j_3 & j_1 & j_2 \\ -m_3 & m_1 & m_2 \end{pmatrix} \langle \Phi \,\|\, O \,\|\, \Psi \rangle \tag{B.229}$$

By using this theorem we can separate those features of a physical process which depend on geometry or symmetry properties (contained in the coupling coefficients) from those which depend on the detailed physical interactions (contained in the reduced matrix elements). From the properties of the coupling coefficients we can deduce the following selection rules,

$$\left\langle \Phi_{j_3}^{m_3} \middle| O_{j_1}^{m_1} \middle| \Psi_{j_2}^{m_2} \right\rangle = 0 \qquad \text{unless} \qquad m_3 = m_1 + m_2 \quad \text{and} \quad \triangle(j_1 j_2 j_3). \tag{B.230}$$

The $3j$-symbols are real, so

$$\left\langle \Phi_{j_3}^{m_3} \middle| O_{j_1}^{m_1} \middle| \Psi_{j_2}^{m_2} \right\rangle^* = (-1)^{j_3 - m_3} \begin{pmatrix} j_3 & j_1 & j_2 \\ -m_3 & m_1 & m_2 \end{pmatrix} \langle \Phi \,\|\, O \,\|\, \Psi \rangle^*\,, \tag{B.231}$$

and since the reduced matrix elements are independent of the projection quantum numbers, we have

$$\sum_{m_2 m_3} \left| \left\langle \Phi_{j_3}^{m_3} \middle| O_{j_1}^{m_1} \middle| \Psi_{j_2}^{m_2} \right\rangle \right|^2 = |\langle \Phi \,\|\, O \,\|\, \Psi \rangle|^2 \sum_{m_2 m_3} \begin{pmatrix} j_3 & j_1 & j_2 \\ -m_3 & m_1 & m_2 \end{pmatrix}^2. \tag{B.232}$$

Now we can use the orthogonality condition of the $3j$–symbols

$$\sum_{m_1 m_2} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j_3' \\ m_1 & m_2 & m_3' \end{pmatrix} = \frac{\delta(j_3, j_3')\delta(m_3, m_3')}{(2j_3+1)}, \tag{B.233}$$

to write

$$\sum_{m_2 m_3} \left| \left\langle \Phi_{j_3}^{m_3} \middle| O_{j_1}^{m_1} \middle| \Psi_{j_2}^{m_2} \right\rangle \right|^2 = \frac{1}{2j_1+1} \left| \langle \Phi \parallel O \parallel \Psi \rangle \right|^2 . \tag{B.234}$$

Since the right hand side of this formula is independent of m_1 and there are just $2j_1+1$ values of m_1, we can perform also the sumation over m_1 to get

$$\sum_{m_1 m_2 m_3} \left| \left\langle \Phi_{j_3}^{m_3} \middle| O_{j_1}^{m_1} \middle| \Psi_{j_2}^{m_2} \right\rangle \right|^2 = \left| \langle \Phi \parallel O \parallel \Psi \rangle \right|^2 . \tag{B.235}$$

The Landé formula.

Let's return to equation (B.228). Setting $j_2 = j$, $m_3 = m$, $m_2 = m'$ and the ratio of reduced matrix elements to C, we could rewrite that equation as

$$\left\langle \Phi_j^m \middle| \mathbf{A} \middle| \Psi_j^{m'} \right\rangle = C \left\langle \Phi_j^m \middle| \mathbf{J} \middle| \Psi_j^{m'} \right\rangle . \tag{B.236}$$

To evaluate the constant C we can do

$$\left\langle \Phi_j^m \middle| \mathbf{A} \cdot \mathbf{J} \middle| \Psi_j^m \right\rangle = \sum_{m'} \left\langle \Phi_j^m \middle| \mathbf{A} \middle| \Psi_j^{m'} \right\rangle \left\langle \Psi_j^{m'} \middle| \mathbf{J} \middle| \Psi_j^m \right\rangle , \tag{B.237}$$

and use the proportionality condition to write

$$\begin{aligned} \sum_{m'} \left\langle \Phi_j^m \middle| \mathbf{A} \middle| \Psi_j^{m'} \right\rangle \left\langle \Psi_j^{m'} \middle| \mathbf{J} \middle| \Psi_j^m \right\rangle &= C \sum_{m'} \left\langle \Phi_j^m \middle| \mathbf{J} \middle| \Psi_j^{m'} \right\rangle \left\langle \Psi_j^{m'} \middle| \mathbf{J} \middle| \Psi_j^m \right\rangle \\ &= C \left\langle \Phi_j^m \middle| \mathbf{J}^2 \middle| \Psi_j^m \right\rangle = C\, j(j+1)\hbar^2. \end{aligned} \tag{B.238}$$

Therefore

$$C = \frac{\left\langle \Phi_j^m \middle| \mathbf{A} \cdot \mathbf{J} \middle| \Psi_j^m \right\rangle}{j(j+1)\hbar^2}, \tag{B.239}$$

and

$$\boxed{\left\langle \Phi_j^m \middle| \mathbf{A} \middle| \Psi_j^{m'} \right\rangle = \frac{\left\langle \Phi_j^m \middle| \mathbf{A} \cdot \mathbf{J} \middle| \Psi_j^m \right\rangle}{j(j+1)\hbar^2} \left\langle \Phi_j^m \middle| \mathbf{J} \middle| \Psi_j^{m'} \right\rangle} \tag{B.240}$$

This special form of the Wigner–Eckart theorem is known as the Landé formula.

B.6 Irreducible tensor operators.

We give here another definition of irreducible tensor operators. An irreducible tensor operator $\mathbf{T}^{(k)}$ of rank k is an object with $2k+1$ components which, under a coordinate rotation $\mathbf{r}' = R\mathbf{r}$, satisfy

$$P_R T_q^{(k)} P_R^{-1} = \sum_{q'} T_{q'}^{(k)} D_{q'q}^{(k)}(R), \qquad q, q' = -k, -k+1, \ldots, k-1, k \tag{B.241}$$

where $D_{q'q}^{(k)}(R)$ are the matrix elements associated with the irreducible representation[14] $D^{(k)}$ of the rotation group $SO(3)$. An example of irreducible tensor operators is the set of spherical

[14] See section B.5.2 for a brief introduction to the rotation matrices $D_{q'q}^{(k)}(R)$. Without going into a very precise mathematical definition, let's just say that in an irreducible representation the corresponding matrices cannot be brought into a block-diagonal form by performing a similarity transformation.

harmonics. When a rotation operator acts on an spherical harmonic eigenstate, $|jm\rangle$, the result can be generally written as a linear combination of states with the same value of j and all possible different values of m:

$$P_R\,|jm\rangle = \sum_{m'} |jm'\rangle\, D^{(j)}_{m'm}(R), \qquad D^{(j)}_{m'm}(R) = \left\langle jm' \,\middle|\, P_R \,\middle|\, jm \right\rangle. \tag{B.242}$$

So doing $j \to k$, $m \to q$ and $m' \to q'$ we have

$$D^{(k)}_{q'q}(R) = \left\langle kq' \,\middle|\, P_R \,\middle|\, kq \right\rangle. \tag{B.243}$$

Let us now see how the definition of irreducible tensor operators given in the previous section is related to the one given here. Remember that infinitesimal rotation by an angle $\delta\theta$ about the z axis can be written to first order as $P_R(\delta\theta, z) = 1 + i\delta\theta J_z$. Therefore, to first order

$$P_R(\delta\theta, z) T^{(k)}_q P^{-1}_R(\delta\theta, z) = (1 + i\delta\theta J_z) T^{(k)}_q (1 - i\delta\theta J_z) = T^{(k)}_q + i\delta\theta \left[J_z,\, T^{(k)}_q\right] \tag{B.244}$$

$$D^{(k)}_{q'q}(\delta\theta, z) = \left\langle kq' \,\middle|\, 1 + i\delta\theta J_z \,\middle|\, kq \right\rangle \tag{B.245}$$

Inserting (B.244) and (B.245) in (B.241), and using the relations $\langle kq'|kq\rangle = \delta_{q'q}$ and $J_z\,|kq\rangle = q\,|kq\rangle$ we obtain

$$T^{(k)}_q + i\delta\theta \left[J_z,\, T^{(k)}_q\right] = \sum_{q'} T^{(k)}_{q'} \left\langle kq' \,\middle|\, 1 + i\delta\theta J_z \,\middle|\, kq \right\rangle = \sum_{q'} T^{(k)}_{q'} \langle kq'|kq\rangle + i\delta\theta \sum_{q'} T^{(k)}_{q'} \langle kq'|J_z|kq\rangle. \tag{B.246}$$

Therefore

$$\left[J_z,\, T^{(k)}_q\right] = \sum_{q'} T^{(k)}_{q'} \left\langle kq' \,\middle|\, J_z \,\middle|\, kq \right\rangle = qT^{(k)}_q. \tag{B.247}$$

In a similar way we can obtain

$$\left[J_x,\, T^{(k)}_q\right] = \sum_{q'} T^{(k)}_{q'} \left\langle kq' \,\middle|\, J_x \,\middle|\, kq \right\rangle, \tag{B.248}$$

$$\left[J_y,\, T^{(k)}_q\right] = \sum_{q'} T^{(k)}_{q'} \left\langle kq' \,\middle|\, J_y \,\middle|\, kq \right\rangle. \tag{B.249}$$

Using

$$J_{+1} = -\frac{1}{\sqrt{2}}(J_x + iJ_y), \quad \text{and} \quad J_{-1} = \frac{1}{\sqrt{2}}(J_x - iJ_y),$$

with the inverse relations

$$J_x = -\frac{1}{\sqrt{2}}(J_{+1} - J_{-1}), \quad \text{and} \quad J_y = \frac{i}{\sqrt{2}}(J_{+1} + J_{-1}),$$

the two previous expressions can be combined into a single one

$$\left[J_{\pm1},\, T^{(k)}_q\right] = \sum_{q'} T^{(k)}_{q'} \left\langle kq' \,\middle|\, J_{\pm1} \,\middle|\, kq \right\rangle. \tag{B.250}$$

Remembering that

$$\langle kq'|J_{\pm1}|kq\rangle = \mp\sqrt{\tfrac{1}{2}\left[k(k+1) - q(q\pm1)\right]}\,\delta_{q',q\pm1} \tag{B.251}$$

we can write all the commutators between the components of $\mathbf{J}$ and the components of $\mathbf{T}^{(k)}$ in the following way

$$\left[J_{\pm1},\, T^{(k)}_q\right] = \mp\sqrt{\tfrac{1}{2}\left[k(k+1) - q(q\pm1)\right]}\,T^{(k)}_{q\pm1}, \qquad \left[J_0,\, T^{(k)}_q\right] = qT^{(k)}_q. \tag{B.252}$$

It can be shown that the transformation law (B.241) follows from the commutation relations (B.252), so they can be considered as an alternative definition of an irreducible tensor. Note that they are the same as the commutation relations (B.205).

The scalar product of two irreducible tensor operators is defined as

$$\mathbf{T}^{(k)} \cdot \mathbf{U}^{(k)} = \sum_q (-1)^q \, T_q^{(k)} U_{-q}^{(k)}. \tag{B.253}$$

The tensor product of two irreducible tensor operators with components $T_q^{(k)}$ and $U_{q'}^{(k')}$ is defined as the irreducible tensor operator with the following components,

$$V_Q^{(K)} = \sum_{qq'} T_q^{(k)} U_{q'}^{(k')} \, \langle kk'qq'|kk'KQ\rangle = \sqrt{2k+1} \sum_{qq'} (-1)^{-k+k'-Q} \begin{pmatrix} k & k' & K \\ q & q' & -Q \end{pmatrix} T_q^{(k)} U_{q'}^{(k')}. \tag{B.254}$$

The values of K and Q are restricted to

$$K = |k-k'|, \ldots, k+k', \qquad Q = q+q' = -K, \ldots, +K. \tag{B.255}$$

From the commutation relations of $\mathbf{T}^{(k)}$ and $\mathbf{U}^{(k')}$ it can be shown that the components of $\mathbf{V}^{(K)}$ satisfy the following commutation relations,

$$\left[J_{\pm 1}, \, V_Q^{(K)}\right] = \mp \sqrt{\tfrac{1}{2}\left[K(K+1) - Q(Q \pm 1)\right]} \, U_{Q\pm 1}^{(K)}, \qquad \left[J_0, \, V_Q^{(K)}\right] = Q V_Q^{(K)}. \tag{B.256}$$

So, indeed, $\mathbf{V}^{(K)}$ is an irreducible tensor operator.

B.7 Vector fields and multipole expansions.

B.7.1 Rotational properties of a vector field.

Let $\mathbf{V}(\mathbf{r})$ be a vector field in 3D-space and let $\mathbf{R}$ be a general proper rotation (a member of the SO(3) group, sometimes also called $\mathrm{O}^+(3)$ group). Consider the following transformations under the $\mathbf{R}$ operator,

$$\mathbf{r} \longrightarrow \mathbf{r}' = \mathbf{R}\mathbf{r}, \qquad \mathbf{V}(\mathbf{r}) \longrightarrow \mathbf{V}'(\mathbf{r}') = \mathbf{R}\Big(\mathbf{V}\big(\mathbf{R}^{-1}\mathbf{r}\big)\Big). \tag{B.257}$$

Think of $\mathbf{R}$ as a (rotation) matrix operating on a three dimensional vector. To know how $\mathbf{V}(\mathbf{r})$ transforms under a rotation we need to apply the unrotated field $\mathbf{V}$ to the unrotated vector $\mathbf{R}^{-1}\mathbf{r}$, and then rotate the whole result according to $\mathbf{R}$. That's what the expression $\mathbf{R}\Big(\mathbf{V}\big(\mathbf{R}^{-1}\mathbf{r}\big)\Big)$ means.[15] We now want to find an operator P_R such that

$$P_R\Big(\mathbf{V}(\mathbf{r})\Big) = \mathbf{R}\Big(\mathbf{V}\big(\mathbf{R}^{-1}\mathbf{r}\big)\Big). \tag{B.258}$$

Consider first an infinitesimal rotation about the z axis by an angle $d\theta$. This is given by

$$R(\delta\theta, z) = \begin{pmatrix} 1 & \delta\theta & 0 \\ -\delta\theta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{B.259}$$

So

$$R^{-1}(\delta\theta, z)\,\mathbf{r} = \begin{pmatrix} 1 & -\delta\theta & 0 \\ \delta\theta & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - y\delta\theta \\ x\delta\theta + y \\ z \end{pmatrix}, \tag{B.260}$$

[15] To know the value of the field on a rotated vector $\mathbf{r}$ (that is, on a vector expressed in the *new* reference frame), we have to apply the specific functional expression of $\mathbf{V}$ on the vector *before* it is rotated $\mathbf{R}^{-1}\mathbf{r}$ (that is, expressed in the *old* reference frame), and then rotate the result back to the *new* reference frame. Note that, in general, the functional form of $\mathbf{V}$ is itself modified by the rotation operation, $\mathbf{V} \to \mathbf{R}\mathbf{V} = \mathbf{V}'$.

$$\mathbf{V}\left(R^{-1}(\delta\theta,z)\,\mathbf{r}\right)=\begin{pmatrix}V_x(R^{-1}\mathbf{r})\\V_y(R^{-1}\mathbf{r})\\V_z(R^{-1}\mathbf{r})\end{pmatrix}, \tag{B.261}$$

$$\begin{aligned}P_R(d\theta,z)\Big(\mathbf{V}(\mathbf{r})\Big)&=\mathbf{R}\Big(\mathbf{V}\big(\mathbf{R}^{-1}\mathbf{r}\big)\Big)=R(\delta\theta,z)\left(\mathbf{V}\left(R^{-1}(\delta\theta,z)\,\mathbf{r}\right)\right)\\&=\begin{pmatrix}1&\delta\theta&0\\-\delta\theta&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}V_x(R^{-1}\mathbf{r})\\V_y(R^{-1}\mathbf{r})\\V_z(R^{-1}\mathbf{r})\end{pmatrix}=\begin{pmatrix}V_x(R^{-1}\mathbf{r})+\delta\theta V_y(R^{-1}\mathbf{r})\\-\delta\theta V_x(R^{-1}\mathbf{r})+V_y(R^{-1}\mathbf{r})\\V_z(R^{-1}\mathbf{r})\end{pmatrix}\\&=\begin{pmatrix}V_x(x-y\delta\theta,x\delta\theta+y,z)+\delta\theta V_y(x-y\delta\theta,x\delta\theta+y,z)\\-\delta\theta V_x(x-y\delta\theta,x\delta\theta+y,z)+V_y(x-y\delta\theta,x\delta\theta+y,z)\\V_z(x-y\delta\theta,x\delta\theta+y,z)\end{pmatrix}.\end{aligned} \tag{B.262}$$

Calling $P_RV_x(\mathbf{r})$ the first component of the previous vector,

$$\Big[P_R(d\theta,z)\Big(\mathbf{V}(\mathbf{r})\Big)\Big]_x=P_RV_x(\mathbf{r})=V_x(x-y\delta\theta,x\delta\theta+y,z)+\delta\theta V_y(x-y\delta\theta,x\delta\theta+y,z), \tag{B.263}$$

and writing to first order

$$P_RV_x(\mathbf{r})=V_x(\mathbf{r})+\frac{\partial V_x}{\partial x}dx+\frac{\partial V_x}{\partial y}dy+\frac{\partial V_x}{\partial z}dz, \tag{B.264}$$

we can deduce comparing[16] both expressions that

$$\begin{aligned}P_RV_x(\mathbf{r})&=V_x(\mathbf{r})+\frac{\partial V_x}{\partial x}(-y\delta\theta)+\frac{\partial V_x}{\partial y}(x\delta\theta)+\delta\theta V_y\\&=V_x(\mathbf{r})+\delta\theta\left(x\frac{\partial}{\partial y}-y\frac{\partial}{\partial x}\right)V_x+V_y\delta\theta=\left(1+\frac{i}{\hbar}\delta\theta L_z\right)V_x+\delta\theta V_y,\end{aligned} \tag{B.265}$$

where $L_z=-i\hbar(x\partial_y-y\partial_x)$ is the third component of the orbital angular momentum. Similarly,

$$P_RV_y(\mathbf{r})=\left(1+\frac{i}{\hbar}\delta\theta L_z\right)V_y-\delta\theta V_x,\qquad P_RV_z(\mathbf{r})=\left(1+\frac{i}{\hbar}\delta\theta L_z\right)V_z. \tag{B.266}$$

The three expressions can be combined into this one

$$P_R(d\theta,z)\mathbf{V}(\mathbf{r})=\left(1+\frac{i}{\hbar}\delta\theta L_z\right)\mathbf{V}(\mathbf{r})+\delta\theta\begin{pmatrix}0&1&0\\-1&0&0\\0&0&0\end{pmatrix}\begin{pmatrix}V_x(\mathbf{r})\\V_y(\mathbf{r})\\V_z(\mathbf{r})\end{pmatrix}, \tag{B.267}$$

Defining

$$S_z=\hbar\begin{pmatrix}0&-i&0\\i&0&0\\0&0&0\end{pmatrix}, \tag{B.268}$$

and $J_z=L_z+S_z$, we have

$$P_R(d\theta,z)\mathbf{V}(\mathbf{r})=\left(1+\frac{i}{\hbar}\delta\theta J_z\right)\mathbf{V}(\mathbf{r}). \tag{B.269}$$

Considering also infinitesimal rotations about the x and y axis would lead us to

$$S_x=\hbar\begin{pmatrix}0&0&0\\0&0&-i\\0&i&0\end{pmatrix},\qquad S_y=\hbar\begin{pmatrix}0&0&i\\0&0&0\\-i&0&0\end{pmatrix}. \tag{B.270}$$

[16] From $f(x+\Delta x)=f(x)+\frac{\partial f}{\partial x}\Delta x$ at first order, you get $V_x(x-y\delta\theta)=V_x(x)+\frac{\partial V_x}{\partial x}(-y\delta\theta)$, also at first order.

It can be recognised that S_x, S_y, S_z are in fact the matrix representation of the spin 1 quantum mechanical operators, $\mathbf{S} = (S_x, S_y, S_z)$. The eigenvalues of S_z are $\lambda = -1, 0, +1$. The normalized eigenvectors for each one of these eigenvalues are

$$\hat{\mathbf{e}}_+ = -\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}_{(\lambda=+1)}, \qquad \hat{\mathbf{e}}_\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}_{(\lambda=0)}, \qquad \hat{\mathbf{e}}_- = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}_{(\lambda=-1)}, \quad \text{(B.271)}$$

which are just the unit vectors in the spherical basis. In an alternative notation, we can also write

$$|SM_s\rangle = \begin{cases} |1,1\rangle = \hat{\mathbf{e}}_+ \\ |10\rangle = \hat{\mathbf{e}}_\mathbf{0} \\ |1-1\rangle = \hat{\mathbf{e}}_- \end{cases}. \quad \text{(B.272)}$$

We have found that the transformation properties of a vector field are governed by an orbital angular momentum operator $\mathbf{L}$, a spin angular momentum operator $\mathbf{S}$, and that the two operators are coupled according to the relation $\mathbf{J} = \mathbf{L} + \mathbf{S}$. Given the specific form of the $\mathbf{S}$ operator found, we conclude that a vector field has spin $S = 1$. Since the transformation of a scalar field is describable entirely on the basis of an orbital angular momentum operator, the spin of a scalar field is zero.

B.7.2 Multipole expansion of the electric and magnetic fields.

The Maxwell equations in empty space are

$$\boldsymbol{\nabla}\mathbf{B} = 0, \qquad \boldsymbol{\nabla}\mathbf{E} = 0, \qquad \boldsymbol{\nabla} \times \mathbf{B} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}, \qquad \boldsymbol{\nabla} \times \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \quad \text{(B.273)}$$

with

$$\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}, \qquad \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \quad \text{(B.274)}$$

where $\mathbf{A}$ is the vector potential. If we take the time-dependence of the field vectors to be of the form $e^{-i\omega t}$ and set $k = \omega/c$, the equations become

$$\boldsymbol{\nabla}\mathbf{B} = 0, \qquad \boldsymbol{\nabla}\mathbf{E} = 0, \qquad \boldsymbol{\nabla} \times \mathbf{B} = -ik\mathbf{E}, \qquad \boldsymbol{\nabla} \times \mathbf{E} = ik\mathbf{B}, \quad \text{(B.275)}$$

$$\mathbf{B} = \boldsymbol{\nabla} \times \mathbf{A}, \qquad \mathbf{E} = ik\mathbf{A}. \quad \text{(B.276)}$$

Working in the Coulomb gauge ($\boldsymbol{\nabla}\mathbf{A} = 0$), and using the vector identity $\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times \mathbf{E}) = \boldsymbol{\nabla}(\boldsymbol{\nabla}\mathbf{E}) - \nabla^2\mathbf{E}$, we can deduce from the previous equations the following set of vector Helmholtz equations.

$$(\nabla^2 + k^2)\mathbf{B}(\mathbf{r}) = 0, \qquad (\nabla^2 + k^2)\mathbf{E}(\mathbf{r}) = 0, \qquad (\nabla^2 + k^2)\mathbf{A}(\mathbf{r}) = 0. \quad \text{(B.277)}$$

A solution to $(\nabla^2 + k^2)\mathbf{B}(\mathbf{r}) = 0$ can be written as

$$(\nabla^2 + k^2)\mathbf{B}_\mathrm{E}(\ell, m, \mathbf{r}) = 0, \quad \text{(B.278)}$$

with

$$\mathbf{B}_\mathrm{E}(\ell, m, \mathbf{r}) = \sqrt{\ell(\ell+1)} f_\ell(kr)\mathbf{Y}_{\ell\ell m}, \quad \text{(B.279)}$$

where $f_\ell(kr)$ are spherical Bessel functions and $\mathbf{Y}_{\ell\ell m}$ are vector spherical harmonics given by

$$\begin{aligned} \mathbf{Y}_{\ell\ell m} &= \frac{1}{\sqrt{\ell(\ell+1)}}\Big(- L_- Y_\ell^m\, \hat{\mathbf{e}}_+ \;+\; L_0 Y_\ell^m\, \hat{\mathbf{e}}_0 \;+\; L_+ Y_\ell^m\, \hat{\mathbf{e}}_-\Big) \\ &= \frac{1}{\sqrt{\ell(\ell+1)}} \sum_q (-1)^q L_q Y_\ell^m\, \hat{\mathbf{e}}_{-q} = \frac{\mathbf{L} Y_\ell^m}{\sqrt{\ell(\ell+1)}}. \end{aligned} \quad \text{(B.280)}$$

$$\hat{\mathbf{e}}_+ = -\frac{1}{\sqrt{2}}(\hat{\mathbf{e}}_x + i\hat{\mathbf{e}}_y), \qquad \hat{\mathbf{e}}_0 = \hat{\mathbf{e}}_z, \qquad \hat{\mathbf{e}}_- = -\frac{1}{\sqrt{2}}(\hat{\mathbf{e}}_x - i\hat{\mathbf{e}}_y). \tag{B.281}$$

Note that $\mathbf{Y}_{\ell\ell m}$ is orthogonal to the radius vector $\mathbf{r}$ because

$$\mathbf{r}\cdot\mathbf{Y}_{\ell\ell m} = \frac{Y_\ell^m}{\sqrt{\ell(\ell+1)}}\,\mathbf{r}\cdot\mathbf{L} = \frac{Y_\ell^m}{\sqrt{\ell(\ell+1)}}\,\mathbf{r}(\mathbf{r}\times\mathbf{p}) = 0. \tag{B.282}$$

The corresponding electric field is, according to the Maxwell equations,

$$\mathbf{E}_{\mathrm{E}}(\ell, m, \mathbf{r}) = \frac{i}{k}\boldsymbol{\nabla}\times\mathbf{B}_{\mathrm{E}}(\ell, m, \mathbf{r}). \tag{B.283}$$

The orthogonality relation $\mathbf{r}\cdot\mathbf{Y}_{\ell\ell m} = 0$ implies that

$$\mathbf{r}\cdot\mathbf{B}_{\mathrm{E}}(\ell, m, \mathbf{r}) = 0, \tag{B.284}$$

therefore an electromagnetic field characterized by $\mathbf{B}_{\mathrm{E}}(\ell, m, \mathbf{r})$ and $\mathbf{E}_{\mathrm{E}}(\ell, m, \mathbf{r})$ is known as a *transverse magnetic* (TM) field of order ℓ. The transversality condition does not apply to $\mathbf{E}_{\mathrm{E}}$, which may have non-vanishing radial components. Such a field is therefore also designated as an *electric multipole* of order ℓ.

Of course, $f_\ell(kr)\mathbf{Y}_{\ell\ell m}$ also satisfies the Helmholtz equation $(\nabla^2 + k^2)\mathbf{E}(\mathbf{r}) = 0$, therefore we have the following new pair of solutions,

$$\mathbf{E}_{\mathrm{M}}(\ell, m, \mathbf{r}) = \sqrt{\ell(\ell+1)}\,f_\ell(kr)\mathbf{Y}_{\ell\ell m}, \tag{B.285}$$

$$\mathbf{B}_{\mathrm{M}}(\ell, m, \mathbf{r}) = -\frac{i}{k}\boldsymbol{\nabla}\times\mathbf{E}_{\mathrm{M}}(\ell, m, \mathbf{r}), \tag{B.286}$$

$$\mathbf{r}\cdot\mathbf{E}_{\mathrm{M}}(\ell, m, \mathbf{r}) = 0. \tag{B.287}$$

Such a radiation field is known as a *magnetic multipole* or a transverse electric (TE) field of order ℓ. In this case $\mathbf{B}_{\mathrm{M}}$ may have radial components.

The general solution to the vector Helmholtz equations for the electromagnetic field can be expressed as a linear combination of electric and magnetic multipoles of all orders,

$$\mathbf{B}(\mathbf{r}) = \sum_{\ell,m}\Big[a_{\mathrm{E}}(\ell, m)\mathbf{B}_{\mathrm{E}}(\ell, m, \mathbf{r}) + a_{\mathrm{M}}(\ell, m)\mathbf{B}_{\mathrm{M}}(\ell, m, \mathbf{r})\Big], \tag{B.288}$$

$$\mathbf{E}(\mathbf{r}) = \sum_{\ell,m}\Big[a_{\mathrm{E}}(\ell, m)\mathbf{E}_{\mathrm{E}}(\ell, m, \mathbf{r}) + a_{\mathrm{M}}(\ell, m)\mathbf{E}_{\mathrm{M}}(\ell, m, \mathbf{r})\Big], \tag{B.289}$$

where the coefficients depend on the sources and the boundary conditions. Since the parity of the vector spherical harmonic $\mathbf{Y}_{\ell\ell m}$ is $(-1)^\ell$, the parities of $\mathbf{B}_{\mathrm{E}}(\ell, m, \mathbf{r})$ and $\mathbf{E}_{\mathrm{M}}(\ell, m, \mathbf{r})$ are also both $(-1)^\ell$. The electric field $\mathbf{E}_{\mathrm{E}}(\ell, m, \mathbf{r})$ is proportional to $\boldsymbol{\nabla}\times\mathbf{B}_{\mathrm{E}}(\ell, m, \mathbf{r})$, therefore its parity must be opposite to that of $\mathbf{B}_{\mathrm{E}}(\ell, m, \mathbf{r})$, that is $(-1)^{\ell+1}$, because the curl operation transforms a polar vector into an axial vector and vice versa. The same argument applies to $\mathbf{B}_{\mathrm{M}}(\ell, m, \mathbf{r})$, which must then have parity $(-1)^{\ell+1}$. Also, since $\mathbf{A}$ is proportional to $\mathbf{E}$, both have the same parity. We have then the following parity relations,

$$\pi\big[\mathbf{B}_{\mathrm{E}}(\ell, m, \mathbf{r})\big] = \pi\big[\mathbf{E}_{\mathrm{M}}(\ell, m, \mathbf{r})\big] = \pi\big[\mathbf{A}_{\mathrm{M}}(\ell, m, \mathbf{r})\big] = (-1)^\ell, \tag{B.290}$$

$$\pi\big[\mathbf{E}_{\mathrm{E}}(\ell, m, \mathbf{r})\big] = \pi\big[\mathbf{B}_{\mathrm{M}}(\ell, m, \mathbf{r})\big] = \pi\big[\mathbf{A}_{\mathrm{E}}(\ell, m, \mathbf{r})\big] = (-1)^{\ell+1}. \tag{B.291}$$

There are no multipoles of order zero because $\mathbf{Y}_{\ell\ell m} = 0$ for $\ell = 0$. The following conventions and notations are often used,

- The parity of an electromagnetic field is defined as the parity of $\mathbf{B}$, $\pi_{\mathrm{em}} = \pi(\mathbf{B})$.
- The polarization of an electromagnetic field is defined as the direction of $\mathbf{E}$.
- The multipoles are named according to the values of 2^ℓ: dipole ($\ell = 1$, $2^\ell = 2$), quadrupole ($\ell = 2$, $2^\ell = 4$), octapole ($\ell = 3$, $2^\ell = 8$), etc. Additional designations are Eℓ for electric multipole of order ℓ and Mℓ for magnetic multipole of order ℓ. Therefore E1, E2 are electric dipole and electric quadrupole, and M1, M2 are magnetic dipole and magnetic quadrupole.

B.8 Multipole expansion of the electromagnetic scalar and vector potentials.

We provide here a brief reminder of multipole expansions for the scalar and vector potential. For more details please see for example [10] or [6]. In this section we use the SI system of units.

The electric potential created by point-like charge q at a distance r is $\phi = \frac{1}{4\pi\epsilon_0}\frac{q}{r}$, but the spatial dependence of the potential created by an arbitrary distribution of charge will be more complicated than just $1/r$. At large distances from the source any potential can be conveniently written as an expansion in powers of $1/r$, called multipole expansion. For large values of r the series will be dominated by multipoles of low order and the higher terms can be ignored.

Consider a localized distribution of charge described by a charge density $\rho(\mathbf{r}_p)$ centered around the origin of coordinates. The potential created by this distribution at large distances can be written as an expansion in spherical harmonics,

$$\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}\frac{4\pi}{2\ell+1}q_{\ell m}\frac{Y_\ell^m(\theta,\varphi)}{r^{\ell+1}}. \tag{B.292}$$

This is called a *multipole expansion*. The $\ell = 0$ term is the monopole term, $\ell = 1$ are dipole terms etc. The constants $q_{\ell m}$ depend on the specific form of the charge distribution $\rho(\mathbf{r}_p)$, and can be determined taking into account that the potential is also given by

$$\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\int\frac{\rho(\mathbf{r}_p)}{|\mathbf{r}-\mathbf{r}_p|}\,d^3\mathbf{r}_p, \tag{B.293}$$

and the expansion of $1/|\mathbf{r}-\mathbf{r}_p|$ given by

$$\frac{1}{|\mathbf{r}-\mathbf{r}_p|} = 4\pi\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}\frac{1}{2\ell+1}\frac{r_<^\ell}{r_>^{\ell+1}}(Y_\ell^m)^*(\theta_p,\varphi_p)Y_\ell^m(\theta,\varphi), \tag{B.294}$$

where the notation $r_>$ means the greater of the two distances (r, r_p), and $r_<$ means the smaller. Since we want to find the potential in a point outside the charge distribution, we take $r_< = r_p$ and $r_> = r$. We have

$$\begin{aligned}\phi(\mathbf{r}) &= \frac{1}{4\pi\epsilon_0}\int\left[4\pi\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}\frac{1}{2\ell+1}\frac{r_p^\ell}{r^{\ell+1}}(Y_\ell^m)^*(\theta_p,\varphi_p)Y_\ell^m(\theta,\varphi)\right]\rho(\mathbf{r}_p)\,d^3\mathbf{r}_p \\ &= \frac{1}{\epsilon_0}\sum_{\ell,m}\frac{1}{2\ell+1}\left[\int(Y_\ell^m)^*(\theta_p,\varphi_p)\,r_p^\ell\,\rho(\mathbf{r}_p)\,d^3\mathbf{r}_p\right]\frac{Y_\ell^m(\theta,\varphi)}{r^{\ell+1}}.\end{aligned} \tag{B.295}$$

The coefficients are then given by

$$q_{\ell m} = \int(Y_\ell^m)^*(\theta_p,\varphi_p)\,r_p^\ell\,\rho(\mathbf{r}_p)\,d^3\mathbf{r}_p. \tag{B.296}$$

These coefficients are called *multipole moments*. It is instructive to write a few of them in Cartesian coordinates to better exhibit their physical interpretation,

$$q_{00} = \frac{1}{\sqrt{4\pi}}\int\rho(\mathbf{r}_p)\,d\mathbf{r}_p = \frac{1}{\sqrt{4\pi}}q \tag{B.297}$$

$$\begin{aligned}q_{11} &= -\sqrt{\frac{3}{8\pi}}\int(x-iy)\rho(\mathbf{r}_p)\,d\mathbf{r}_p = -\sqrt{\frac{3}{8\pi}}(p_x-ip_y), \\ q_{10} &= \sqrt{\frac{3}{4\pi}}\int z\rho(\mathbf{r}_p)\,d\mathbf{r}_p = \sqrt{\frac{3}{4\pi}}p_z\end{aligned} \tag{B.298}$$

$$
\begin{aligned}
q_{22} &= \frac{1}{4}\sqrt{\frac{15}{2\pi}}\int (x-iy)^2\rho(\mathbf{r}_p)\,d\mathbf{r}_p = \frac{1}{12}\sqrt{\frac{15}{2\pi}}(Q_{11}-2iQ_{12}-Q_{22}),\\
q_{21} &= -\sqrt{\frac{15}{8\pi}}\int z(x-iy)\rho(\mathbf{r}_p)\,d\mathbf{r}_p = -\frac{1}{3}\sqrt{\frac{15}{8\pi}}(Q_{13}-iQ_{23}),\\
q_{20} &= \frac{1}{2}\sqrt{\frac{5}{4\pi}}\int (3z^2-r^2)\rho(\mathbf{r}_p)\,d\mathbf{r}_p = \frac{1}{2}\sqrt{\frac{5}{4\pi}}Q_{33}.
\end{aligned}
\tag{B.299}
$$

We have written only the moments with $m \geq 0$, since for a real charge density the moments with negative m are related by $q_{\ell,-m} = (-1)^m q^*_{\ell m}$. In the previous equations q is the total charge, or monopole moment, $\mathbf{p}$ is the electric dipole moment,

$$
\mathbf{p} = \int \mathbf{r}_p\,\rho(\mathbf{r}_p)\,d\mathbf{r}_p, \tag{B.300}
$$

and Q_{ij} is the traceless quadrupole moment tensor,

$$
Q_{ij} = \int \left[3(x_i)_p(x_j)_p - r_p^2\delta_{ij}\right]\rho(\mathbf{r}_p)\,d\mathbf{r}_p. \tag{B.301}
$$

The ℓth multipole coefficients ($2\ell+1$ of them in number) are linear combinations of the corresponding multipoles expressed in rectangular coordinates. The expansion of $\phi(\mathbf{r})$ in rectangular coordinates can be derived by a direct Taylor series expansion of $1/|\mathbf{r}-\mathbf{r}_p|$,

$$
\phi(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\left[\frac{q}{r} + \frac{\mathbf{pr}}{r^3} + \frac{1}{2}\sum_{i,j} Q_{ij}\frac{x_i x_j}{r^5} + \ldots\right] \tag{B.302}
$$

The monopole and dipole terms are

$$
\phi_{\text{mon}}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{q}{r}, \qquad \phi_{\text{dip}}(\mathbf{r}) = \frac{1}{4\pi\epsilon_0}\frac{\mathbf{pr}}{r^3} = \frac{1}{4\pi\epsilon_0}\frac{\mathbf{p}\hat{\mathbf{r}}}{r^2}. \tag{B.303}
$$

The vector potential created by a current of intensity I circulating through a closed loop of arbitrary shape is given by the following line integral around the loop

$$
\mathbf{A}(\mathbf{r}) = \frac{\mu_0 I}{4\pi}\oint \frac{1}{|\mathbf{r}-\mathbf{r}_p|}\,\mathrm{d}\mathbf{r}_p. \tag{B.304}
$$

Expanding $1/|\mathbf{r}-\mathbf{r}_p|$ as

$$
\frac{1}{|\mathbf{r}-\mathbf{r}_p|} = \frac{1}{\sqrt{r^2+r_p^2+2rr_p\cos\theta_{rr_p}}} = \frac{1}{r}\sum_{n=0}^{\infty}\left(\frac{r_p}{r}\right)^n P_n(\cos\theta_{rr_p}), \tag{B.305}
$$

we obtain a multipole expansion for $\mathbf{A}(\mathbf{r})$,

$$
\mathbf{A}(\mathbf{r}) = \frac{\mu_0 I}{4\pi}\left[\frac{1}{r}\oint \mathrm{d}\mathbf{r}_p + \frac{1}{r^2}\oint r_p\cos\theta_{rr_p}\,\mathrm{d}\mathbf{r}_p + \frac{1}{r^3}\oint r_p^2\left(\frac{3}{2}\cos^2\theta_{rr_p} - \frac{1}{2}\right)\mathrm{d}\mathbf{r}_p + \ldots\right]. \tag{B.306}
$$

The monopole term (which goes like $1/r$) is always zero because the integral is just the total vector displacement around a closed loop, $\oint \mathrm{d}\mathbf{r}_p = 0$. This, of course, is in agreement with $\boldsymbol{\nabla} B = 0$. The dipole term is

$$
\mathbf{A}_{\text{dip}}(\mathbf{r}) = \frac{\mu_0 I}{4\pi r^2}\oint r_p\cos\theta_{rr_p}\,\mathrm{d}\mathbf{r}_p = \frac{\mu_0 I}{4\pi r^2}\oint (\hat{\mathbf{r}}\mathbf{r}_p)\,\mathrm{d}\mathbf{r}_p. \tag{B.307}
$$

This expression can be written in a different way using the identity

$$
\oint (\hat{\mathbf{r}}\mathbf{r}_p)\,\mathrm{d}\mathbf{r}_p = -\hat{\mathbf{r}}\times\int \mathrm{d}\mathbf{a}_p \tag{B.308}
$$

where $\mathbf{a}_p$ is the *vector area* of the loop. If the loop is flat, then the magnitude of $\mathbf{a}_p$ is the area enclosed, with the direction assigned by the usual right hand rule (fingers in the direction of the current). We can then write

$$\mathbf{A}_{\rm dip}(\mathbf{r}) = \frac{\mu_0}{4\pi}\frac{\boldsymbol{\mu}\times\hat{\mathbf{r}}}{r^2} = \frac{\mu_0}{4\pi}\frac{\boldsymbol{\mu}\times\mathbf{r}}{r^3} = \frac{1}{4\pi\epsilon_0 c^2}\frac{\boldsymbol{\mu}\times\mathbf{r}}{r^3}, \tag{B.309}$$

where we have replaced μ_0 by c^2 using the relation $\mu_0\epsilon_0 = 1/c^2$, and $\boldsymbol{\mu}$ is the *magnetic dipole moment*

$$\boldsymbol{\mu} \equiv I\int \mathrm{d}\mathbf{a}_p = I\mathbf{a}_p \tag{B.310}$$

According to Eqs. (B.16) and (B.30), a vector potential and a magnetic dipole moment can be converted between the Gaussian and SI systems of units using the relations

$$\mathbf{A}_{\rm si} = \sqrt{\frac{\mu_0}{4\pi}}\mathbf{A}_{\rm g}, \qquad \boldsymbol{\mu}_{\rm si} = \sqrt{\frac{4\pi}{\mu_0}}\boldsymbol{\mu}_{\rm g}. \tag{B.311}$$

Therefore,

$$\mathbf{A}_{\rm g} = \sqrt{\frac{4\pi}{\mu_0}}\mathbf{A}_{\rm si} = \sqrt{\frac{4\pi}{\mu_0}}\frac{\mu_0}{4\pi}\frac{\boldsymbol{\mu}_{\rm si}\times\mathbf{r}}{r^3} = \sqrt{\frac{\mu_0}{4\pi}}\frac{\boldsymbol{\mu}_{\rm si}\times\mathbf{r}}{r^3} = \frac{\boldsymbol{\mu}_{\rm g}\times\mathbf{r}}{r^3}. \tag{B.312}$$

The vector potential generated by a magnetic dipole moment can then be written in both systems of units as

$$\mathbf{A}_{\rm si} = \frac{\mu_0}{4\pi}\frac{\boldsymbol{\mu}_{\rm si}\times\mathbf{r}}{r^3}, \qquad \mathbf{A}_{\rm g} = \frac{\boldsymbol{\mu}_{\rm g}\times\mathbf{r}}{r^3}. \tag{B.313}$$

B.9 Dirac equation up to order v^2/c^2.

Following the notation in Chapter 1, we start from the Dirac equation written in the form

$$(E' - q\phi)\psi_u = \frac{1}{2m}(\boldsymbol{\sigma\pi})K(\boldsymbol{\sigma\pi})\,\psi_u, \tag{B.314}$$

where

$$K = \frac{2mc^2}{E' + 2mc^2 - q\phi}. \tag{B.315}$$

Now we approximate K as

$$K = \frac{1}{1 + \dfrac{E' - q\phi}{2mc^2}} \approx 1 - \frac{E' - q\phi}{2mc^2}. \tag{B.316}$$

Inserting this in equation (B.314) we get

$$(E' - q\phi)\,\psi_u = \frac{1}{2m}\left[(\boldsymbol{\sigma\pi})^2 - (\boldsymbol{\sigma\pi})\frac{E' - q\phi}{2mc^2}(\boldsymbol{\sigma\pi})\right]\,\psi_u. \tag{B.317}$$

To proceed simplifying this equation it is convenient to define the operator

$$\boldsymbol{\Omega} = 1 + \frac{(\boldsymbol{\sigma\pi})^2}{8m^2c^2}. \tag{B.318}$$

To order v^2/c^2 we have

$$\boldsymbol{\Omega}^{-1} = 1 - \frac{(\boldsymbol{\sigma\pi})^2}{8m^2c^2}. \tag{B.319}$$

We define $\psi = \boldsymbol{\Omega}\psi_u$ and multiply (B.317) on the left by $\boldsymbol{\Omega}^{-1}$. Taking into account that $\psi_u = \boldsymbol{\Omega}^{-1}\psi$ we get

$$\boldsymbol{\Omega}^{-1}(E' - q\phi)\,\boldsymbol{\Omega}^{-1}\psi = \frac{1}{2m}\boldsymbol{\Omega}^{-1}\left[(\boldsymbol{\sigma\pi})^2 - (\boldsymbol{\sigma\pi})\frac{E' - q\phi}{2mc^2}(\boldsymbol{\sigma\pi})\right]\,\boldsymbol{\Omega}^{-1}\psi. \tag{B.320}$$

Let's inspect one by one the different terms of the previous equation only to order v^2/c^2.

$$\boldsymbol{\Omega}^{-1}(E'-q\phi)\,\boldsymbol{\Omega}^{-1}\psi \;\approx\; (E'-q\phi)\,\psi \;-\; \frac{(\boldsymbol{\sigma\pi})^2}{8m^2c^2}(E'-q\phi)\,\psi \;-\; (E'-q\phi)\frac{(\boldsymbol{\sigma\pi})^2}{8m^2c^2}\,\psi.$$

$$\begin{aligned}\frac{1}{2m}\boldsymbol{\Omega}^{-1}(\boldsymbol{\sigma\pi})^2\boldsymbol{\Omega}^{-1}\,\psi \;&\approx\; \frac{1}{2m}(\boldsymbol{\sigma\pi})^2\,\psi-\frac{1}{2m}\frac{(\boldsymbol{\sigma\pi})^2}{8m^2c^2}(\boldsymbol{\sigma\pi})^2\,\psi-\frac{1}{2m}(\boldsymbol{\sigma\pi})^2\frac{(\boldsymbol{\sigma\pi})^2}{8m^2c^2},\psi\\ &=\; \frac{1}{2m}(\boldsymbol{\sigma\pi})^2\,\psi-\frac{1}{8m^3c^2}(\boldsymbol{\sigma\pi})^4\,\psi.\end{aligned}$$

$$\frac{1}{2m}\boldsymbol{\Omega}^{-1}(\boldsymbol{\sigma\pi})\frac{E'-q\phi}{2mc^2}(\boldsymbol{\sigma\pi})\boldsymbol{\Omega}^{-1}\,\psi \;\approx\; \frac{1}{2m}(\boldsymbol{\sigma\pi})\frac{E'-q\phi}{2mc^2}(\boldsymbol{\sigma\pi})\,\psi \;=\; (\boldsymbol{\sigma\pi})\frac{E'-q\phi}{4m^2c^2}(\boldsymbol{\sigma\pi})\,\psi.$$

Putting everything together:

$$\begin{aligned}(E'-q\phi)\,\psi \;-\; \frac{(\boldsymbol{\sigma\pi})^2}{8m^2c^2}(E'-q\phi)\,\psi \;-\; (E'-q\phi)\frac{(\boldsymbol{\sigma\pi})^2}{8m^2c^2}\,\psi&\\ =\; \frac{1}{2m}(\boldsymbol{\sigma\pi})^2\,\psi-\frac{1}{8m^3c^2}(\boldsymbol{\sigma\pi})^4\,\psi \;-\; (\boldsymbol{\sigma\pi})\frac{E'-q\phi}{4m^2c^2}(\boldsymbol{\sigma\pi})\,\psi.&\end{aligned}\tag{B.321}$$

$$(E'-q\phi)\,\psi \;=\; \left[\frac{(\boldsymbol{\sigma\pi})^2}{2m}-\frac{(\boldsymbol{\sigma\pi})^4}{8m^3c^2}+\frac{(\boldsymbol{\sigma\pi})^2}{8m^2c^2}(E'-q\phi)-(\boldsymbol{\sigma\pi})\frac{E'-q\phi}{4m^2c^2}(\boldsymbol{\sigma\pi})+(E'-q\phi)\frac{(\boldsymbol{\sigma\pi})^2}{8m^2c^2}\right]\psi.\tag{B.322}$$

The following relations are useful to further simplify this equation.

$$\begin{aligned}(\boldsymbol{\sigma\pi})(E'-q\phi)-(E'-q\phi)(\boldsymbol{\sigma\pi}) &= E'(\boldsymbol{\sigma\pi})-q(\boldsymbol{\sigma\pi})\phi-E'(\boldsymbol{\sigma\pi})+q\phi(\boldsymbol{\sigma\pi})\\ &= q\phi(\boldsymbol{\sigma\pi})-q(\boldsymbol{\sigma\pi})\phi = q\phi\boldsymbol{\sigma}\,(\mathbf{p}-q\mathbf{A})-q\boldsymbol{\sigma}\,(\mathbf{p}-q\mathbf{A})\,\phi\\ &= q\phi(\boldsymbol{\sigma p})-q^2\phi(\boldsymbol{\sigma A})-q(\boldsymbol{\sigma p})\phi+q^2(\boldsymbol{\sigma A})\phi = q\phi(\boldsymbol{\sigma p})-q(\boldsymbol{\sigma p})\phi.\end{aligned}\tag{B.323}$$

But, remembering that the operator $\mathbf{p}$ acts on everything that stands to its right,

$$\big[\phi(\boldsymbol{\sigma p})-(\boldsymbol{\sigma p})\phi\big]\,\psi=\phi(\boldsymbol{\sigma p})\psi-[(\boldsymbol{\sigma p})\phi]\,\psi-\phi\,[(\boldsymbol{\sigma p})\psi]=-\,[(\boldsymbol{\sigma p})\phi]\,\psi,\tag{B.324}$$

which means that

$$\phi(\boldsymbol{\sigma p})-(\boldsymbol{\sigma p})\phi=i\hbar\boldsymbol{\sigma}\,(\boldsymbol{\nabla}\phi)\,.\tag{B.325}$$

Therefore

$$(\boldsymbol{\sigma\pi})(E'-q\phi)-(E'-q\phi)(\boldsymbol{\sigma\pi})=i\hbar q\boldsymbol{\sigma}\,(\boldsymbol{\nabla}\phi)\,.\tag{B.326}$$

We need also to use the following relations,

$$\begin{aligned}[\boldsymbol{\pi}\,(\boldsymbol{\nabla}\phi)-(\boldsymbol{\nabla}\phi)\,\boldsymbol{\pi}]\,\psi=[(\mathbf{p}-q\mathbf{A})\,(\boldsymbol{\nabla}\phi)]\,\psi \;-\; [(\boldsymbol{\nabla}\phi)\,(\mathbf{p}-q\mathbf{A})]\,\psi&\\ =(\mathbf{p}\,\boldsymbol{\nabla}\phi)\,\psi+(\boldsymbol{\nabla}\phi)\,(\mathbf{p}\psi)-q\mathbf{A}\,(\boldsymbol{\nabla}\phi)\,\psi-(\boldsymbol{\nabla}\phi)\,(\mathbf{p}\psi)+q\,(\boldsymbol{\nabla}\phi)\,\mathbf{A}\psi=(\mathbf{p}\,\boldsymbol{\nabla}\phi)\,\psi.&\end{aligned}$$

This means that

$$[\boldsymbol{\pi}\,(\boldsymbol{\nabla}\phi)-(\boldsymbol{\nabla}\phi)\,\boldsymbol{\pi}]=-i\hbar\boldsymbol{\nabla}\,(\boldsymbol{\nabla}\phi)\,.\tag{B.327}$$

Also, using[17] $\mathbf{p}\times\boldsymbol{\nabla}\phi=-i\hbar\,(\boldsymbol{\nabla}\times\boldsymbol{\nabla}\phi)-(\boldsymbol{\nabla}\phi\times\mathbf{p})$, and $\boldsymbol{\nabla}\times\boldsymbol{\nabla}\phi=0$, we have

$$\begin{aligned}[\boldsymbol{\pi}\times(\boldsymbol{\nabla}\phi)]-[(\boldsymbol{\nabla}\phi)\times\boldsymbol{\pi}] &= (\mathbf{p}-q\mathbf{A})\times(\boldsymbol{\nabla}\phi) \;-\; (\boldsymbol{\nabla}\phi)\times(\mathbf{p}-q\mathbf{A})\\ &=\; \mathbf{p}\times(\boldsymbol{\nabla}\phi)-q\mathbf{A}\times(\boldsymbol{\nabla}\phi)-(\boldsymbol{\nabla}\phi)\times\mathbf{p}+q\,(\boldsymbol{\nabla}\phi)\times\mathbf{A}\\ &=\; \mathbf{p}\times(\boldsymbol{\nabla}\phi)-(\boldsymbol{\nabla}\phi)\times\mathbf{p}+2q\,(\boldsymbol{\nabla}\phi)\times\mathbf{A}\\ &=\; -i\hbar\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\phi)-2\,(\boldsymbol{\nabla}\phi)\times\mathbf{p}+2q\,(\boldsymbol{\nabla}\phi)\times\mathbf{A} \;=\; -2\,(\boldsymbol{\nabla}\phi)\times(\mathbf{p}-q\mathbf{A}).\end{aligned}$$

[17] The term $\mathbf{p}\times(\boldsymbol{\nabla}\phi)$ is worked out in the same way we have obtained $\mathbf{p}\times\mathbf{A}=-i\hbar\,(\boldsymbol{\nabla}\times\mathbf{A})-\mathbf{A}\times\mathbf{p}$, and using the fact that $\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\phi)=0$.

So

$$[\boldsymbol{\pi}\times(\boldsymbol{\nabla}\phi)]-[(\boldsymbol{\nabla}\phi)\times\boldsymbol{\pi}]=-2\,(\boldsymbol{\nabla}\phi)\times\boldsymbol{\pi}. \tag{B.328}$$

Using the previous two results and the relation $(\boldsymbol{\sigma}\mathbf{a})(\boldsymbol{\sigma}\mathbf{b})=\mathbf{ab}+i\boldsymbol{\sigma}(\mathbf{a}\times\mathbf{b})$, we find

$$\begin{aligned}(\boldsymbol{\sigma}\boldsymbol{\pi})\,(\boldsymbol{\sigma}\boldsymbol{\nabla}\phi)-(\boldsymbol{\sigma}\boldsymbol{\nabla}\phi)\,(\boldsymbol{\sigma}\boldsymbol{\pi})&=\boldsymbol{\pi}\,(\boldsymbol{\nabla}\phi)+i\boldsymbol{\sigma}\,[\boldsymbol{\pi}\times(\boldsymbol{\nabla}\phi)]-(\boldsymbol{\nabla}\phi)\,\boldsymbol{\pi}-i\boldsymbol{\sigma}\,[(\boldsymbol{\nabla}\phi)\times\boldsymbol{\pi}]\\&=\boldsymbol{\pi}\,(\boldsymbol{\nabla}\phi)-(\boldsymbol{\nabla}\phi)\,\boldsymbol{\pi}+i\boldsymbol{\sigma}\Big[\,[\boldsymbol{\pi}\times(\boldsymbol{\nabla}\phi)]-[(\boldsymbol{\nabla}\phi)\times\boldsymbol{\pi}]\,\Big]\\&=-i\hbar\boldsymbol{\nabla}\,(\boldsymbol{\nabla}\phi)-2i\boldsymbol{\sigma}\,(\boldsymbol{\nabla}\phi)\times\boldsymbol{\pi}\;=\;-i\hbar\boldsymbol{\nabla}\,(\boldsymbol{\nabla}\phi)+2i(\boldsymbol{\sigma}\boldsymbol{\pi})\times(\boldsymbol{\nabla}\phi)\,.\end{aligned} \tag{B.329}$$

Collecting everything together,

$$\begin{aligned}(\boldsymbol{\sigma}\boldsymbol{\pi})^2(E'-q\phi)+(E'-q\phi)(\boldsymbol{\sigma}\boldsymbol{\pi})^2&=(\boldsymbol{\sigma}\boldsymbol{\pi})\Big[(\boldsymbol{\sigma}\boldsymbol{\pi})(E'-q\phi)-(E'-q\phi)(\boldsymbol{\sigma}\boldsymbol{\pi})\Big]\\&-\Big[(\boldsymbol{\sigma}\boldsymbol{\pi})(E'-q\phi)-(E'-q\phi)(\boldsymbol{\sigma}\boldsymbol{\pi})\Big](\boldsymbol{\sigma}\boldsymbol{\pi})+2(\boldsymbol{\sigma}\boldsymbol{\pi})(E'-q\phi)(\boldsymbol{\sigma}\boldsymbol{\pi})\\&=(\boldsymbol{\sigma}\boldsymbol{\pi})\Big[i\hbar q\boldsymbol{\sigma}\,(\boldsymbol{\nabla}\phi)\,\Big]-\Big[i\hbar q\boldsymbol{\sigma}\,(\boldsymbol{\nabla}\phi)\,\Big](\boldsymbol{\sigma}\boldsymbol{\pi})+2(\boldsymbol{\sigma}\boldsymbol{\pi})(E'-q\phi)(\boldsymbol{\sigma}\boldsymbol{\pi})\\&=i\hbar q\Big[(\boldsymbol{\sigma}\boldsymbol{\pi})\,(\boldsymbol{\sigma}\boldsymbol{\nabla}\phi)-(\boldsymbol{\sigma}\boldsymbol{\nabla}\phi)\,(\boldsymbol{\sigma}\boldsymbol{\pi})\Big]+2(\boldsymbol{\sigma}\boldsymbol{\pi})(E'-q\phi)(\boldsymbol{\sigma}\boldsymbol{\pi})\\&=i\hbar q\Big[-i\hbar\boldsymbol{\nabla}\,(\boldsymbol{\nabla}\phi)+2i(\boldsymbol{\sigma}\boldsymbol{\pi})\times(\boldsymbol{\nabla}\phi)\Big]+2(\boldsymbol{\sigma}\boldsymbol{\pi})(E'-q\phi)(\boldsymbol{\sigma}\boldsymbol{\pi})\\&=\hbar^2q\boldsymbol{\nabla}\,(\boldsymbol{\nabla}\phi)-2\hbar q(\boldsymbol{\sigma}\boldsymbol{\pi})\times(\boldsymbol{\nabla}\phi)+2(\boldsymbol{\sigma}\boldsymbol{\pi})(E'-q\phi)(\boldsymbol{\sigma}\boldsymbol{\pi}).\end{aligned} \tag{B.330}$$

Now we insert this into equation (B.322),

$$\begin{aligned}(E'-q\phi)\,\psi\;=\;\Big[&\frac{(\boldsymbol{\sigma}\boldsymbol{\pi})^2}{2m}-\frac{(\boldsymbol{\sigma}\boldsymbol{\pi})^4}{8m^3c^2}+\frac{\hbar^2q}{8m^2c^2}\boldsymbol{\nabla}\,(\boldsymbol{\nabla}\phi)-\frac{2\hbar q}{8m^2c^2}(\boldsymbol{\sigma}\boldsymbol{\pi})\times(\boldsymbol{\nabla}\phi)\\&+\frac{2}{8m^2c^2}(\boldsymbol{\sigma}\boldsymbol{\pi})(E'-q\phi)(\boldsymbol{\sigma}\boldsymbol{\pi})-\frac{1}{4m^2c^2}(\boldsymbol{\sigma}\boldsymbol{\pi})(E'-q\phi)(\boldsymbol{\sigma}\boldsymbol{\pi})\Big]\,\psi,\end{aligned}$$

$$(E'-q\phi)\,\psi\;=\;\left[\frac{(\boldsymbol{\sigma}\boldsymbol{\pi})^2}{2m}-\frac{(\boldsymbol{\sigma}\boldsymbol{\pi})^4}{8m^3c^2}+\frac{\hbar^2q}{8m^2c^2}\boldsymbol{\nabla}\,(\boldsymbol{\nabla}\phi)-\frac{\hbar q}{4m^2c^2}(\boldsymbol{\sigma}\boldsymbol{\pi})\times(\boldsymbol{\nabla}\phi)\right]\,\psi. \tag{B.331}$$

Following equation (1.82) we have

$$\frac{1}{2m}(\boldsymbol{\sigma}\boldsymbol{\pi})^2=\frac{1}{2m}\left(\mathbf{p}-\frac{q}{c}\mathbf{A}\right)^2-\frac{q\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A})\,.$$

Furthermore, to order v^2/c^2 we can write

$$\frac{1}{c^2}(\boldsymbol{\sigma}\boldsymbol{\pi})^4=\frac{1}{c^2}\left[\boldsymbol{\sigma}\left(\mathbf{p}-\frac{q}{c}\mathbf{A}\right)\right]^4\;\approx\;\frac{1}{c^2}(\boldsymbol{\sigma}\mathbf{p})^4=\frac{1}{c^2}p^4,$$

$$\begin{aligned}\frac{1}{c^2}\boldsymbol{\pi}\times(\boldsymbol{\nabla}\phi)=\frac{1}{c^2}\left(\mathbf{p}-\frac{q}{c}\mathbf{A}\right)\times(\boldsymbol{\nabla}\phi)\;&\approx\;\frac{1}{c^2}\mathbf{p}\times(\boldsymbol{\nabla}\phi)\\&=\;-\frac{i}{c^2}\hbar\boldsymbol{\nabla}\times(\boldsymbol{\nabla}\phi)-\frac{i}{c^2}(\boldsymbol{\nabla}\phi)\times\mathbf{p}\;=\;-\frac{1}{c^2}(\boldsymbol{\nabla}\phi)\times\mathbf{p}.\end{aligned}$$

Therefore we finally arrive at

$$\begin{aligned}(E'-q\phi)\,\psi\;=\;\Big[&\frac{1}{2m}\left(\mathbf{p}-\frac{q}{c}\mathbf{A}\right)^2-\frac{q\hbar}{2mc}\boldsymbol{\sigma}\cdot(\boldsymbol{\nabla}\times\mathbf{A})\\&-\frac{p^4}{8m^3c^2}+\frac{\hbar^2q}{8m^2c^2}\boldsymbol{\nabla}\,(\boldsymbol{\nabla}\phi)+\frac{\hbar q}{4m^2c^2}\boldsymbol{\sigma}\big[\,(\boldsymbol{\nabla}\phi)\times\mathbf{p}\big]\Big]\,\psi.\end{aligned} \tag{B.332}$$

B.10 Entropy, temperature and the Boltzmann factor.

This is a very terse summary of fundamental concepts. For more details and a full treatment the reader is referred to the standard textbooks on the subject, for example *Thermal Physics, C. Kittel and H. Kroemer, 1980.*

Thermal physics can be developed following a logical structure where the main concepts are entropy, temperature, the Boltzmann factor, the chemical potential, the Gibbs factor, and the distribution functions. The fundamental logical assumption of the theory is that a closed system can be in any of its accessible quantum states with equal probability. Quantum states are either accessible or inaccessible to the system, and the system is equally likely to be in any one accessible state as in any other accessible state. The entropy is a measure of the number of quantum states accessible to a system. For a system with n allowed quantum states the entropy is defined as the natural logarithm of n,

$$\sigma \;=\; \log(n). \tag{B.333}$$

The entropy is a function of the energy U, the number of particles N, and the volume V of the system, because these parameters enter in the determination of n (other parameters may enter as well). The definition $\sigma = \log(n)$ is more convenient than $\sigma = n$, since it is much easier (and natural) to deal with the combined entropy of two systems as $\sigma_1 + \sigma_2$ rather than $n_1 n_2$.

When two systems are brought into thermal contact they will transfer energy into one another while keeping the total energy constant. The initial number of combined states before contact, $n_1 n_2$, may increase after contact due to the flow of energy between the systems in one or other direction. The fundamental assumption favours the energy sharing between the two systems that maximizes the number of accessible states of the combined system. This statement is the kernel of the law of increase of entropy (second law of thermodynamics). After thermal contact, the combined system evolves until the maximum value for the number of accessible quantum states is reached. It can be shown that this maximum is attained when the value of $(\partial\sigma/\partial U)_{N,V}$ for one system is equal to the value of the same quantity for the second system,

$$\left(\frac{\partial\sigma_1}{\partial U_1}\right)_{N_1,V_1} \;=\; \left(\frac{\partial\sigma_2}{\partial U_2}\right)_{N_2,V_2}. \tag{B.334}$$

The partial derivatives here are evaluated keeping the number of particles and volume, N_i and V_i, fixed. This equality property for two systems in thermal contact is just the property we expect of the temperature. Accordingly, the fundamental temperature τ is defined by the relation

$$\frac{1}{\tau} \;\equiv\; \left(\frac{\partial\sigma}{\partial U}\right)_{N,V}. \tag{B.335}$$

The use of $1/\tau$ ensures that the energy flows from high τ to low τ. The Kelvin temperature T is proportional to τ, with $\tau = k_{\rm B}T$, where $k_{\rm B}$ is the Boltzmann constant. The conventional entropy S is given by $S = k_{\rm B}\sigma$, so

$$S = k_{\rm B}\log(n). \tag{B.336}$$

Consider now a small system with only two states, one at energy 0 and one at energy ε, that is placed in thermal contact with a large system called *reservoir*. The total energy of the combined two systems is U_0. When the small system is in the state of energy 0, the reservoir will have U_0 energy and $n(U_0)$ states available to it. When the small system is in the state of energy ε, the reservoir will have $U_0 - \varepsilon$ energy and $n(U_0 - \varepsilon)$ states available to it. By the fundamental assumption, the ratio of the probability of finding the small system with energy ε to the probability of finding it with energy 0 is given by

$$\frac{P(\varepsilon)}{P(0)} = \frac{n(U_0-\varepsilon)}{n(U_0)} = \frac{\exp[\sigma(U_0-\varepsilon)]}{\exp[\sigma(U_0)]} = \exp\big[\sigma(U_0-\varepsilon) - \sigma(U_0)\big]. \tag{B.337}$$

By expanding the reservoir entropy σ in a Taylor series and using the definition of fundamental temperature we arrive, neglecting terms of higher order, at

$$\sigma(U_0 - \varepsilon) \approx \sigma(U_0) - \varepsilon \left(\frac{\partial \sigma}{\partial U} \right)_{U_0} = \sigma(U_0) - \frac{\epsilon}{\tau}. \tag{B.338}$$

Therefore $\sigma(U_0 - \varepsilon) - \sigma(U_0) = -\varepsilon/\tau$ and

$$\frac{P(\varepsilon)}{P(0)} = e^{-\varepsilon/\tau}. \tag{B.339}$$

This is Boltzmann's result.

Bibliography

[1] H. A. Bethe and E. E. Salpeter. *Quantum Mechanics of one- and two-electron atoms.* Springer-Verlag, 1957.

[2] B. H. Bransden and C. J. Joachain. *Physics of Atoms and Molecules (second edition).* Prentice Hall, 2003.

[3] Sánchez del Río (coordinador). *Física Cuántica.* Pirámide, 2017.

[4] R. Eisberg and R. Resnick. *Quantum Physics of Atoms, Molecules, Solids, Nuclei and Particles.* Wiley, 1985.

[5] M. Fox. *A Student's Guide to Atomic Physics.* Cambridge University Press, 2018.

[6] D. J. Griffiths. *Introduction to Electrodynamics (fourth edition).* Cambridge University Press, 2017.

[7] D. J. Griffiths and D. F. Schroeter. *Introduction to Quantum Mechanics (third edition).* Cambridge University Press, 2018.

[8] David Griffiths. *Introduction to Elementary Particles.* Wiley-VCH, 2008.

[9] H. Haken and H. C. Wolf. *The Physics of Atoms and Quanta.* Springer–Verlag., 2004.

[10] J. D. Jackson. *Classical Electrodynamics.* John Wiley & Sons, Inc., 1998.

[11] L. D. Landau and E. M. Lifshitz. *Quantum Mechanics; non-relatitivistic theory.* Pergamon Press, 1977.

[12] A. Rubbia. *Phenomenology of Particle Physics.* Cambridge University Press, 2022.

[13] J. J. Sakurai and J. Napolitano. *Modern Quantum Mechanics (second edition).* Cambridge University Press, 2017.

[14] M. Tanabashi et al. Review of particle physics. *Phys. Rev. D*, D98:030001, 2018.

[15] S. Weinberg. *Lectures on Quantum Mechanics (second edition).* Cambridge University Press, 2015.

[16] Mitchel Weissbluth. *Atoms and Molecules.* Academic Press, Inc., 1978.