

8.421 AMO I LECTURE NOTES

Contents

Preface	1
1 The Two-State System: Resonance	2
1.1 Introduction	2
1.2 Resonance Studies and Q.E.D.	3
1.2.1 The language of resonance: a classical damped system	4
1.3 Magnetic Resonance: Classical Spin in Time-Varying B-Field	6
1.3.1 The classical motion of spins in a static magnetic field	6
1.3.2 Rotating coordinate transformation	7
1.3.3 Larmor's theorem	7
1.3.4 Motion in a Rotating Magnetic Field	8
1.3.5 Rapid Adiabatic Passage: Landau-Zener Crossing	10
1.3.6 Adiabatic Following in a Magnetic Trap	12
1.4 Resonance Of Quantized Spin	15
1.4.1 Expectation value of magnetic moment behaves classically	15
1.4.2 The Rabi transition probability	16
1.4.3 The Hamiltonian of a quantized spin $\frac{1}{2}$	17
1.4.4 Quantum Mechanical Solution for Resonance in a Two-State System	18
1.4.5 Interaction representation	19
1.4.6 Two-state problem	19
1.4.7 Solution via rotating frame	21
1.4.8 Rapid adiabatic passage - Quantum treatment	24
1.4.9 Adiabatic passage - Detailed calculation	26
1.5 Density Matrix	28
1.5.1 General results	28
1.5.2 Density matrix for two level system	30
1.5.3 Phenomological treatment of relaxation: Bloch equations	32
1.5.4 Introduction: Electrons, Protons, and Nuclei	33
2 Atoms: Some Basics	35
2.1 Spectroscopic Notation	36
2.2 One-Electron Atoms	37
2.2.1 The Bohr Atom	37
2.2.2 Radial Schrödinger equation for central potentials	38
2.2.3 Radial equation for hydrogen	40
2.2.4 Expectation values of $1/r^k$ for hydrogen wavefunction	42

2.3	One Electron Atoms with Cores: The Quantum Defect	43
2.3.1	Phenomenology	43
2.3.2	Explanation	44
2.3.3	Quantum defects for a model atom	46
2.4	Metrology and Precision Measurement and Units	48
2.4.1	Dimensions and Dimensional Analysis	48
2.4.2	SI units	49
2.4.3	Metrology	52
2.5	Universal Units and Fundamental Constants	53
2.6	Atomic Units	54
2.7	The Fine Structure Constant	55
3	Fine Structure and Lamb Shift	56
3.1	Fine Structure	56
3.1.1	Kinetic contribution	57
3.1.2	Spin-Orbit Interaction	57
3.1.3	The Darwin Term	61
3.1.4	Evaluation of the fine structure interaction	61
3.2	The Lamb Shift	62
4	Effects of the Nucleus on Atomic Structure	64
4.1	Hyperfine interaction	65
4.1.1	Introduction	65
4.1.2	Classical analysis of the magnetic hyperfine interaction	65
4.1.3	Hyperfine structure at zero magnetic field	69
4.1.4	Electric quadrupole interaction	71
4.1.5	Multipole expansion	71
4.1.6	Theoretical restrictions on multipole orders	74
4.1.7	Simple examples for a non-zero quadrupole moment	77
4.1.8	Calculating quadrupole moment tensor	78
4.1.9	Diagonalization of the electric quadrupole interaction	79
4.1.10	Order of magnitude of hyperfine structure	85
4.2	Isotope Effects	86
4.2.1	Mass effect	86
4.2.2	Volume effect	87
5	Atoms in Magnetic Fields	89
5.1	The Landé g -factor	89
5.1.1	Magnetic moment of circulating charge (classical)	89
5.1.2	Intrinsic electron spin and magnetic moment	90
5.1.3	Vector model of the Landé g -factor	90

iii	0.0
-----	-----

5.2	Hyperfine structure in an applied field	92
5.2.1	Low field	92
5.2.2	High field	92
5.2.3	General solution	93
6	Atoms in Electric Fields	93
6.1	Atoms in a Static Electric Field	96
6.2	Some Results of Stationary Perturbation Theory	98
6.3	Perturbation Theory of Polarizability	99
6.4	Beyond the quadratic Stark effect	100
6.5	Field ionization	102
6.6	Atoms in an Oscillating Electric Field	104
6.7	Oscillator Strength	106
6.8	DC Stark Effect of a diatomic molecule	109
7	Interaction of an Atom with an Electro-Magnetic Field	113
7.1	Introduction: Spontaneous and Stimulated Emission	113
7.2	Quantum Theory of Absorption and Emission	116
7.3	The classical E-M field	116
7.4	Interaction of an electromagnetic wave and an atom	117
7.5	Quantization of the radiation field	119
7.6	Interaction of a two-level system and a single mode of the radiation field	121
7.7	Absorption and emission	123
7.8	Spontaneous emission rate	125
7.9	Line Strength	126
7.10	Excitation by narrow and broad band light sources	127
7.11	Higher-order radiation processes	129
7.12	Selection rules	131

Preface

The original incarnation of these notes was developed to accompany the lectures in the MIT graduate courses in atomic physics. AMO I was created in the late 1960s as a one-term introductory course to prepare graduate students for research in atomic physics in the Physics department. Over the years Dan Kleppner and David Pritchard changed the contents of the course to reflect new directions of research, though the basic concepts remained as a constant thread. With the growth of interest in atom cooling and quantum gases, a second one-term course, AMO II, was designed by Wolfgang Ketterle in the late 1990s and presented with AMO I in alternating years. We still teach AMO I in the traditional way. These lecture notes combine the (g)olden notes of Dan and Dave. As part of the Joint Harvard/MIT Center for Ultracold Atoms summer school in Atomic Physics in 2002, John Doyle got involved and improved the notes. They have been circulated since and at some point were put into the form of an AMO Wiki. At this moment in time, I'd like to resurrect them in their traditional paper form, and only carefully add topics as I see fit.

Martin Zwierlein,

Spring 2022

Chapter 1

The Two-State System: Resonance

1.1 Introduction

The cornerstone of major areas of contemporary Atomic, Molecular and Optical Physics (AMO Physics) is the study of atomic and molecular systems through their resonant interaction with applied oscillating electromagnetic fields. The thrust of these studies has evolved continuously since Rabi performed the first resonance experiments in 1938. In the decade following World War II the edifice of quantum electrodynamics was constructed largely in response to resonance measurements of unprecedented accuracy on the properties of the electron and the fine and hyperfine structure of simple atoms. At the same time, nuclear magnetic resonance and electron paramagnetic resonance were developed and quickly became essential research tools for chemists and solid state physicists. Molecular beam magnetic and electric resonance studies yielded a wealth of information on the properties of nuclei and molecules, and provided invaluable data for the nuclear physicist and physical chemist. With the advent of lasers and laser spectroscopy these studies evolved into the creation of new species, such as Rydberg atoms, to studies of matter in ultra intense fields, to fundamental studies in the symmetries of physics, to new types of metrology. With the advent of laser cooling and trapping, these techniques led to the creation of novel atomic quantum fluids, from Bose-Einstein condensates to strongly interacting Fermi gases.

Resonance techniques may be used not only to learn about the structure of a system, but also to prepare it in a precisely controlled way. Because of these two facets, resonance studies have led physicists through a fundamental change in attitude - from the passive study of atoms to the active control of their internal quantum state and their interactions with the radiation field. The chief technical legacy of the early work on resonance spectroscopy is the family of lasers which have sprung up like the brooms of the sorcerer's apprentice. The scientific applications of these devices have been prodigious. They caused the resurrection of physical optics - what we now call quantum optics - and turned it into one of the liveliest fields in physics. They have had a similar impact on atomic and molecular spectroscopy. In addition, lasers have led to new families of physical studies such as single particle spectroscopy, multiphoton excitation, cavity quantum electrodynamics, and laser cooling and trapping. This chapter is about the interactions of a two-state system with a sinusoidally oscillating field whose frequency is close to the natural resonance frequency of the system. The term "two-level" system is sometimes used, but this is less accurate than the term two-state, because the levels could be degenerate, comprising several states. However, its misusage is so widespread that we adopt it anyway. The oscillating field will be treated classically, and the linewidth of both states will be taken as zero until near the end of the chapter where relaxation will be treated phenomenologically. The organization of the material is historical because this

happens to be also a logical order of presentation. The classical driven oscillator is discussed first, then the magnetic resonance of a classical spin, and then a quantized spin. The density matrix is introduced last and used to treat systems with damping - this is a useful prelude to the application of resonance ideas at optical frequencies and to the many real systems which have damping.

1.2 Resonance Studies and Q.E.D.

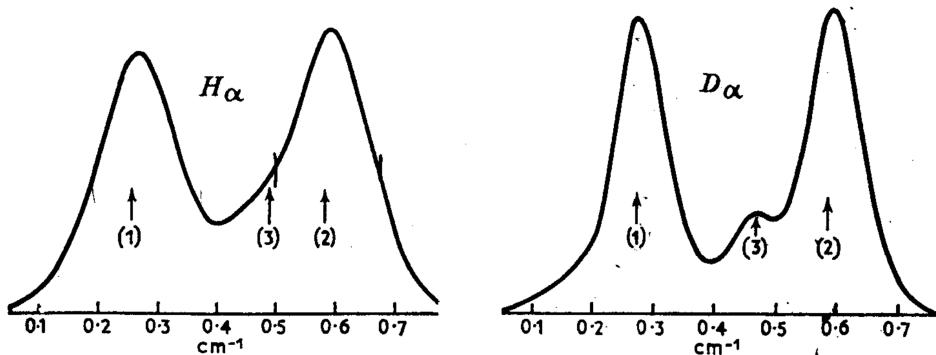


Figure 1. Spectral profile of the H_{α} line of atomic hydrogen by conventional absorption spectroscopy. Components 1) and 2) arise from the fine structure splitting. The possibility that a third line lies at position 3) was suggested to indicate that the Dirac theory might need to be revised. (From “The Spectrum of Atomic Hydrogen”-Advances. G.W. Series ed., World Scientific, 1988).

One characteristic of atomic resonance is that the results, if you can obtain them at all, are generally of very high accuracy, so high that the information is qualitatively different from other types. The hydrogen fine structure is a good example.

In the late 1930s there was extensive investigation of the Balmer series of hydrogen, ($n > 2 \rightarrow n = 2$). The Dirac Theory was thought to be in good shape, but some doubts were arising. Careful study of the Balmer-alpha line ($n = 3 \rightarrow n = 2$) showed that the line shape might not be given accurately by the Dirac Theory.

Pasternack, in 1939, suggested that the $^2S_{1/2}$ and $^2P_{1/2}$ states were not degenerate, but that the energy of the $^2S_{1/2}$ state was greater than the Dirac value by $\sim .04 \text{ cm}^{-1}$ (or, in frequency, $\sim 1200 \text{ MHz}$). However, there was no rush to throw out the Dirac theory on such flimsy evidence.

In 1947, Lamb found a splitting between the $^2S_{1/2}$ and $^2P_{1/2}$ levels using a resonance method. The experiment is one of the classics of physics. Although his very first observation was relatively crude, it was nevertheless accurate to one percent. He found

$$S_H = \frac{1}{h} [E(^2S_{1/2}) - E(^2P_{1/2})] = 1050(10) \text{ MHz} \quad (1.1)$$

The inadequacy of the Dirac theory was inescapably demonstrated.

The magnetic moment of the electron offers another example. In 1925, Uhlenbeck and Goudsmit suggested that the electron has intrinsic spin angular momentum $s = 1/2$ (in units of \hbar) and magnetic moment

$$|\mu_e| = \frac{e\hbar}{2m} = \mu_B \quad (1.2)$$

where μ_B is the Bohr magneton (and $\mu_e = -\mu_B$ is negative). The evidence was based on studies of the multiplicity of atomic lines (in particular, the Zeeman structure). The proposal was revolutionary, but the accuracy of the prediction that $|\mu_e| = \mu_B$ was poor, essentially one significant figure. According to the Dirac theory, $|\mu_e| = \mu_B$, exactly. However, our present understanding is

$$\frac{|\mu_e|}{\mu_B} - 1 = 1.1596521884(43) \times 10^{-3} \quad (\text{experiment, U. of Washington}) \quad (1.3)$$

This result is in good agreement with theory, the limiting factor in the comparison being possible doubts about the value of the fine structure constant.

The Lamb shift and the departure of $|\mu_e|$ from μ_B resulted in the award of the 1955 Nobel prize to Lamb and Kusch, and provided the experimental basis for the theory of quantum electrodynamics for which Feynman, Schwinger and Tomonaga received the Nobel Prize in 1965.

1.2.1 The language of resonance: a classical damped system

Because the terminology of classical resonance, as well as many of its features, are carried over into quantum mechanics, we start by reviewing an elementary resonant system. Consider a harmonic oscillator composed of a series RLC circuit. The charge obeys

$$\ddot{q} + \gamma\dot{q} + \omega_0^2 q = 0 \quad (1.4)$$

where $\gamma = R/L$, $\omega_0^2 = 1/LC$. Assuming that the system is underdamped (i.e. $\gamma^2 < 4\omega_0^2$), the solution for q is a linear combination of

$$\exp\left(-\frac{\gamma}{2}t\right) \exp(\pm i\omega't) \quad (1.5)$$

where $\omega' = \omega_0\sqrt{1 - \gamma^2/4\omega_0^2}$. If $\omega_0 \gg \gamma$, which is often the case, we have $\omega' \equiv \omega_0$. The energy in the circuit is

$$W = \frac{1}{2C}q^2 + \frac{1}{2}L\dot{q}^2 = W_0 e^{-\gamma t} \quad (1.6)$$

where $W_0 = W(t = 0)$. The decay time of the stored energy is $\tau = \frac{1}{\gamma}$. If the circuit is driven by a voltage $E_0 e^{i\omega t}$, the steady state solution is $q_0 e^{i\omega t}$ where

$$q_0 = \frac{E_0}{2\omega_0 L} \frac{1}{(\omega_0 - \omega + i\gamma/2)}. \quad (1.7)$$

(We have made the usual resonance approximation: $\omega_0^2 - \omega^2 \approx 2\omega_0(\omega_0 - \omega)$.) The average power delivered to the circuit is

$$P = \frac{1}{2} \frac{E_0^2}{R} \frac{1}{1 + \left(\frac{\omega - \omega_0}{\gamma/2}\right)^2} \quad (1.8)$$

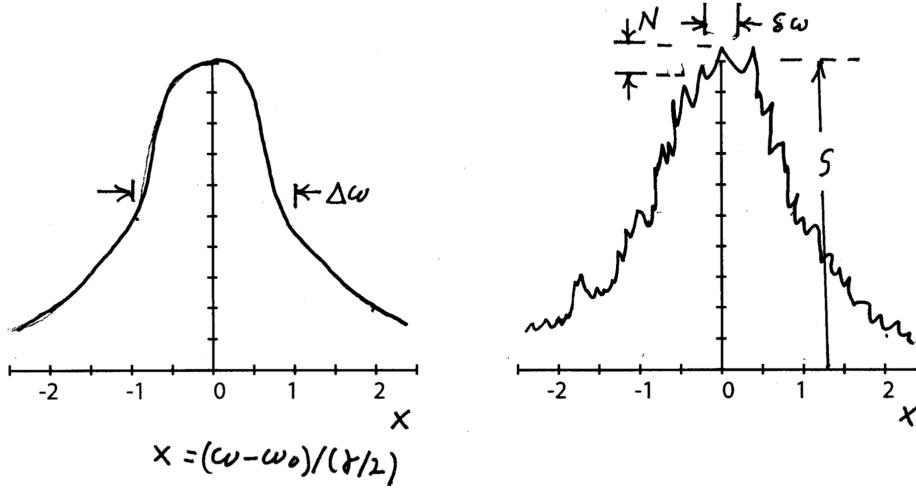


Figure 2. Sketch of a Lorentzian curve, the universal response curve for damped oscillators and for many atomic systems. The width of the curve (full width at half maximum) is $\Delta\omega = \gamma$, where γ is the decay constant. The time constant for decay is $\tau = \gamma$. In the presence of noise (right), the frequency precision with which the center can be located, $\delta\omega$, depends on the signal-to-noise ratio, S/N : $\delta\omega = \Delta\omega/(S/N)$.

The plot of P vs ω (Fig. 2) is universal resonance curve often called a “Lorentzian curve”. The full width at half maximum (“FWHM”) is $\Delta\omega = \gamma$. The quality factor of the oscillator is

$$Q = \frac{\omega_0}{\Delta\omega} \quad (1.9)$$

Note that the decay time of the free oscillator and the linewidth of the driven oscillator obey

$$\tau\Delta\omega = 1 \quad (1.10)$$

This can be regarded as an uncertainty relation. Assuming that energy and frequency are related by $E = \hbar\omega$ then the uncertainty in the energy is $\Delta E = \hbar\Delta\omega$ and

$$\tau\Delta E = \hbar \quad (1.11)$$

It is important to realize that the Uncertainty Principle merely characterizes the spread of individual measurements. Ultimate precision depends on the experimenter's skill: the Uncertainty Principle essentially sets the scale of difficulty for his or her efforts.

The precision of a resonance measurement is determined by how well one can “split” the resonance line. This depends on the signal to noise ratio (S/N)

(see Fig. 2). As a rule of thumb, the uncertainty $\delta\omega$ in the location of the center of the line is

$$\delta\omega = \frac{\Delta\omega}{S/N} \quad (1.12)$$

In principle, one can make $\delta\omega$ arbitrarily small by acquiring enough data to achieve the required statistical accuracy. In practice, systematic errors eventually limit the precision. Splitting a line by a factor of 10^4 is a formidable task which has only been achieved a few times, most notably in the measurement of the Lamb shift. A factor of 10^3 , however, is not uncommon, and 10^2 is child's play.

1.3 Magnetic Resonance: Classical Spin in Time-Varying B-Field

1.3.1 The classical motion of spins in a static magnetic field

Note: angular momentum will always be expressed in a form such as $\hbar\mathbf{J}$, where the vector \mathbf{J} is dimensionless. The interaction energy and equation of motion of a classical spin in a static magnetic field are given by

$$W = -\boldsymbol{\mu} \cdot \mathbf{B}, \quad (1.13)$$

$$\mathbf{F} = -\nabla W = \nabla(\boldsymbol{\mu} \cdot \mathbf{B}), \quad (1.14)$$

$$\text{torque} = \boldsymbol{\mu} \times \mathbf{B} \quad (1.15)$$

In a uniform field, $\mathbf{F} = 0$. The torque equation ($d\hbar\mathbf{J}/dt = \text{torque}$) gives

$$\frac{d\hbar\mathbf{J}}{dt} = \boldsymbol{\mu} \times \mathbf{B}. \quad (1.16)$$

Since $\boldsymbol{\mu} = \gamma\hbar\mathbf{J}$ (where γ is called the gyromagnetic ratio - not to be confused with the different meaning of γ in the previous section), we have

$$\frac{d\mathbf{J}}{dt} = \gamma\mathbf{J} \times \mathbf{B} = -\gamma\mathbf{B} \times \mathbf{J}. \quad (1.17)$$

To see that the motion of \mathbf{J} is a pure precession about \mathbf{B} , imagine that \mathbf{B} is along \hat{z} and that the spin, \mathbf{J} , is tipped at an angle θ from this axis, and then rotated at an angle $\phi(t)$ from the x -axis (i.e., θ and ϕ are the conventionally chosen angles in spherical coordinates). The torque, $-\gamma\mathbf{B} \times \mathbf{J}$, has no component along \mathbf{J} (that is, along \hat{r}), nor along $\hat{\theta}$ (because the $\mathbf{J} - \mathbf{B}$ plane contains $\hat{\theta}$), hence $-\gamma\mathbf{B} \times \mathbf{J} = -\gamma B|\mathbf{J}| \sin(\theta)\hat{\phi}$. This implies that \mathbf{J} maintains constant magnitude and constant tipping angle θ . Generally, for an infinitesimal change $d\phi$ the component of $d\mathbf{J}$ in the direction of $\hat{\phi}$ is $|\mathbf{J}| \sin(\theta)d\phi$, and so we can see that $\phi(t) = -\gamma B t$. This solution shows that the moment precesses with angular velocity

$$\Omega_L = -\gamma B \quad (1.18)$$

where Ω_L is called the *Larmor frequency*.

For electrons, $\gamma_e = -2\pi \times 2.8 \text{ MHz/G} \approx -2\mu_B$, for protons, $\gamma_p = 2\pi \times 4.2 \text{ kHz/G}$. Note that Planck's constant does not appear in the equation of motion: the motion is classical.

Note: G stands for gauss - it is part of the gaussian (cgs) system of units, and ubiquitous in atomic physics labs, as it is a much more typical laboratory field than the SI unit for magnetic field, 1 tesla. $10^4 \text{ G} = 1 \text{ T}$.

1.3.2 Rotating coordinate transformation

A second way to find the motion is to look at the problem in a rotating coordinate system. If some vector \mathbf{A} rotates with angular velocity $\boldsymbol{\Omega}$, then

$$\frac{d\mathbf{A}}{dt} = \boldsymbol{\Omega} \times \mathbf{A}. \quad (1.19)$$

If the rate of change of the vector in a system rotating at $\boldsymbol{\Omega}$ is $(d\mathbf{A}/dt)_{\text{rot}}$, then the rate of change in an inertial system is the motion *in* plus the motion *of* the rotating coordinate system.

$$\left(\frac{d\mathbf{A}}{dt} \right)_{\text{inert}} = \left(\frac{d\mathbf{A}}{dt} \right)_{\text{rot}} + \boldsymbol{\Omega} \times \mathbf{A}. \quad (1.20)$$

The operator prescription for transforming from an inertial to a rotating system is thus

$$\left(\frac{d \cdot}{dt} \right)_{\text{rot}} = \left(\frac{d \cdot}{dt} \right)_{\text{inert}} - \boldsymbol{\Omega} \times \cdot. \quad (1.21)$$

Applying this to Eq. 1.17 gives

$$\left(\frac{d\mathbf{J}}{dt} \right)_{\text{rot}} = \gamma \mathbf{J} \times \mathbf{B} - \boldsymbol{\Omega} \times \mathbf{J} = \gamma \mathbf{J} \times (\mathbf{B} + \boldsymbol{\Omega}/\gamma). \quad (1.22)$$

If we let

$$\mathbf{B}_{\text{eff}} = \mathbf{B} + \boldsymbol{\Omega}/\gamma, \quad (1.23)$$

Eq. 1.22 becomes

$$\left(\frac{d\mathbf{J}}{dt} \right)_{\text{rot}} = \gamma \mathbf{J} \times \mathbf{B}_{\text{eff}}. \quad (1.24)$$

If $\mathbf{B}_{\text{eff}} = 0$, \mathbf{J} is constant in the rotating system. The condition for this is

$$\boldsymbol{\Omega} = -\gamma \mathbf{B} \quad (1.25)$$

as we have previously found in Eq. 1.18.

1.3.3 Larmor's theorem

Treating the effects of a magnetic field on a magnetic moment by transforming to a rotating coordinate system is closely related to Larmor's theorem, which asserts that the effect of a magnetic field on a free charge can be eliminated by a suitable rotating coordinate transformation.

Consider the motion of a particle of mass m , charge q , under the influence

of an applied force \mathbf{F}_0 and the Lorentz force due to a static field \mathbf{B} :

$$\mathbf{F} = \mathbf{F}_0 + q\mathbf{v} \times \mathbf{B}. \quad (1.26)$$

Now consider the motion in a rotating coordinate system. By applying Eq. 1.21 twice to \mathbf{r} , we have

$$\ddot{\mathbf{r}}_{\text{rot}} = \ddot{\mathbf{r}}_{\text{inert}} - 2\boldsymbol{\Omega} \times \mathbf{v}_{\text{rot}} - \boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}). \quad (1.27)$$

$$\mathbf{F}_{\text{rot}} = \mathbf{F}_{\text{inert}} - 2m(\boldsymbol{\Omega} \times \mathbf{v}_{\text{rot}}) - m\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}), \quad (1.28)$$

where \mathbf{F}_{rot} is the apparent force in the rotating system, and $\mathbf{F}_{\text{inert}}$ is the true or inertial force. Substituting Eq. 1.26 gives

$$\mathbf{F}_{\text{rot}} = \mathbf{F}_{0,\text{inert}} + q\mathbf{v} \times \mathbf{B} + 2m\mathbf{v} \times \boldsymbol{\Omega} - m\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{r}). \quad (1.29)$$

If we choose $\boldsymbol{\Omega} = -(q/2m)\mathbf{B}$, and take $\mathbf{B} = \hat{z}B$, we have

$$\mathbf{F}_{\text{rot}} = \mathbf{F}_{0,\text{inert}} - m \left(\frac{qB}{2m} \right)^2 \hat{z} \times (\hat{z} \times \mathbf{r}). \quad (1.30)$$

The last term is usually small. If we drop it we have

$$\mathbf{F}_{\text{rot}} = \mathbf{F}_{0,\text{inert}} \quad (1.31)$$

The effect of the magnetic field is removed by going into a system rotating at the Larmor frequency $qB/2m$.

Although Larmor's theorem is suggestive of the rotating co-ordinate transformation, Eq. 1.22, it is important to realize that the two transformations, though identical in form, apply to fundamentally different systems. A magnetic moment is not necessarily charged- for example a neutral atom can have a net magnetic moment, and the neutron possesses a magnetic moment in spite of being neutral - and it experiences no net force in a uniform magnetic field. Furthermore, the rotating co-ordinate transformation is exact for a magnetic moment, whereas Larmor's theorem for the motion of a charged particle is only valid when the $\propto B^2$ term is neglected.

1.3.4 Motion in a Rotating Magnetic Field

Exact resonance: Consider a moment μ precessing about a static field \mathbf{B}_0 , which we take to lie along the z axis. Its motion might be described by

$$\mu_z = \mu \cos \theta, \quad \mu_x = \mu \sin \theta \cos \omega_0 t, \quad \mu_y = -\mu \sin \theta \sin \omega_0 t \quad (1.32)$$

where ω_0 is the Larmor frequency, and θ is the angle the moment makes with \mathbf{B}_0 .

Now suppose we introduce a magnetic field \mathbf{B}_1 which rotates in the x-y plane at the Larmor frequency $\omega_0 = -\gamma B_0$. The magnetic field is

$$\mathbf{B}(t) = B_1(\hat{x} \cos \omega_0 t - \hat{y} \sin \omega_0 t) + B_0 \hat{z}. \quad (1.33)$$

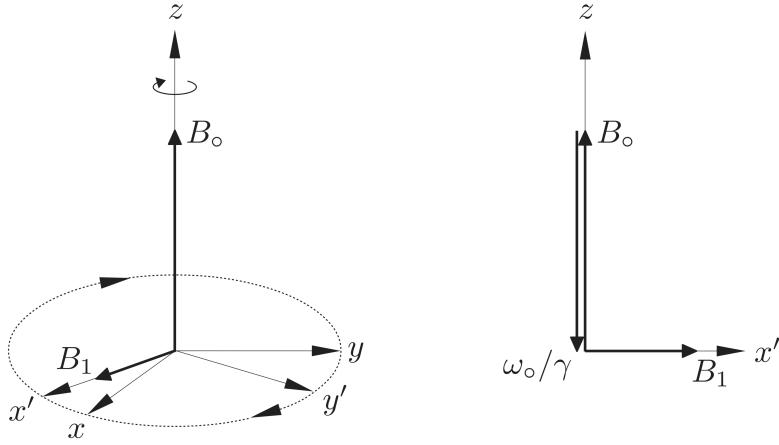


Figure 3. Rotating coordinate transformation to the primed system that is co-rotating with B_1 at ω , with x' chosen to lie along B_1 . For the exact resonance case of $\omega = \omega_0$ considered here, the effective field around which the moment precesses is equal to B_1 .

The problem is to find the motion of μ . The solution is simple in a rotating coordinate system (see Fig. 3). Let system $(\hat{x}', \hat{y}', \hat{z}' = \hat{z})$ precess around the z -axis at rate $-\omega_0$. In this system the field \mathbf{B}_1 is stationary (and \hat{x}' is chosen to lie along \mathbf{B}_1), and we have

$$\mathbf{B}_{\text{eff}}(t) = \mathbf{B}(t) - (\omega_0/\gamma) \hat{z} = B_1 \hat{x}' + (B_0 - \omega_0/\gamma) \hat{z} = B_1 \hat{x}'. \quad (1.34)$$

The effective field is static and has the value of B_1 . The moment precesses about the field at rate

$$\omega_R = \gamma B_1, \quad (1.35)$$

often called the *Rabi* frequency, in honor of Rabi's invention of the resonance technique.

If the moment initially lies along the z axis, then its tip traces a circle in the $\hat{y}' - \hat{z}$ plane. At time t it has precessed through an angle $\phi = \omega_R t$. The moment's z -component is given by

$$\mu_z(t) = \mu \cos \omega_R t. \quad (1.36)$$

At time $T = \pi/\omega_R$, the moment points along the negative z -axis: it has “turned over”.

Off-resonant behavior: Now suppose that the field B_1 rotates at frequency $\omega \neq \omega_0$. In a coordinate frame rotating with B_1 the effective field is

$$\mathbf{B}_{\text{eff}} = B_1 \hat{x}' + (B_0 - \omega/\gamma) \hat{z}. \quad (1.37)$$

The effective field lies at angle θ with the z -axis, as shown in Fig. 4. The field is static, and the moment precesses about it at rate (called the *effective*

Rabi frequency)

$$\Omega_R = \gamma B_{\text{eff}} = \gamma \sqrt{(B_0 - \omega/\gamma)^2 + B_1^2} = \sqrt{(\omega_0 - \omega)^2 + \omega_R^2} \quad (1.38)$$

where $\omega_0 = \gamma B_0$, $\omega_R = \gamma B_1$, as before.

Assume that μ points initially along the $+z$ -axis. Finding $\mu_z(t)$ is a straightforward problem in geometry. The moment precesses about B_{eff} at rate Ω_R , sweeping a circle as shown. The radius of the circle is $\mu \sin \theta$, where $\sin \theta = B_1 / \sqrt{(B_0 - \omega/\gamma)^2 + B_1^2} = \omega_R / \sqrt{(\omega - \omega_0)^2 + \omega_R^2}$. In time t the tip sweeps through angle $\phi = \Omega_R t$. The z -component of the moment is $\mu_z(t) = \mu \cos \alpha$ where α is the angle between the moment and the z -axis after it has precessed through angle ϕ . As the drawing shows, $\cos \alpha$ is found from $A^2 = 2\mu^2(1 - \cos \alpha)$. Since $A = 2\mu \sin \theta \sin(\Omega_R t/2)$, we have $4\mu^2 \sin^2 \theta \sin^2(\Omega_R t/2) = 2\mu^2(1 - \cos \alpha)$ and

$$\begin{aligned} \mu_z(t) &= \mu \cos \alpha = \mu(1 - 2 \sin^2 \theta \sin^2 \Omega_R t/2) \\ &= \mu \left[1 - 2 \frac{\omega_R^2}{(\omega - \omega_0)^2 + \omega_R^2} \sin^2 \frac{1}{2} \sqrt{(\omega - \omega_0)^2 + \Omega_R^2} t \right] \end{aligned} \quad (1.39)$$

$$= \mu \left[1 - 2(\omega_R/\Omega_R)^2 \sin^2(\Omega_R t/2) \right] \quad (1.40)$$

The z -component of μ oscillates in time, but unless $\omega = \omega_0$, the moment never completely inverts. The rate of oscillation depends on the magnitude of the rotating field; the amplitude of oscillation depends on the frequency difference, $\omega - \omega_0$, relative to ω_R . The quantum mechanical result will turn out to be identical.

1.3.5 Rapid Adiabatic Passage: Landau-Zener Crossing

Adiabatic rapid passage is a technique for inverting a spin population by sweeping the system through resonance. Either the frequency of the oscillating field or the transition frequency (e.g., by changing the applied magnetic field) is slowly varied. The principle is qualitatively simple in the rotating coordinate system.

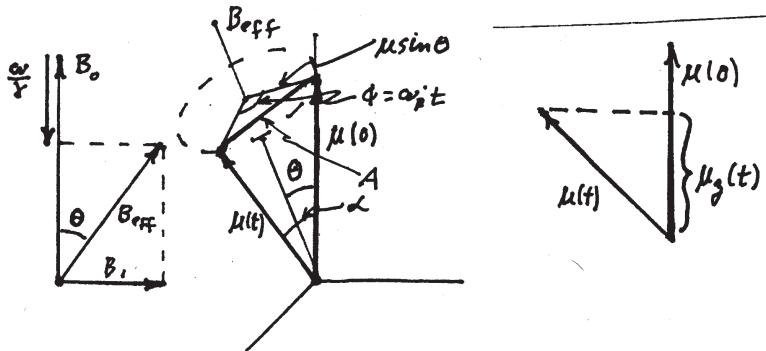


Figure 4. Constructions for viewing spin motion in a coordinate system rotating below the resonance frequency.

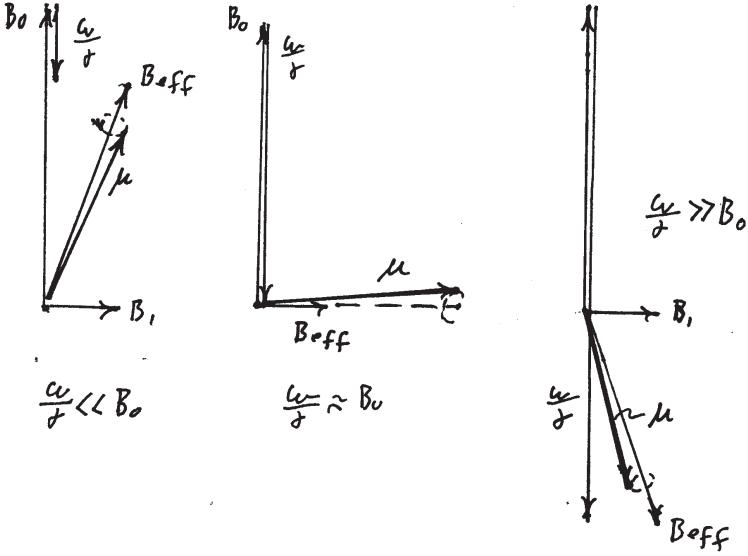


Figure 5. Motion of precessing moment in a rotating coordinate system whose frequency is swept from below resonance to above resonance.

The problem can also be solved analytically. In this section we give the qualitative argument and then, after having treated the quantum spin 1/2, we will present the analytic quantum result. The solution is of quite general interest because this physical situation arises frequently, for example in inelastic scattering, where it is called a curve crossing.

Consider a moment μ in the presence of a static magnetic field B_0 and a perpendicular field B_1 rotating at some frequency ω , originally far from resonance: $\omega \ll \gamma B_0$. In the frame rotating with B_1 the magnetic moment “sees” an effective field B_{eff} whose direction is nearly parallel to B_0 . A magnetic moment μ initially parallel to B_0 precesses around B_{eff} , making only a small angle with B_{eff} , as shown in Fig. 5.

If ω is *slowly* swept through resonance, μ will continue to precess tightly around B_{eff} , as shown in Figs. 5b,c. and will follow its direction adiabatically. In Fig. 5 the effective field now points in the $-\hat{z}$ direction, because $\omega \gg \gamma B_0$. Since the spin still precesses tightly around B_{eff} , its direction in the laboratory system has “flipped” from $+\hat{z}$ to $-\hat{z}$. The laboratory field B_0 remains unchanged, so this represents a transition from spin up to spin down.

The requirement for μ to follow the effective field $B_{\text{eff}}(t)$ is that, at all times, the magnetic moment always precesses tightly around B_{eff} . That means, within one precession period $2\pi/\Omega_R$, the angle θ that B_{eff} makes with \hat{z} must not have advanced more than a few degrees, i.e.

$$\Delta\theta = \dot{\theta} \cdot \Delta t = \dot{\theta} \frac{2\pi}{\Omega_R} \ll 2\pi \quad (1.41)$$

In other words, the Larmor frequency (the generalized Rabi frequency) $\Omega_R =$

γB_{eff} must be large compared to $\dot{\theta}$, the rate at which $\mathbf{B}_{\text{eff}}(t)$ is changing direction.

This requirement is most severe near exact resonance where $\theta = \pi/2$.

Using $B_{z,\text{eff}}(t) = B_0 - \omega(t)/\gamma$ we have in this case (from geometry)

$$|\dot{\theta}_{\max}| = \frac{1}{B_1} \frac{dB_{z,\text{eff}}(t)}{dt} = \frac{1}{B_1} \frac{1}{\gamma} \frac{d\omega}{dt} \ll \gamma B_1, \quad (1.42)$$

or using $\omega_R = \gamma B_1$,

$$\frac{d\omega}{dt} \ll \omega_R^2. \quad (1.43)$$

In this example we have shown that a slow change from $\omega \ll \gamma B_0$ to $\omega \gg \gamma B_0$ will flip the spin; the same argument shows that the reverse direction of slow change will also flip the spin.

Note that since the change of ω in one Rabi period must be much smaller than the Rabi frequency itself, and since we need to sweep the detuning $\omega - \omega_0$ from much larger than ω_R to much smaller than $-\omega_R$ to complete the inversion, it is clear that the inversion will be slower than an on-resonance π -pulse.

In the examples below (Fig. 6), we see how the magnetic moment follows the changing magnetic field \mathbf{B}_{eff} less and less faithfully as the ramp speed $\dot{\omega}$ is increased beyond ω_R^2 . We will come back to the Landau-Zener problem after having introduced the quantum spin 1/2 case.

1.3.6 Adiabatic Following in a Magnetic Trap

In a slight digression, an important example of adiabatic following of magnetic moments is given by the trapping of neutral atoms in magnetic traps. First, since Maxwell's equations forbid the existence of a magnetic field maximum in free space (there are no magnetic charges), we can only create magnetic field minima in free space, and therefore only trap atoms in so-called "low-field seeking" states, so for which the projection of the magnetic moment on the local magnetic field direction points opposite the field. For the atom to stay trapped, the magnetic moment must always be able to follow adiabatically the direction of the local magnetic field. Otherwise, "spin-flips", whereby the atoms flip their magnetic moments to be aligned with the field, lowering their energy and then leaving the trap (by seeking the high field right at the coils). A classic example is the quadrupole magnetic trap, formed by two parallel coils with currents running in anti-Helmholtz configuration. This results in a magnetic field zero right in the center between the coils, and a gradient field away from the zero:

$$\mathbf{B} = B' \begin{pmatrix} x \\ y \\ -2z \end{pmatrix}.$$

If an atom moves through the region close to the zero of magnetic field, the local Larmor frequency may become low compared to the rate of change of the field direction. For adiabatic following, the rate of change of the field angle should satisfy:

$$\dot{\theta} \ll \Omega_L$$

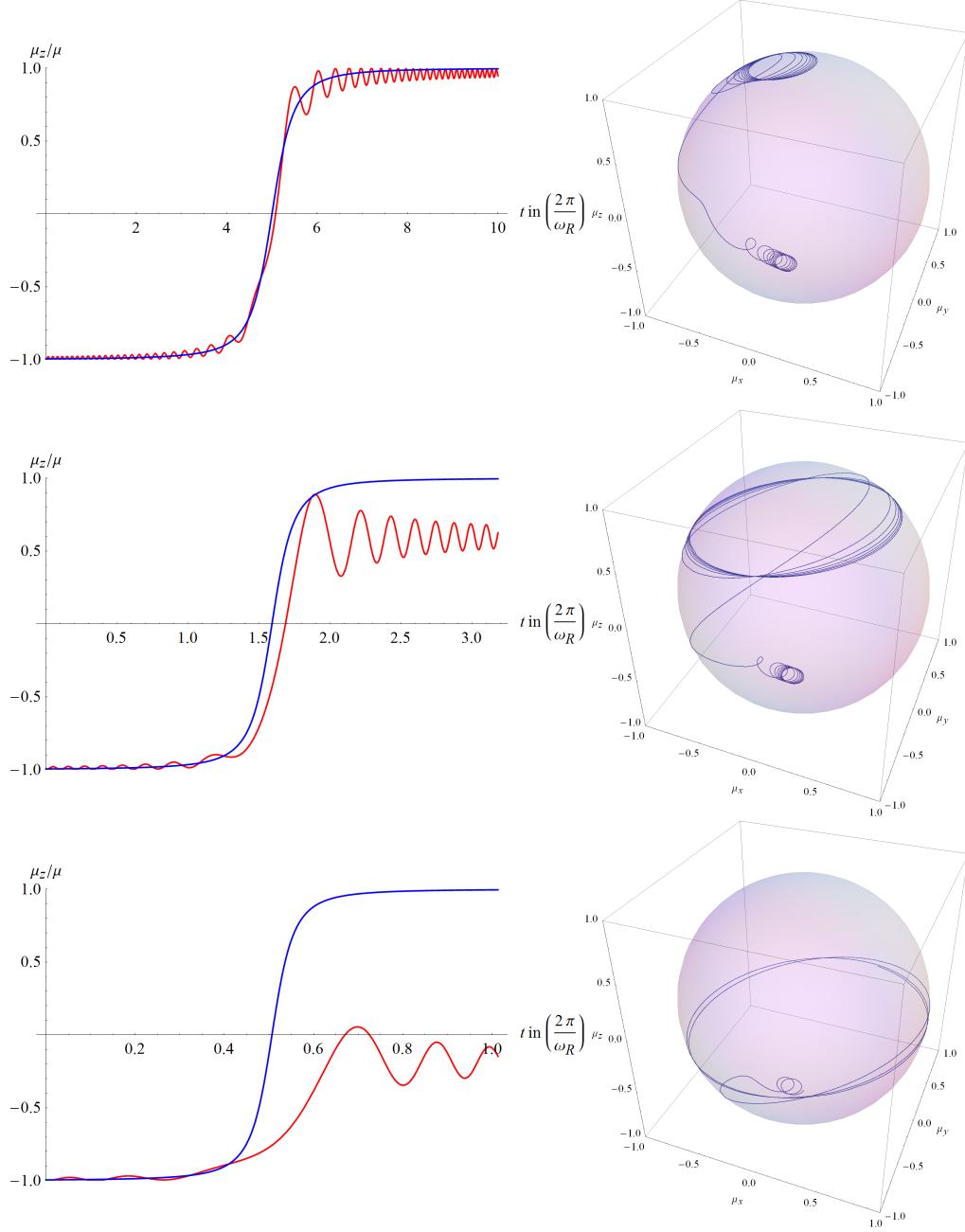


Figure 6. Top row: Adiabatic sweep. Start of sweep $\delta = -10\omega_R$, end of sweep $\delta = +10\omega_R$, sweep rate $\dot{\omega} = \frac{20\omega_R}{10\frac{2\pi}{\omega_R}} = \frac{1}{\pi}\omega_R^2$. Left: z-component of the magnetic moment versus time t measured in Rabi periods $2\pi/\omega_R$ (red curve). Blue is $\cos\theta = \frac{\delta(t)}{\sqrt{\delta(t)^2+\omega_R^2}}$, i.e. the cosine of the angle that \mathbf{B}_{eff} makes with the z-axis. Right: Path traced out by the tip of the magnetic moment in real space (or, for a two-level system in quantum mechanics described by a pseudo-spin 1/2: the path traced out by the tip of the Bloch vector on the Bloch sphere). Middle row: sweep with imperfect adiabaticity. The sweep rate here is $\dot{\omega} = \pi\frac{20\omega_R}{10\frac{2\pi}{\omega_R}} = \omega_R^2$. Bottom row: non-adiabatic sweep. Sweep rate $\dot{\omega} = \pi^2\frac{20\omega_R}{10\frac{2\pi}{\omega_R}} = \pi\omega_R^2$.

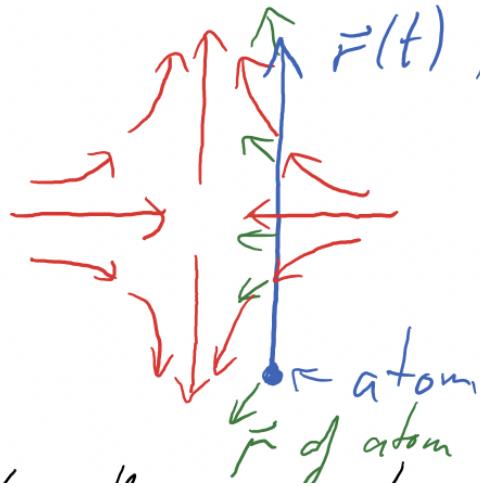


Figure 7. An atom moving in the quadrupole magnetic trap, with its magnetic moment attempting to follow the direction of the local magnetic field.

The field experienced by the atom changes as the atom is moving through space. We have (approximately, i.e. not worrying about factors on the order of two):

$$\dot{\theta} = \frac{\dot{B}}{B} = \frac{B'v}{B} = \frac{B}{r}.$$

Here, v is the atom velocity and r is the atom position. So, the adiabaticity condition is

$$\dot{\theta} \ll \Omega_L \implies \frac{v}{r} \ll \Omega_L = \gamma B = \gamma B' r.$$

So the “danger zone” where the atom may no longer be able to follow adiabatically, and therefore undergo spin flips, is the region of distances from the field zero smaller than

$$r = \sqrt{\frac{v}{\gamma B'}}.$$

This region is called the “Majorana hole”. As the atom cloud is cooled (e.g. through radiofrequency evaporation), the velocity decreases like \sqrt{T} , the hole radius like $T^{1/4}$, but the cloud size $R \propto T/\mu B'$ decreases faster. So eventually, the cloud becomes comparable in size to the “Majorana hole” and severe losses in the number of trapped atoms are observed.

There are several solutions to this that were developed in order to keep atoms in a magnetic trap. Wolfgang Ketterle at MIT used a repulsive laser beam to “plug the hole”. This is called the “plug trap”. Eric Cornell at JILA added a time varying magnetic bias field that moves the magnetic field zero in a circular orbit of radius larger than the cloud size, and at a rate much faster than the atoms can follow (typically: 10 kHz, compared to the frequency of atomic motion of 100 Hz), but still much slower than the Larmor frequency (1 MHz), so that atoms stay trapped in a region of non-zero magnetic field (1 G) where their moments can adiabatically follow the rotating bias field. This was called the TOP (time orbiting potential) trap. These two completely different solutions

both led to the realization of Bose-Einstein condensation.

1.4 Resonance Of Quantized Spin

1.4.1 Expectation value of magnetic moment behaves classically

Before solving the quantum mechanical problem of a magnetic moment in a time varying field, it is worthwhile demonstrating that its motion is classical. By “its motion is classical” we mean the time evolution of the expectation value of the magnetic moment operator obeys the classical equation of motion. Specifically, we shall show that

$$\frac{d}{dt}\langle\hat{\mu}\rangle = \gamma\langle\hat{\mu}\rangle \times \mathbf{B}. \quad (1.44)$$

Proof: Recall the Heisenberg equations of motion for an operator \hat{O} :

$$\frac{d}{dt}\hat{O} = \frac{i}{\hbar}[\hat{H}, \hat{O}] + \frac{\partial\hat{O}}{\partial t}. \quad (1.45)$$

If the operator is not explicitly time dependent the last term vanishes.

The interaction of $\hat{\mu}$ with a static field $B_0\hat{z}$ is

$$\hat{H} = -\hat{\mu} \cdot \mathbf{B}_0 = -\gamma\hat{\mathbf{J}} \cdot \mathbf{B}_0 = -\gamma B_0\hat{J}_z, \quad (1.46)$$

Note that $\hat{\mathbf{J}}$ has dimensions of angular momentum. Thus

$$\frac{d\hat{\mu}}{dt} = -i\gamma B_0[\hat{J}_z, \hat{\mu}]/\hbar. \quad (1.47)$$

Using $\hat{\mu} = \gamma\hat{\mathbf{J}}$, we can rewrite this as

$$\frac{d\hat{\mathbf{J}}}{dt} = -i\gamma B_0[\hat{J}_z, \hat{\mathbf{J}}]/\hbar. \quad (1.48)$$

The commutation rules for $\hat{\mathbf{J}}$ are $[\hat{J}_x, \hat{J}_y] = i\hbar\hat{J}_z$, etc., or $\hat{\mathbf{J}} \times \hat{\mathbf{J}} = i\hbar\hat{\mathbf{J}}$ (this is a shorthand way of writing $[\hat{J}_i, \hat{J}_j] = \epsilon_{ijk}\hat{J}_k$.) Hence

$$\frac{d\hat{J}_x}{dt} = \gamma B_0\hat{J}_y \quad (1.49)$$

$$\frac{d\hat{J}_y}{dt} = -\gamma B_0\hat{J}_x \quad (1.50)$$

$$\frac{d\hat{J}_z}{dt} = 0 \quad (1.51)$$

These describe the uniform precession of $\hat{\mathbf{J}}$ about the z -axis at a rate $-\gamma B_0$. In particular, taking the expectation value, we have:

$$\frac{d}{dt}\langle\hat{\mathbf{J}}\rangle = \gamma\langle\hat{\mathbf{J}}\rangle \times \mathbf{B} \quad (1.52)$$

and since $\hat{\mu} = \gamma \hat{\mathbf{J}}$, this directly yields Eq. 1.44:

$$\frac{d}{dt} \langle \hat{\mu} \rangle = \gamma \langle \hat{\mu} \rangle \times \mathbf{B}. \quad (1.53)$$

Thus the quantum mechanical and classical equation of motion are identical. This fact underlies the great utility of classical magnetic resonance in providing intuition about resonance in quantum spin systems.

A few comments on this result:

- the result is valid for any angular momentum operator, so also for spin 1/2...
- ... and therefore for any two-level system that can be mapped onto an effective spin 1/2.
- it is valid for the case of several angular momenta within an atom coupled to a total angular momentum \mathbf{F} (as long as the magnetic field \mathbf{B} is not large enough to “break” the coupling).
- it is valid also for a system of N two-level systems symmetrically coupled to an external field. In this case, we have an effective angular momentum $L = N/2$. Spin precession in this case can be understood as Dicke superradiance, the constructive interference of N “aligned” particles.

1.4.2 The Rabi transition probability

For a spin 1/2 particle we can push the classical solution further and obtain the amplitudes and probabilities for each state. Consider ¹ $\langle \mu_z \rangle / \hbar = \gamma \langle J_z \rangle = \gamma m$, where m is the usual “magnetic” quantum number. For a spin 1/2 particle m has the value +1/2 or -1/2. Let the probabilities for having these values be P_+ and P_- respectively. Then

$$\langle J_z \rangle = \frac{1}{2} P_+ - \frac{1}{2} P_-, \quad (1.54)$$

or, since $P_+ + P_- = 1$,

$$\langle J_z \rangle = \frac{1}{2}(1 - 2P_-), \quad (1.55)$$

$$\langle \mu_z \rangle = \frac{1}{2}\gamma\hbar(1 - 2P_-). \quad (1.56)$$

If μ lies along the z axis at $t = 0$, then $\mu_z(0) = \gamma\hbar/2$, and we have

$$\mu_z(t) = \mu_z(0)(1 - 2P_-). \quad (1.57)$$

In this case, P_- is the probability that a spin in state $m = +1/2$ at $t = 0$ has made a transition to $m = -1/2$ at time t , $P_{\uparrow \rightarrow \downarrow}(t)$. So for the Rabi problem, the case of a static field $\mathbf{B}_0 = \omega_0 \hat{z}/\gamma$ and a rotating magnetic field \mathbf{B}_1 rotating

¹We will in the following drop hats $\hat{\cdot}$ over operators for simplicity.

at frequency ω about the z-axis, switched on at time $t = 0$, we see immediately, comparing Eq. 1.57 with 1.39,

$$P_{\uparrow \rightarrow \downarrow}(t) = \frac{\omega_R^2}{\omega_R^2 + (\omega - \omega_0)^2} \sin^2 \left(\frac{1}{2} \sqrt{\omega_R^2 + (\omega - \omega_0)^2} t \right) \quad (1.58)$$

$$P_{\uparrow \rightarrow \downarrow}(t) = (\omega_R/\Omega_R)^2 \sin^2(\Omega_R t/2) \quad (1.59)$$

This result is known as the *Rabi transition probability*. It is important enough to memorize. We have derived it from a classical correspondence argument, but it can also be derived quantum mechanically. In fact, such a treatment is essential for a complete understanding of the system.

1.4.3 The Hamiltonian of a quantized spin $\frac{1}{2}$

Let us investigate the time dependence of the wave function for a quantized spin $\frac{1}{2}$ system with moment $\boldsymbol{\mu} = \gamma \hbar \mathbf{S}$ that is placed in a uniform magnetic field $\mathbf{B}_0 = \omega_0 \hat{\mathbf{z}}/\gamma$ and, starting at $t = 0$, subject to a field $\mathbf{B}_R(t)$ which rotates in the $x-y$ plane with frequency ω . These fields are the same as the fields discussed in the preceding section on the motion of a classical spin and a time-varying field. The only difference is that now we are discussing their effect on a quantized system, so we must use Schrödinger's equation rather than the laws of classical Electricity and Magnetism to discuss the dynamics of the system.

The basis states are (using the standard column vector representation):

$$|g\rangle \equiv |1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad (1.60)$$

$$|e\rangle \equiv |2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (1.61)$$

with state $|1\rangle$ lower in energy.

The unperturbed Hamiltonian is (note: we will take $\omega_0 = \gamma B_0$ to be positive, and $\boldsymbol{\mu}$ is an operator)

$$\begin{aligned} H_0 &= -\boldsymbol{\mu} \cdot \mathbf{B}_0 = -\hbar S_z \omega_0 \\ &= -\frac{1}{2} \hbar \omega_0 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = -\frac{1}{2} \hbar \omega_0 \sigma_z \end{aligned} \quad (1.62)$$

where σ_z is a Pauli spin matrix.

The energies are

$$\begin{aligned} E_1 &= -\hbar \omega_0/2; & \omega_1 &= -\omega_0/2 \\ E_2 &= +\hbar \omega_0/2; & \omega_2 &= +\omega_0/2 \end{aligned} \quad (1.63)$$

Suppose we have a spin initially aligned along \hat{x} :

$$|\psi(t=0)\rangle = \frac{1}{\sqrt{2}}(|g\rangle + |e\rangle).$$

Then

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}}\left(e^{i\omega_0 t/2}|g\rangle + e^{-i\omega_0 t/2}|e\rangle\right) = \frac{1}{\sqrt{2}}e^{i\omega_0 t/2}\left(|g\rangle + e^{-i\omega_0 t}|e\rangle\right).$$

This corresponds to rotation about the equator of the Bloch sphere at a rate ω_0 : the state of the spin 1/2 rotates from being aligned along \hat{x} to being aligned along $-\hat{y}$, then $-\hat{x}$, then \hat{y} etc., so in the same sense as the classical magnetic moment.

We will now switch on the magnetic field \mathbf{B}_R which rotates in the $x - y$ plane, with the goal of making it co-rotate with the spin when driven at $\omega = \omega_0$.

The perturbation Hamiltonian $H'(t)$ is written best in terms of $\omega_R = \gamma B_R$ where B_R is the magnetic field which rotates in the $x - y$ plane (in the magnetic resonance community the subscript R is often replaced by 1):

$$\begin{aligned} H'(t) &= -\boldsymbol{\mu} \cdot \mathbf{B}_R(t) \\ &= -\boldsymbol{\mu} \cdot (\omega_R/\gamma)[\hat{x} \cos \omega t - \hat{y} \sin \omega t] \\ &= -\omega_R[S_x \cos \omega t - S_y \sin \omega t] \\ &= -\frac{\hbar\omega_R}{2} \left[\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cos \omega t - \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \sin \omega t \right] \\ &= -\frac{\hbar\omega_R}{2} \begin{pmatrix} 0 & e^{i\omega t} \\ e^{-i\omega t} & 0 \end{pmatrix} \end{aligned} \quad (1.64)$$

where we used $\boldsymbol{\mu} \cdot \hat{x} = \gamma S_x$. In the penultimate line we have replaced the operators S_x and S_y with $(\hbar/2) \sigma_x$ and $(\hbar/2) \sigma_y$ where σ_x and σ_y are Pauli spin matrices. The perturbation matrix element is just the entry $H'_{21}(t)$ in row 2 and column 1:

$$\langle 2 | H' | 1 \rangle = -\frac{\hbar\omega_R}{2} e^{-i\omega t} \quad (1.65)$$

We have thus derived the following Hamiltonian for the spin $\frac{1}{2}$ problem:

$$H = \frac{\hbar}{2} \begin{pmatrix} -\omega_0 & -\omega_R e^{+i\omega t} \\ -\omega_R e^{-i\omega t} & \omega_0 \end{pmatrix} \quad (1.66)$$

1.4.4 Quantum Mechanical Solution for Resonance in a Two-State System

As has been emphasized, a two-state system coupled by a periodic interaction is an archetype for large areas of atomic/optical physics. The quantum mechanical solution can be achieved by a variety of approaches, the most elegant of which is the dressed atom picture in which the atom and radiation field constitute a single quantum system and one finds its eigenstates. That approach will be introduced later. Here we follow a rather different approach, less elegant, but capable of being generalized to a variety of problems including multi-level resonance and

radiative decay in the presence of oscillating fields. The starting point is the interaction representation.

1.4.5 Interaction representation

We consider a complete set of eigenstates to a Hamiltonian H_0 , $\psi = |1\rangle, |2\rangle, \dots$, such that

$$H_0 |j\rangle = E_j |j\rangle. \quad (1.67)$$

The problem is to find the behavior of the system under an interaction $V(t)$, i.e. to find solutions to

$$i\hbar \frac{\partial \psi(t)}{\partial t} = (H_0 + V(t))\psi(t) \quad (1.68)$$

In the interaction representation we take

$$\psi(t) = \sum_j a_j(t) |j\rangle e^{-iE_j t/\hbar}. \quad (1.69)$$

Schrödinger's equation yields

$$i\hbar \dot{a}_k = \sum_j \langle k | V | j \rangle a_j e^{i(E_j - E_k)t/\hbar} = \sum_j V_{kj} a_j e^{i\omega_{jk} t}. \quad (1.70)$$

Given any set of initial conditions, $a_j(0), j = 1, 2, 3, \dots$, these equations can be integrated to find $\psi(t)$. Often, this is done iteratively, following a perturbative approach. For the two-state system, with V periodic, one can obtain an exact solution.

1.4.6 Two-state problem

We consider a two-state system

$$\psi = a_1 |1\rangle + a_2 |2\rangle, \quad (1.71)$$

with $|a_1|^2 + |a_2|^2 = 1$. We assume $E_2 > E_1$, and introduce $\hbar\omega_{12} = \hbar\omega_0 = E_2 - E_1$. Without loss of generality, we let $E_1 = -\hbar\omega_0/2$; $E_2 = \hbar\omega_0/2$. We take the interaction to be of the form $V_{11} = V_{22} = 0$, and

$$V_{12} = \frac{1}{2}\hbar\omega_R (e^{-i\omega t}), \quad (1.72)$$

Eq. 1.70 gives

$$\begin{aligned} i\dot{a}_1 &= \frac{1}{2}\omega_R (e^{-i\omega t}) e^{+i\omega_0 t} a_2, \\ i\dot{a}_2 &= \frac{1}{2}\omega_R (e^{i\omega t}) e^{-i\omega_0 t} a_1. \end{aligned} \quad (1.73)$$

Introducing $\delta \equiv \omega - \omega_0$, Eqs. 1.73 becomes

$$\begin{aligned} i\dot{a}_1 &= \frac{1}{2}\omega_R e^{-i\delta t} a_2, \\ i\dot{a}_2 &= \frac{1}{2}\omega_R e^{i\delta t} a_1. \end{aligned} \quad (1.74)$$

We can eliminate the explicit time dependence by making the substitution

$$\begin{aligned} a_1 &= e^{-i\delta t/2} b_1, \\ a_2 &= e^{+i\delta t/2} b_2. \end{aligned} \quad (1.75)$$

Eqs. 1.74 become

$$\begin{aligned} \dot{b}_1 - i\frac{\delta}{2}b_1 &= \frac{-i}{2}\omega_R b_2, \\ \dot{b}_2 + i\frac{\delta}{2}b_2 &= \frac{-i}{2}\omega_R b_1. \end{aligned} \quad (1.76)$$

These equations describe periodic behavior, so that it is natural to try solutions of the form

$$\begin{aligned} b_1 &= \sum_j B_j e^{i\alpha_j t}, \\ b_2 &= \sum_j C_j e^{i\alpha_j t}. \end{aligned} \quad (1.77)$$

Substituting these in Eq. 1.76 yields

$$\begin{aligned} \left(\alpha_j - \frac{\delta}{2}\right)B_j + \left(\frac{\omega_R}{2}\right)C_j &= 0, \\ \left(\frac{\omega_R}{2}\right)B_j + \left(\alpha_j + \frac{\delta}{2}\right)C_j &= 0. \end{aligned} \quad (1.78)$$

for which the determinental equation yields two eigenfrequencies

$$\alpha_{1,2} = \pm \frac{1}{2}\sqrt{\omega_R^2 + \delta^2} = \pm \frac{1}{2}\Omega_R, \quad (1.79)$$

where Ω_R is the generalized Rabi frequency:

$$\Omega_R = \sqrt{\omega_R^2 + \delta^2}. \quad (1.80)$$

From Eqs. 1.78:

$$C_1 = -B_1 \frac{\omega_R}{\Omega_R + \delta}, \quad C_2 = B_2 \frac{\omega_R}{\Omega_R - \delta}. \quad (1.81)$$

By combining this result with Eq. 1.76 and Eq. 1.77 we obtain

$$\begin{aligned} a_1(t) &= e^{-i(\delta-\Omega_R)t/2}B_1 + e^{-i(\delta+\Omega_R)t/2}B_2 \\ a_2(t) &= -\left(\frac{\omega_R}{\Omega_R + \delta}\right)e^{i(\delta+\Omega_R)t/2}B_1 + \left(\frac{\omega_R}{\Omega_R - \delta}\right)e^{i(\delta-\Omega_R)t/2}B_2 \end{aligned} \quad (1.82)$$

The solution contains two arbitrary constants, B_1 and B_2 , which permit fitting the boundary condition for the two amplitudes.

If the system is in state 1 at $t = 0$, then $a_1(1) = 1, a_2(0) = 0$, and

$$\begin{aligned} B_1 &= \frac{1}{2} \frac{\Omega_R + \delta}{\Omega_R}, \quad B_2 = \frac{1}{2} \frac{\Omega_R - \delta}{\Omega_R} \\ a_2(t) &= \frac{1}{2} \frac{\omega_R}{\Omega_R} \left[-e^{i(\delta+\Omega_R)t/2} + e^{i(\delta-\Omega_R)t/2} \right] \\ &= -i \frac{\omega_R}{\Omega_R} e^{i\delta t/2} \sin(\Omega_R t/2). \end{aligned} \quad (1.83)$$

The probability of being in state 2 at time t is,

$$P_2(t) = |a_2(t)|^2 = \frac{\omega_R^2}{\Omega_R^2} \sin^2 \left[\frac{1}{2} \Omega_R t \right], \quad (1.84)$$

which is identical to the classical result for the Rabi resonance formula, derived earlier.

If we introduce the parameter θ defined by

$$\begin{aligned} \cos \theta &= \left(\frac{\Omega_R + \delta}{2\Omega_R} \right)^{1/2} = \left(\frac{1}{2} \left(1 + \frac{\delta}{\Omega_R} \right) \right)^{1/2} \\ \sin \theta &= \left(\frac{\Omega_R - \delta}{2\Omega_R} \right)^{1/2} = \left(\frac{1}{2} \left(1 - \frac{\delta}{\Omega_R} \right) \right)^{1/2} \end{aligned} \quad (1.85)$$

Then we have

$$\psi(t) = a_1(t) |1\rangle + a_2(t) |2\rangle \quad (1.86)$$

$$\begin{aligned} &= [\cos^2 \theta e^{-i(\delta-\Omega_R)t/2} + \sin^2 \theta e^{-i(\delta+\Omega_R)t/2}] |1\rangle \\ &\quad + [\cos \theta \sin \theta] [e^{i(\delta+\Omega_R)t/2} - e^{i(\delta-\Omega_R)t/2}] |2\rangle \\ &= e^{-i\delta t/2} \left[\cos \left(\frac{\Omega_R t}{2} \right) + i \frac{\delta}{\Omega_R} \sin \left(\frac{\Omega_R t}{2} \right) \right] |1\rangle \\ &\quad - i e^{+i\delta t/2} \frac{\omega_R}{\Omega_R} \sin \left(\frac{\Omega_R t}{2} \right) |2\rangle \end{aligned} \quad (1.87)$$

1.4.7 Solution via rotating frame

The preceding solution seems a bit “brute force” and required some ad hoc “tricks” like the substitution from the a to the b amplitudes. We will here see very clearly what is behind the mathematics: It is again frame rotation.

In the very same spirit of the classical solution, we want to rotate our spin 1/2 into a frame in which the magnetic fields appear stationary. In this frame the Hamiltonian will be time independent and can be solved easily. Now just like the classical angular momentum L_z generated rotation of vectors about the z-axis, the spin angular momentum operator S_z generates rotation of spinors about the z-axis. The rotation operator for rotation about \hat{z} by an angle θ is thus

$$T = e^{-i\hat{S}_z\theta}.$$

We can find this result from integration of infinitesimal rotations.

To keep aligned with the rotating magnetic field, we want $\theta = -\omega t$, then we have

$$T = e^{i\omega t \sigma_z/2} = \begin{pmatrix} e^{i\omega t/2} & 0 \\ 0 & e^{-i\omega t/2} \end{pmatrix}.$$

Our states transform by

$$|\psi\rangle = T |\tilde{\psi}\rangle.$$

We get a new term in the Schrödinger equation according to

$$i\hbar \frac{d}{dt} |\psi\rangle = i\hbar \dot{T} |\tilde{\psi}\rangle + T i\hbar \frac{d}{dt} |\tilde{\psi}\rangle = H |\psi\rangle = HT |\tilde{\psi}\rangle.$$

So,

$$i\hbar \frac{d}{dt} |\tilde{\psi}\rangle = (T^\dagger HT - i\hbar T^\dagger \dot{T}) |\tilde{\psi}\rangle.$$

The Hamiltonian in the rotating frame is then

$$\tilde{H} = T^\dagger \left(H - i\hbar \frac{d}{dt} \right) T.$$

H_0 and T are both diagonal, so

$$T^\dagger H_0 T = H_0.$$

For H_1 , T removes the time-dependence:

$$T^\dagger H_1 T = -\frac{\hbar\omega_R}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -\frac{\hbar\omega_R}{2} \sigma_x.$$

This is what we set out to do, since we chose the frame to rotate with the field. The time derivative term is

$$T^\dagger \left(-i\hbar \frac{d}{dt} T \right) = \frac{\hbar\omega}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \frac{\hbar\omega}{2} \sigma_z.$$

We thus see the \hat{z} field effectively reduced in amplitude, as occurred for our classical system. The new Hamiltonian is thus

$$\tilde{H} = -\frac{\hbar}{2} \begin{pmatrix} -\delta & \omega_R \\ \omega_R & \delta \end{pmatrix} \text{ with } \delta = \omega - \omega_0.$$

We can also write

$$\tilde{H} = \frac{\hbar\delta}{2}\sigma_z - \frac{\hbar\omega_R}{2}\sigma_x = \frac{1}{2}\mathbf{h} \cdot \boldsymbol{\sigma} \text{ for } \mathbf{h} = \begin{pmatrix} -\hbar\omega_R \\ 0 \\ \hbar\delta \end{pmatrix}.$$

The Hamiltonian is now time-independent, and represents the interaction of a spin 1/2 with a static magnetic field. The eigenvectors are thus simply the two states representing the spin aligned and anti-aligned with \mathbf{h} (+ and -). The eigenvalues are

$$E_{\pm} = \pm\frac{1}{2}|\mathbf{h}| = \pm\frac{\hbar}{2}\sqrt{\omega_R^2 + \delta^2} = \pm\frac{\hbar\Omega_R}{2}.$$

We can obtain the eigenvectors from the original eigenvectors $|g\rangle \equiv |\uparrow\rangle$ and $|e\rangle \equiv |\downarrow\rangle$ which described a spin 1/2 aligned and anti-aligned with the z-axis, simply by rotating about the axis

$$\hat{\mathbf{z}} \times \mathbf{h} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \times \begin{pmatrix} -\hbar\omega_R \\ 0 \\ \hbar\delta \end{pmatrix} = -\begin{pmatrix} 0 \\ \hbar\omega_R \\ 0 \end{pmatrix} = \hbar\omega_R \hat{\mathbf{y}}$$

by the angle ϕ satisfying

$$\cos\phi = \hat{\mathbf{z}} \cdot \hat{\mathbf{h}} = \frac{\delta}{\Omega_R}.$$

Note that

$$\tan\phi = -\frac{\omega_R}{\delta}.$$

Rotation about the y-axis is generated by S_y , again through the exponential

$$R = e^{-i\phi S_y} = e^{-i\phi\sigma_y/2}.$$

This is equivalent to

$$R = \cos\frac{\phi}{2}\mathbb{I} - i\sin\frac{\phi}{2}\sigma_y = \begin{pmatrix} \cos\frac{\phi}{2} & -\sin\frac{\phi}{2} \\ \sin\frac{\phi}{2} & \cos\frac{\phi}{2} \end{pmatrix}.$$

We can thus find the eigenstates, properly aligned and anti-aligned with the direction of \mathbf{h} :

$$|+\rangle_h = R|\uparrow\rangle = \begin{pmatrix} \cos\phi/2 \\ \sin\phi/2 \end{pmatrix} = \cos(\phi/2)|\uparrow\rangle + \sin(\phi/2)|\downarrow\rangle \quad (1.88)$$

$$|-\rangle_h = R|\downarrow\rangle = \begin{pmatrix} -\sin\phi/2 \\ \cos\phi/2 \end{pmatrix} = -\sin(\phi/2)|\uparrow\rangle + \cos(\phi/2)|\downarrow\rangle \quad (1.89)$$

Note that $\cos(\phi/2) = \sqrt{\frac{1+\cos\phi}{2}} = \sqrt{\frac{\Omega_R+\delta}{2\Omega_R}}$ and $\sin(\phi/2) = \sqrt{\frac{1-\cos\phi}{2}} = \sqrt{\frac{\Omega_R-\delta}{2\Omega_R}}$.

Their time-evolution is simply, since $E_+ = -\frac{\hbar\Omega_R}{2}$ and $E_- = \frac{\hbar\Omega_R}{2}$:

$$|+(t)\rangle = e^{i\Omega_R t/2} |+\rangle_h \quad (1.90)$$

$$|-(t)\rangle = e^{-i\Omega_R t/2} |-\rangle_h \quad (1.91)$$

Now say we start with $|\psi(t=0)\rangle = |\uparrow\rangle$, that is

$$|\psi(t=0)\rangle = R^T |+\rangle = \cos(\phi/2) |+\rangle - \sin(\phi/2) |-\rangle$$

We immediately obtain the time evolution

$$\begin{aligned} |\psi(t)\rangle &= \cos \frac{\phi}{2} e^{+i\Omega_R t/2} |+\rangle - \sin \frac{\phi}{2} e^{-i\Omega_R t/2} |-\rangle \\ &= \cos \frac{\phi}{2} e^{+i\Omega_R t/2} (\cos(\phi/2) |\uparrow\rangle + \sin(\phi/2) |\downarrow\rangle) \\ &\quad - \sin \frac{\phi}{2} e^{-i\Omega_R t/2} (-\sin(\phi/2) |\uparrow\rangle + \cos(\phi/2) |\downarrow\rangle) \\ &= \cos\left(\frac{\Omega_R t}{2}\right) |\uparrow\rangle + i \sin\left(\frac{\Omega_R t}{2}\right) \cos\phi |\uparrow\rangle + i \sin\left(\frac{\Omega_R t}{2}\right) \sin\phi |\downarrow\rangle \\ &= \left(\cos\left(\frac{\Omega_R t}{2}\right) + i \frac{\delta}{\Omega_R} \sin\left(\frac{\Omega_R t}{2}\right) \right) |\uparrow\rangle - i \frac{\omega_R}{\Omega_R} \sin\left(\frac{\Omega_R t}{2}\right) |\downarrow\rangle. \end{aligned}$$

This is identical to what we had found before. We can also consider the time evolution operator directly

$$U(t) = e^{-i\tilde{H}t/\hbar} = \cos\left(\frac{\Omega_R t}{2}\right) \mathbb{1} + i \sin\left(\frac{\Omega_R t}{2}\right) \hat{\mathbf{h}} \cdot \boldsymbol{\sigma} \quad (1.92)$$

$$= \cos\left(\frac{\Omega_R t}{2}\right) \mathbb{1} + i \sin\left(\frac{\Omega_R t}{2}\right) \left(\frac{\delta}{\Omega_R} \sigma_z - \frac{\omega_R}{\Omega_R} \sigma_x \right). \quad (1.93)$$

from which the above result for $|\psi(t)\rangle = U(t) |\uparrow\rangle$ immediately follows. The probability of a spin flip is

$$p_{\downarrow}(t) = |\langle \downarrow | \psi(t) \rangle|^2 = \frac{\omega_R^2}{\Omega_R^2} \sin^2\left(\frac{\Omega_R t}{2}\right).$$

This is indeed just the same as the z -component of a classical magnetic moment.

1.4.8 Rapid adiabatic passage - Quantum treatment

We revisit the Landau Zener problem but now for a two-state system or equivalently a spin 1/2 system. The problem can be solved rigorously. Consider a spin 1/2 system in a magnetic field \mathbf{B}_{eff} with energies

$$W_{\pm} = \pm \frac{1}{2} \hbar \gamma B_{\text{eff}}. \quad (1.94)$$

For a uniform field B_0 (with $B_1 = 0$), the effective field in the rotating frame is $B_0 - \omega/\gamma$, and

$$W_{\pm} = \pm \frac{1}{2} \hbar(\omega_0 - \omega), \quad (1.95)$$

where $\omega_0 = \gamma B_0$. As ω is swept through resonance, the two states move along their changing eigenenergies. The energies change, but the states do not. There is no coupling between the states, so a spin initially in one or the other will remain so indefinitely no matter how ω changes relative to ω_0 .

In the presence of a rotating field, B_1 , however, the energy levels look quite different: instead of intersecting lines they form non-intersecting hyperbolas separated by energy $\hbar\omega_R$. If the system moves along these hyperbolas, then an \uparrow spin will adiabatically convert into \downarrow and a \downarrow spin will turn into \uparrow .

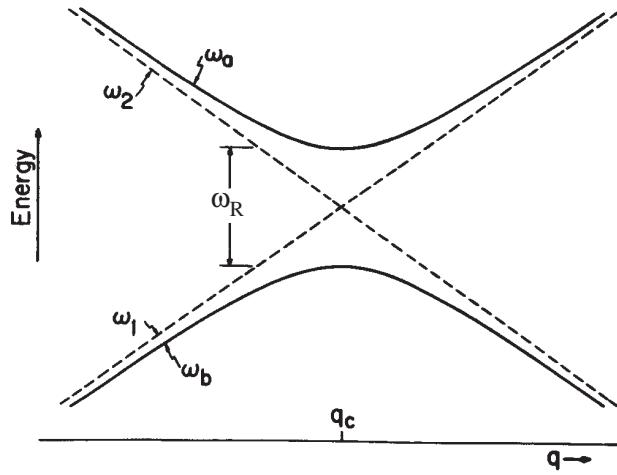


Figure 8. An avoided crossing. If the parameter q that governs the energy levels is swept slowly, the system can adiabatically change its state (“flip its spin”). If q is swept rapidly, it may instead jump across the gap and no spin flip occurs.

Whether or not the system follows an energy level adiabatically depends on how rapidly the energy is changed, compared to the minimum energy separation. To cast the problem in quantum mechanical terms, imagine two non-interacting states whose energy separation, ω , depends on some parameter q which varies linearly in time, and vanishes for some value q_c . Now add a perturbation having an off-diagonal matrix element V which is independent of q , so that the energies at q_c are $\pm V$, as shown in Fig. 8 ($\omega_R = 2V$). The probability that the system will “jump” from one adiabatic level to the other after passing through the “avoided crossing” (i.e., the probability of non-adiabatic behavior) is

$$P_{na} = e^{-2\pi\Gamma} \quad (1.96)$$

where

$$\Gamma = \frac{|V|^2}{\hbar^2} \left[\frac{d\omega}{dt} \right]^{-1} \quad (1.97)$$

This result was originally obtained by Landau and Zener. The jumping of a system as it travels across an avoided crossing is called the Landau-Zener effect. Further description, and reference to the initial papers, can be found in [5]. Inserting the parameters for our magnetic field problem, we have

$$P_{na} = \exp\left\{-\frac{\pi}{2} \frac{\omega_R^2}{d\omega/dt}\right\} \quad (1.98)$$

Note that the factor in the exponential is related to the inequality in Eq. 1.43. When Eq. 1.43 is satisfied, the exponent is large and the probability of non-adiabatic behavior is exponentially small.

Incidentally the “rapid” in adiabatic rapid passage is something of a misnomer. The technique was originally developed in nuclear magnetic resonance in which thermal relaxation effects destroy the spin polarization if one does not invert the population sufficiently rapidly. In the absence of such relaxation processes one can take as long as one pleases to traverse the anticrossings, and the slower the crossing the less the probability of jumping.

1.4.9 Adiabatic passage - Detailed calculation

Our initial Hamiltonian in the rotating frame is (

$$H_0 = \frac{\hbar}{2} \begin{pmatrix} -\delta & 0 \\ 0 & \delta \end{pmatrix}.$$

The full Hamiltonian is

$$H = \frac{\hbar}{2} \begin{pmatrix} -\delta & \omega_R \\ \omega_R & \delta \end{pmatrix}.$$

This has eigenstates $|\pm\rangle$, which have energies separated by ω_R .

We start with $-\delta \ll \omega_R$, so the eigenstates are initially $|\uparrow\rangle$ and $|\downarrow\rangle$. If we change the system adiabatically, states connect smoothly according to the adiabatic theorem, so the system will go from $|\uparrow\rangle$ to $|\downarrow\rangle$. However, for a non-adiabatic change, the system may jump across the avoided crossing, staying in the same state. The probability of a non-adiabatic transition is

$$p_{na} = e^{-\frac{\pi}{2} \frac{\omega_R^2}{\omega}}.$$

We thus see the previous adiabaticity criteria appear.

Now, let us derive this. Suppose we are in a perturbative limit, $p_{\text{flip}} \ll 1$. Since the drive frequency changes, the Hamiltonian is

$$H = \frac{\hbar}{2} \begin{pmatrix} \omega_0 & \omega_R e^{-i\phi(t)} \\ \omega_R e^{+i\phi(t)} & -\omega_0 \end{pmatrix}.$$

In the frame rotating with the drive (in a time-dependent manner),

$$\tilde{H} = \frac{\hbar}{2} \begin{pmatrix} -\delta(t) & \omega_R \\ \omega_R & \delta(t) \end{pmatrix},$$

where $\delta(t) = \dot{\phi}(t) - \omega_0$. Suppose instead we rotate the frame at ω_0 . We find

$$\bar{H} = \frac{\hbar}{2} \begin{pmatrix} 0 & \omega_R e^{-i(\phi(t)-\omega_0 t)} \\ \omega_R e^{+i(\phi(t)-\omega_0 t)} & 0 \end{pmatrix}.$$

The second Hamiltonian is actually easier to deal with. We know that

$$|\psi(t)\rangle = a(t)|\uparrow\rangle + b(t)|\downarrow\rangle$$

for eigenstates of this Hamiltonian. Acting with \bar{H} on this,

$$\dot{a} = -i\frac{\omega_R}{2}e^{-i(\phi(t)-\omega_0 t)}b \text{ and } \dot{b} = -i\frac{\omega_R}{2}e^{+i(\phi(t)-\omega_0 t)}a.$$

Let us use a linear sweep, $\delta(t) = \dot{\phi} - \omega_0 = \alpha t$. So, $\phi(t) = \omega_0 t + \frac{1}{2}\alpha t^2$. Then

$$\dot{a} = -i\frac{\omega_R}{2}e^{-i\alpha t^2/2}b \text{ and } \dot{b} = -i\frac{\omega_R}{2}e^{+i\alpha t^2/2}a.$$

Suppose we start in $|\downarrow\rangle$, and have weak coupling, so at all times $b \approx 1$. Then

$$\dot{a} \approx -i\frac{\omega_R}{2}e^{-i\alpha t^2/2}.$$

So, a will only grow significantly when $\alpha t^2 \leq 1$. That is, the phase needs to be fairly constant for a to accumulate. This occurs for $\Delta t \sim \frac{1}{\sqrt{\alpha}} = \frac{1}{\sqrt{\omega}}$ around the resonance. This implies that $a \approx \omega_R \Delta t \approx \frac{\omega_R}{\sqrt{\omega}}$. The probability of a flip is thus

$$p_{\text{flip}} = |a|^2 \sim \frac{\omega_R^2}{\dot{\omega}},$$

which is proportional to what we expect. We can do the integral exactly since the \dot{a} equation is Gaussian. We find

$$a(t) = -i\omega_R \int_0^\infty dt e^{-i\alpha t^2/2} = -ie^{-i\pi/4} \omega_R \sqrt{\frac{\pi}{2\alpha}}.$$

This gives a flip probability of

$$p_{\text{flip}} = |a|^2 = \frac{\pi}{2} \frac{\omega_R^2}{\dot{\omega}},$$

which is exactly the correct limit of the Landau-Zener formula.

We can perform the calculation also fully non-perturbatively. Note in passing that we could have solved the problem of a classical spin as well, as the result will have to be identical, but it is even more difficult than the quantum case as it involves three coupled equations instead of two.

Returning to the spin-1/2 system, we start with the coupled equations

$$\dot{a} = -i\frac{\omega_R}{2}e^{-i\alpha t^2/2}b \text{ and } \dot{b} = -i\frac{\omega_R}{2}e^{+i\alpha t^2/2}a.$$

We have boundary conditions $a(-\infty) = 0$, $b(-\infty) = 1$. Taking another derivative of a ,

$$\ddot{a} = -i\frac{\omega_R}{2}e^{-i\alpha t^2/2}\dot{b} - i\alpha t \dot{a}.$$

So

$$\ddot{a} = -\frac{\omega_R^2}{4}a - i\alpha t \dot{a}.$$

We substitute $a = e^{-i\alpha t^2/4}c$, and find

$$\ddot{c} + \left(\frac{\omega_R^2}{4} - i\frac{\alpha}{2} + \frac{\alpha^2}{4}t^2 \right) c = 0.$$

This is known as the Weber solution. We make it resemble a harmonic oscillator by introducing

$$z = \sqrt{\alpha}e^{-i\pi/4}t \implies dz = \sqrt{\alpha}e^{-i\pi/4}dt.$$

(the same substitution was actually necessary for the gaussian integral of the perturbative calculation). From this, our equation becomes

$$\frac{d^2c}{dz^2} + \left(i\frac{\omega_R^2}{4\alpha} + \frac{1}{2} - \frac{z^2}{4} \right) c = 0.$$

This can be compared to the Schrödinger equation for the harmonic oscillator

$$-\frac{\hbar^2}{2m}\psi'' + \frac{1}{2}m\omega^2x^2\psi = E\psi \implies \frac{d^2\psi}{d\tilde{x}^2} + \left(n + \frac{1}{2} - \frac{\tilde{x}^2}{4} \right) \psi = 0.$$

when expressing space in units of the harmonic oscillator length $x = \tilde{x}\sqrt{\hbar/m\omega}$, and writing energy $E = \hbar\omega(n + 1/2)$.

Here, we have $n = i\frac{\omega_R^2}{4\alpha}$. The parabolic cylinder functions are ultimately the solution, and accept complex arguments:

$$c(z) = \frac{\omega_R}{2\sqrt{\alpha}}e^{-\pi\omega_R^2/16\alpha}D_{-1-i\frac{\omega_R^2}{4\alpha}}(iz).$$

We eventually find

$$|a(\infty)|^2 = 1 - e^{-\pi\omega_R^2/2\alpha}.$$

This is exactly the spin-flip probability that we expect.

1.5 Density Matrix

1.5.1 General results

The *density matrix* provides a way of treating the time evolution of a quantized system which offers several advantages over the usual time dependent expansion,

$$|\psi(t)\rangle = \sum_n c_n(t) |\psi_n\rangle = \sum_n \langle\psi_n | \psi(t)\rangle |\psi_n\rangle \quad (1.99)$$

plus Schrödinger equation. It provides a natural way to express coherences and to find the expectation value of operators which do not commute with the Hamiltonian, it treats *pure quantum states* and *statistical mixtures* on an equal footing, and it allows straightforward determination of the time evolution of the system even when it is affected by incoherent processes such as damping, addition or subtraction of atoms (from the system) or interactions with other quantized systems not accessible to measurement (eg. collisions).

An operator, A, with matrix elements

$$A_{nm} \equiv \langle\psi_n | A | \psi_m\rangle \quad (1.100)$$

has expectation value at time t

$$\langle A \rangle_t \equiv \langle\psi(t) | A | \psi(t)\rangle \quad (1.101)$$

$$= \sum_{mn} c_m^*(t) c_n(t) A_{nm} \text{(using Eq. 1.99).} \quad (1.102)$$

Clearly the correlation between the c_m and c_n coefficients is important - physically it reflects the coherence between the amplitude for being in states m and n . These correlations are naturally dealt with by the density operator

$$\rho(t) = \overline{|\psi(t)\rangle\langle\psi(t)|} \quad (1.103)$$

because its matrix elements are

$$\rho_{nm}(t) = c_m^*(t) c_n(t) \quad (1.104)$$

The bar here indicates an ensemble average over identically (but not necessarily completely) prepared systems. An ensemble average is essential to treat probabilities (eg. only the ensemble average of spin projections of atoms from an oven is zero although each atom will have $m = +1/2$ or $-1/2$ when measured), and an ensemble average is always implicit in using a density matrix. For notational simplicity, the averaging bar will be eliminated from here on.

The density matrix permits easy evaluation of expectation values: combining Eqs. 1.99 and Eq. 1.104 gives

$$\langle A \rangle_t = \sum_{nm} \rho_{nm} A_{mn} = Tr(\rho(t)A) \quad (1.105)$$

where Tr is the trace, i.e., the sum of the diagonal elements. Eq. 1.105 for $\langle A \rangle$ really involves two sums: the ensemble average in the preparation of the systems, and the usual quantum mechanical sum over the basis to find the expectation value.

The time evolution of the density matrix is determined by a first order

differential equation which is obtained by applying Schrödinger's equation to the time derivative of Eq. 1.103;

$$i\hbar\dot{\rho} = H\rho - \rho H \equiv [H, \rho]. \quad (1.106)$$

This reflects changes in ρ due solely to the interactions (eg. radiation, dc fields) included in the Hamiltonian - additional terms may be added to account for collisions, loss of atoms, damping, etc.

The density matrix operator also provides a convenient test for a properly normalized system (sum of all probabilities, $p_n = c_n^* c_n$, equal to unity)

$$\text{Tr}(\rho(t)) = 1, \quad (1.107)$$

and

$$\text{Tr}\rho^2 \leq \text{Tr}\rho. \quad (1.108)$$

where the equality implies a pure quantum state.

We always have in mind that ρ is to be used on a statistical ensemble of systems similarly prepared. If this preparation is sufficient to force the system into a pure state [so that Eq. 1.99 holds for each member of the ensemble], then the ensemble average is superfluous - if the preparation is insufficient, then there will be random phases between some of the c'_n 's in Eq. 1.100 and some ensemble averages of $c_n^* c_m$ will have modulus less than $|c_n||c_m|$. If no relative phase information is present in the ensemble the ensemble is termed a "mixture" (except it is pure if only one $|c_n|^2$ is non-zero).

1.5.2 Density matrix for two level system

The density matrix for a two level system is

$$\rho = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{21} & \rho_{22} \end{pmatrix} \text{ with } \rho_{12}^* = \rho_{21}^* \quad (1.109)$$

We shall consider a two level system in which $E_1 = \hbar\omega_0/2$ and $E_2 = -\hbar\omega_0/2$ where ω_0 is constant, and we shall subject it to an off-diagonal perturbation of arbitrary strength and time dependence: $\langle 1 | H' | 2 \rangle = (V_1 - iV_2)/2$. Thus

$$H_0 = \frac{\hbar}{2} \begin{pmatrix} \omega_0 & 0 \\ 0 & -\omega_0 \end{pmatrix} = \frac{\hbar\omega_0}{2} \sigma_z \quad (1.110)$$

and

$$H' = \frac{1}{2} \begin{pmatrix} 0 & V_1 - iV_2 \\ V_1 + iV_2 & 0 \end{pmatrix} = \frac{V_1}{2} \sigma_x + \frac{V_2}{2} \sigma_y \quad (1.111)$$

so

$$H = \frac{1}{2} \begin{pmatrix} \hbar\omega_0 & V_1 - iV_2 \\ V_1 + iV_2 & \hbar\omega_0 \end{pmatrix} = \frac{1}{2} [V_1 \sigma_x + V_2 \sigma_y + \hbar\omega_0 \sigma_z] \quad (1.112)$$

is the full Hamiltonian (the σ 's are Pauli spin matrices). This is a general enough system to encompass most two-level systems which are encountered in resonance physics.

Before solving for $\dot{\rho}$ (which we could do by grinding away using Eq. 1.106) we shall change variables in the density matrix:

$$\rho = \frac{1}{2} \begin{pmatrix} r_0 + r_3 & r_1 - ir_2 \\ r_1 + ir_2 & r_0 - r_3 \end{pmatrix} = \frac{1}{2} [r_0 I + r_1 \sigma_x + r_2 \sigma_y + r_3 \sigma_z] \quad (1.113)$$

There is no loss of generality in this substitution (it has 4 independent quantities just as ρ does), and it makes the physical constraints on ρ manifest, e.g.

$$Tr(\rho) = r_0 = 1 \quad (1.114)$$

and $\rho_{12} = \rho_{21}^*$ obviously.

Now we have expressed both H and ρ in terms of Pauli spin matrices. We can now solve the equation of motion for $\rho(t)$,

$$i\hbar\dot{\rho} = [H, \rho], \quad (1.115)$$

by using the cyclic commutation relations $[\sigma_j, \sigma_{j+1}] = 2i\sigma_{j+2}$ and then equating the coefficients of σ_x, σ_y , and σ_z , (rather than having to grind out the matrix products term by term):

$$\begin{aligned} \sigma_x : \dot{r}_1 &= \frac{1}{\hbar} V_2 r_3 - \omega_0 r_2 \\ \sigma_y : \dot{r}_2 &= \omega_0 r_1 - \frac{1}{\hbar} V_1 r_3 \\ \sigma_z : \dot{r}_3 &= \frac{1}{\hbar} V_1 r_2 - \frac{1}{\hbar} V_2 r_1 \end{aligned} \quad (1.116)$$

These final results can be summarized by using the vector representation due [1]. Define

$$\boldsymbol{\omega} = \frac{1}{\hbar} V_1 \hat{\mathbf{x}} + \frac{1}{\hbar} V_2 \hat{\mathbf{y}} + \omega_0 \hat{\mathbf{z}} \quad \text{and} \quad \hat{\mathbf{r}} = r_1 \hat{\mathbf{x}} + r_2 \hat{\mathbf{y}} + r_3 \hat{\mathbf{z}} \quad (1.117)$$

Using these definitions it is easy to see that Eq. 1.116, 1.116, 1.116 become

$$\frac{d\mathbf{r}}{dt} = \boldsymbol{\omega} \times \mathbf{r} \quad (1.118)$$

1.118 proves that the time evolution of the density matrix for our very general 2-level system is isomorphic to the behavior of a classic magnetic moment in a magnetic field which points along $\boldsymbol{\omega}$ (??). Our previous discussion showing that the quantum mechanical spin obeyed this equation also is therefore superfluous for spin 1/2 systems.)

One consequence of 1.118 is that \mathbf{r} is always perpendicular to $\dot{\mathbf{r}}$ so that $| \mathbf{r} |$ does not change with time. This implies that if ρ is initially a pure state, ρ remains forever in a pure state no matter how violently $\boldsymbol{\omega}$ is gyrated, because (recall $\sigma_i^2 = I$)

$$Tr\rho^2 = \frac{1}{2}(r_0^2 + r_1^2 + r_2^2 + r_3^2) = \frac{1}{2}(|\mathbf{r}|^2 + r_0^2) \quad (1.119)$$

[1.119](#) will be satisfied for all time since $|\mathbf{r}|^2$ doesn't change and, the state will remain pure. In general it is not possible to decrease the purity (coherence) of a system with a Hamiltonian like the one in [1.112](#). Since real coherences do, in fact die out, we shall have to add relaxation processes to our description in order to approach reality. The density matrix formulation makes this easy to do, and this development will be done in the next part of the section.

1.5.3 Phenomenological treatment of relaxation: Bloch equations

Statistical mechanics tells us the form which the density matrix will ultimately take, but it does not tell us how the system will get there or how long it will take. All we know is that ultimately the density matrix will thermalize to

$$\rho^T = \frac{1}{Z}e^{-H_0/kT}, \quad (1.120)$$

where Z is the partition function.

Since the interactions which ultimately bring thermal equilibrium are incoherent processes, the density matrix formulation seems like a natural way to treat them. Unfortunately in most cases these interactions are sufficiently complex that this is done phenomenologically. For example, the equation of motion for the density matrix [1.116](#) might be modified by the addition of a damping term:

$$\dot{\rho} = \frac{1}{i\hbar}[H, \rho] - (\rho - \rho^T)/T_e \quad (1.121)$$

which would (in the absence of a source of non-equilibrium interactions) drive the system to equilibrium with time constant T_e .

This equation is not sufficiently general to describe the behavior of most systems studied in resonance physics, which exhibit different decay times for the energy and phase coherence, called T_1 and T_2 respectively.

- T_1 - decay time for population differences between non-degenerate levels, eg. for r_3 (also called the energy decay time)
- T_2 - decay time for coherences (between either degenerate or non-degenerate states), i.e. for r_1 or r_2 .

The reason is that, in general, it requires a weaker interaction to destroy coherence (the relative phase of the coefficients of different states) than to destroy the population difference, so some relaxation processes will relax only the phase, resulting in $T_2 < T_1$. (caution: certain types of collisions violate this generality.)

The effects of thermal relaxation with the two decay times described above are easily incorporated into the vector model for the 2-level system since the z-component of the population vector \mathbf{r} [1.117](#) represents the population difference

and r_x and r_y represent coherences (i.e. off-diagonal matrix elements of ρ): The results (which modify 1.119) are

$$\dot{r}_z = \frac{1}{\hbar}(\boldsymbol{\omega} \times \mathbf{r})_2 - (r_z - r_z^T)/T_1 \quad (1.122)$$

$$\dot{r}_{x,y} = \frac{1}{\hbar}(\boldsymbol{\omega} \times \mathbf{r})_{x,y} - (r_{x,y} - r_{x,y}^T)/T_2 \quad (1.123)$$

(r_z^T is determined from 1.120). For a magnetic spin system \mathbf{r} corresponds directly to the magnetic moment $\boldsymbol{\mu}$. The above equations were first introduced by Bloch [4] in this context and are known as the Bloch equations.

The addition of phenomenological decay times does not generalize the density matrix enough to cover situations where atoms (possibly state-selected) are added or lost to a system. This situation can be covered by the addition of further terms to $\dot{\rho}$. Thus a calculation on a resonance experiment in which state-selected atoms are added to a two-level system through a tube which also permits atoms to leave (eg. a hydrogen maser) might look like:

$$\begin{aligned} \dot{\rho} = & \frac{1}{i\hbar}[\rho, H] - \begin{pmatrix} (\rho_{11} - \rho_{11T})/T_1 & \rho_{12}/T_2 \\ \rho_{21}/T_2 & (\rho_{22} - \rho_{22T})/T_1 \end{pmatrix} \\ & + R \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} - \rho/T_{\text{escape}} - \rho/T_{\text{collision}} \end{aligned} \quad (1.124)$$

R is the rate of addition of state-selected atoms

The last two terms express effects of atom escape from the system and of collisions (e.g. spin exchange) that can't easily be incorporated in T_1 and T_2 .

The terms representing addition or loss of atoms will not have zero trace, and consequently will not maintain $\text{Tr}(\rho) = 1$. Physically this is reasonable for systems which gain or lose atoms; the application of the density matrix to this case shows its power to deal with complicated situations. In most applications of the above equation, one looks for a steady state solution (with $\dot{\rho} = 0$), so this does not cause problems.

1.5.4 Introduction: Electrons, Protons, and Nuclei

The two-level system is basic to atomic physics because it approximates accurately many physical systems, particularly systems involving resonance phenomena. All two-level systems obey the same dynamical equations: thus to know one is to know all. The archetype two level system is a spin-1/2 particle such as an electron, proton or neutron. The spin motion of an electron or a proton in a magnetic field, for instance, displays the total range of phenomena in a two level system. To slightly generalize the subject, however, we shall also include the motion of atomic nuclei. Here is a summary of their properties.

MASS

electron	$m = 0.91 \times 10^{-31} \text{ kg}$
proton	$M_p = 1.67 \times 10^{-27} \text{ kg}$
neutron	M_p
nuclei	$M = AM_p$
	$A = N + Z = \text{mass number}$
	$Z = \text{atomic number}$
	$N = \text{neutron number}$

CHARGE

electron	-e	$e = 1.60 \times 10^{-19} \text{ C}$
proton	+e	
neutron	0	
nucleus	Ze	

ANGULAR MOMENTUM

electron	$S = \hbar/2$
proton	$I = \hbar/2$
neutron	$I = \hbar/2$
nuclei	even A: $I/\hbar = 0, 1, 2, \dots$
	odd A: $I/\hbar = 1/2, 3/2, \dots$

STATISTICS

electrons	Fermi-Dirac
nuclei:	even A, Bose-Einstein
	odd A, Fermi-Dirac

ELECTRON MAGNETIC
MOMENT

$$\begin{aligned}\mu_e &= \gamma_e S = -g_s \mu_B S / \hbar \\ \gamma_e &= \text{gyromagnetic ratio} = e/m = 2\pi \times 2.80 \times 10^4 \text{ MHz T}^{-1} \\ g_s &= \text{free electron g-factor} = 2 \text{ (Dirac Theory)} \\ \mu_B &= \text{Bohr magneton} = e\hbar/2m = 0.93 \times 10^{-24} \text{ JT}^{-1} \text{ (erg/gauss)}\end{aligned}$$

(Note that μ_e is negative. We show this explicitly by taking g_s to be positive, and writing $\mu_e = -g_s \mu_B S$)

NUCLEAR MAGNETIC MOMENTS

$\mu_{nuc} = \gamma_{nuc}\hbar I = g_{nuc}\mu_N I/\hbar$	
γ_I	= gyromagnetic ratio of the nucleus
μ_N	= nuclear magneton = $e\hbar/2Mc = \mu_B(m/M_p)$
proton	$g_p = 5.6, \gamma_p = 2\pi \times 42.6 \text{ MHz T}^{-1}$
neutron	$g_n = -3.7$

References

- [1] R.P. Feynman, F.L. Vernon, and R.W. Hellwarth, J. Appl. **28**, 49 (1957).
- [2] N.F. Ramsey, Phys. Rev. **76**, 996 (1949).
- [3] S.R. Lundeen and F.M. Pipkin, Phys. Rev. Lett. **34**, 1368 (1975).
- [4] F. Bloch, Phys. Rev. **70**, 460 (1946).
- [5] Rubbermark et al., Phys. Rev. A **23**, 3107 (1981).
- [6] A.H. Autler and C.H. Townes, Phys. Rev. **100**, 703, 1955.
- [7] B.R. Mollow, Phys. Rev. A ??, 1522.
- [8] B.R. Mollow, Phys. Rev. A **5**, 2217, 1972.
- [9] R.E. Grove, F.Y. Wu, S. Ezekiel, Phys. Rev. A **15**, 227.
- [10] R.M. Whitley and C.R. Stroud, Phys. Rev. A **14**, 1498.
- [11] C. Delsart and J.C. Keller, J. Phys. B **9**, 2769.
- [12] J.L. Picque and J. Pinard, J. Phys. B **9**, L77.
- [13] F.Y. Wu, S. Ezekiel, M. Ducloy, B.R. Mollow, Phys. Rev. Lett. **38**, 1077.

Chapter 2

Atoms: Some Basics

This chapter comprises some miscellaneous topics in basic atomic physics. They do not form a unified exposition and the discussions are in some cases quite informal. The topics are intended to provide helpful reviews and to introduce a number of useful concepts.

2.1 Spectroscopic Notation

Neutral atoms consist of a heavy nucleus with charge Z surrounded by Z electrons. Positively charged atomic ions generally have structure similar to the atom with the same number of electrons except for a scale factor; negative ions lack the attractive Coulomb interaction at large electron-core separation and hence have few if any bound levels. Thus the essential feature of an atom is its number of electrons, and their mutual arrangement as expressed in the quantum numbers.

An isolated atom has two good angular momentum quantum numbers, J and M_J . (This is strictly true only for atoms whose nuclei have spin $I = 0$. However, J is never significantly destroyed by coupling to I in ground state atoms.) In zero external field the atomic Hamiltonian possesses rotational invariance which implies that each J level is degenerate with respect to the $2J + 1$ states with specific M_J (traditional atomic spectroscopists call these states “sublevels”). For each J, M_J an atom will typically have a large number of discrete energy levels (plus a continuum) which may be labeled by other quantum numbers.

If Russel-Saunders coupling ($L - S$ coupling) is a good description of the atom, then L and S , where

$$\mathbf{L} = \sum_{i=1}^N \ell_i \quad (2.1)$$

$$\mathbf{S} = \sum_{i=1}^N \mathbf{s}_i \quad (2.2)$$

are nearly good quantum numbers and may be used to distinguish the levels. In this case the level is designated by a *Term* symbolized $^{2S+1}L_J$ where $2S + 1$ and J are written numerically and L is designated with this letter code:

$L:$	O	1	2	3	4	...
Letter:	S	P	D	F	G	

The Letters stand for Sharp, Principal, Diffuse, and Fundamental - adjectives applying to the spectral lines of one electron atoms. The term symbol is frequently preceded by the n value of the outermost electron, e.g. the $3^2S_{1/2}$ ground state of Na.

This discussion of the term symbol has been based on an external view of the atom. Alternatively one may have or assume knowledge of the internal structure - the quantum numbers of each electron. These are specified as the *configuration*, e.g.

$$1s^2 2s^2 2p^2, \quad (2.3)$$

a product of symbols of the form $n\ell^k$ which represents k electrons in the orbital n, ℓ . n is the principal quantum number, which characterizes the radial motion and has the largest influence on the energy. n and m are written numerically,

but the *spdf* ... coding is used for ℓ . An example of the configuration for Ca is $1s^2 2s^2 2p^6 3s 3d$ which is frequently abbreviated $3d$. In general each configuration leads to several terms which may be split apart by several eV. The above Ca configuration gives rise to 1D_2 and $^3D_{1,2,3}$ for example.

In classifying levels, the term is generally more important than the configuration because it determines the behavior of an atom when it interacts with E or B fields. Selection rules, for instance, generally deal with $\Delta J, \Delta L, \Delta S$. Furthermore the configuration may not be pure - if two configurations give rise to the same term (and have the same parity) then intra-atomic electrostatic interactions can mix them together. This process, called configuration interaction, results in shifts in the level positions and intensities of special lines involving them as well as in correlations in the motions of the electrons within the atoms.

2.2 One-Electron Atoms

In trying to understand some new phenomenon for the first time it is common sense (and good science) to study it in the simplest situation where it is manifest. With atoms it is evident that hydrogen is of paramount simplicity and, much of the “fundamental” physics which has been discovered in atoms has been discovered in hydrogen (but Na and other alkalis have taken over since tunable lasers arrived). In this chapter we shall use the phrase “one electron atom” to include not only atoms which are isoelectronic with H (e.g. He^+ , Li^{++} ... etc.) but also atoms with one electron which is far less weakly bound than all the others so that the inner electrons may be considered collectively as a core whose interaction with the active electron may be adequately described by parameters such as a scattering length for low energy electrons, polarizability, etc.

2.2.1 The Bohr Atom

We briefly review the Bohr atom— a model that was soon obsolete, but which nevertheless provided the major impetus for developing quantum mechanics. Balmer’s empirical formula of 1885 had reproduced Angstrom’s observations of spectral lines in hydrogen to 0.1 Å accuracy, but it was not until 1913 that Bohr gave an explanation for this based on a quantized mechanical model of the atom. This model involved the postulates of the Bohr Atom:

- Electron and proton are point charges whose interaction is coulombic at all distances.
- Electron moves in circular orbit about the center of mass in *stationary states* with orbital angular momentum $L = n\hbar$.

These two postulates give the energy levels:

$$-E_n = \frac{1}{2} \left(\frac{m(e^2/4\pi\epsilon_0)^2}{\hbar^2} \frac{M}{M+m} \right) / n^2 = hcR_H/n^2 \quad (2.4)$$

where $m(e^2/4\pi\epsilon_o)^2/\hbar^2$ is the hartree², $M/(M+m)$ is the reduced mass factor, and R_H is the Rydberg in H: ($R_H = 1.09677576 \times 10^5 \text{ cm}^{-1}$), and $hcR_H = 13.6 \text{ eV}$.

- One quantum of radiation is emitted when the system changes between these energy levels.
- The wave number of the radiation is given by the Bohr frequency criterion:

$$\sigma_{n \rightarrow m} = (E_n - E_m)/(hc)$$

(Note that the wave number is a spectroscopic unit defined as the number of wavelengths per cm, $\sigma = 1/\lambda$. It is important not to confuse the wave number with the magnitude of the *wave vector* \mathbf{k} which defines a traveling wave of the form $\exp(i(\mathbf{k} \cdot \mathbf{r} - \omega t))$. The magnitude of the wave vector is 2π times the wave number.)

The mechanical spirit of the Bohr atom was extended by Sommerfeld in 1916 using the Wilson-Sommerfeld quantization rule,

$$\oint p_i dq_i = n_i h \quad (2.5)$$

where q_i and p_i are conjugate coordinate and momentum pairs for each degree of freedom of the system. This extension yielded elliptical orbits which were found to have an energy nearly degenerate with respect to the orbital angular momentum for a particular principal quantum number n . The dependency was lifted by a relativistic correction whose splitting was in agreement with the observed fine structure of hydrogen. (This was a great cruel coincidence in physics. The mechanical description ultimately had to be completely abandoned, in spite of the excellent agreement of theory and experiment.) Although triumphant in hydrogen, simple mechanical models of helium or other two-electron atoms failed, and real progress in understanding atoms had to await the development of quantum mechanics.

2.2.2 Radial Schrödinger equation for central potentials

Stationary solutions of the time dependent Schrödinger equation

$$i\hbar \frac{\partial \psi(\mathbf{r}, \mathbf{t})}{\partial t} = H(\mathbf{r})\psi(\mathbf{r}, \mathbf{t}) \quad (2.6)$$

(H is the Hamiltonian operator) can be represented as

$$\psi(\mathbf{r}, t) = e^{-iE_n t/\hbar} \psi_n(\mathbf{r}), \quad (2.7)$$

where n stands for all quantum numbers necessary to label the state. This leads to the time-independent Schrödinger equation

²There is often confusion about the capitalization of units named after people. The rule is that the unit is spelled lower case, but the abbreviation is upper case. Thus, 2.4 newton, or 2.4 N.

$$[H(\mathbf{r}) - E_n]\psi_n(\mathbf{r}) = \mathbf{0} \quad (2.8)$$

A pervasive application of this equation in atomic physics is for the case of a spherically symmetric one-particle system of mass μ . In this case the Hamiltonian is

$$\begin{aligned} H &\equiv \text{Kinetic Energy} + \text{Potential Energy} \\ &= p^2/2\mu + V(r) = \frac{-\hbar^2\nabla^2}{2\mu} + V(r) \end{aligned}$$

$$= -\frac{\hbar^2}{2\mu} \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] + V(r) \quad (2.9)$$

where the kinetic energy operator ∇^2 has been written in spherical coordinates. Because V is spherically symmetric, the angular dependence of the solution is characteristic of spherically symmetric systems in general and may be factored out:

$$\psi_{n\ell m}(\mathbf{r}) = R_{n\ell}(r)Y_{\ell m}(\theta, \phi) \quad (2.10)$$

$Y_{\ell m}$ are the *spherical harmonics* and ℓ is the eigenvalue of the operator for the orbital angular momentum, \mathbf{L} ,

$$L^2 Y_{\ell m} = \ell(\ell + 1)\hbar^2 Y_{\ell m} \quad (2.11)$$

and m is the eigenvalue of the projection of L on the quantization axis (which may be chosen at will)

$$L_z Y_{\ell m} = m\hbar Y_{\ell m} \quad (2.12)$$

With this substitution Eq. 2.10 the time independent radial Schrödinger equation becomes

$$\frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR_{n\ell}}{dr} \right) + \left[\frac{2\mu}{\hbar^2} [E_{n\ell} - V(r)] - \frac{\ell(\ell + 1)}{r^2} \right] R_{n\ell} = 0 \quad (2.13)$$

This is the equation which is customarily solved for the hydrogen atom's radial wave functions. For most applications (atoms, scattering by a central potential, diatomic molecules) it is more convenient to make a further substitution.

$$R_{n\ell}(r) = y_{n\ell}(r)/r \quad (2.14)$$

which leads to

$$\frac{d^2 y_{n\ell}(r)}{dr^2} + \left[\frac{2\mu}{\hbar^2} [E_{n\ell} - V(r)] - \frac{\ell(\ell + 1)}{r^2} \right] y_{n\ell}(r) = 0 \quad (2.15)$$

with the boundary condition $y_{n\ell}(0) = 0$. This equation is identical with the time independent Schrödinger equation for a particle of mass μ in an effective

one dimensional potential,

$$V^{\text{eff}}(r) = V(r) + L^2/(2\mu r^2) = V(r) + \frac{\hbar^2 \ell(\ell+1)}{2\mu r^2} \quad (2.16)$$

The term $\hbar^2 \ell(\ell+1)/2\mu r^2$ is called the *centrifugal potential*, and adds to the actual potential the kinetic energy of the circular motion that must be present to conserve angular momentum.

2.2.3 Radial equation for hydrogen

The quantum treatment of hydrogenic atoms or ions appears in many textbooks and we present only a summary³. For hydrogen Eq. 2.15 becomes

$$\frac{d^2 y_{n\ell}(r)}{dr^2} + \left[\frac{2\mu}{\hbar^2} \left[E_n + \frac{e^2}{r} \right] - \frac{\ell(\ell+1)}{r^2} \right] y_{n\ell}(r) = 0 \quad (2.17)$$

First look at this as $r \rightarrow 0$, the dominant terms are

$$\frac{d^2 y}{dr^2} - \frac{\ell(\ell+1)}{r^2} y = 0 \quad (2.18)$$

for any value of E . It is easily verified that the two independent solutions are

$$y \sim r^{\ell+1} \text{ and } y \sim r^{-\ell} \quad (2.19)$$

For $\ell \geq 1$ the only normalizable solution is $y \sim r^{\ell+1}$.

Question: What happens to these arguments for $\ell = 0$? What implications does this have for the final solution? Messiah has a good discussion of this.

We look next at the solution for $r \rightarrow \infty$ where we may investigate a simpler equation if $\lim_{r \rightarrow \infty} rV(r) = 0$. For large r :

$$\frac{d^2 y}{dr^2} + \frac{2\mu E}{\hbar^2} y = 0 \quad (2.20)$$

If $E > 0$, this equation has oscillating solutions corresponding to a free particle. For $E < 0$ the equation has exponential solutions, but only the decaying exponential is physically acceptable (i.e. normalizable)

$$R(r) = y(r)/r = \frac{1}{r} e^{(-2\mu E/\hbar^2)^{1/2} r} \quad (2.21)$$

When $E < 0$, it is possible to obtain physically reasonable solutions to Eq. 2.15 (or indeed *any* bound state problem) only for certain discrete values of E , its eigenvalues. This situation arises from the requirement that the radial solution

³The most comprehensive treatment of hydrogen is the classic text of Bethe and Salpeter, *The Quantum Mechanics of One- and Two-Electron Atoms*, H. A. Bethe and E. E. Salpeter, Academic Press (1957). Messiah is also excellent.

be normalizable, which requires that $\int_0^\infty y^2(r)dr = 1$, or alternatively, that $y(r)$ vanishes sufficiently rapidly at large r .

Eq. 2.17 is a prescription for generating a function $y(r)$ for arbitrary $E < 0$ given y and dy/dr at any point. This can be solved exactly in hydrogen. For other central potentials, one can find the eigenvalues and eigenstates by computation. One proceeds as follows: Select a trial eigenvalue, $E < 0$. Starting at large r a “solution” of the form of Eq. 2.21 is selected and extended in to some intermediate value of r , r_m . At the origin one must select the solution of the form $y_\ell \sim r^{\ell+1}$ Eq. (2.19); this “solution” is then extended out to r_m . The two “solutions” may be made to have the same value at r_m by multiplying one by a constant; however, the resulting function is a valid solution only if the first derivative is continuous at r_m , and this occurs only for a discrete set of $E_{n\ell}$. The procedure described here is, in fact, the standard Numerov-Cooley algorithm for finding bound states. Its most elegant feature is a procedure for adjusting the trial eigenvalue using the discontinuity in the derivative that converges to the correct energy very rapidly.

For the hydrogen atom, the eigenvalues can be determined analytically. The substitution

$$y_\ell(r) = r^{\ell+1} e^{(-2\mu E/\hbar^2)^{1/2}r} v_\ell(r) \quad (2.22)$$

leads to a particularly simple equation. To make it dimensionless, one changes the variable from r to x

$$x = [2(-2\mu E)^{1/2}/\hbar]r \quad (2.23)$$

so the exponential in Eq. 2.22 becomes $e^{-1/2x}$, and defines

$$v = \frac{\hbar}{(-2\mu E)^{1/2}a_o} \quad (2.24)$$

where a_o is the Bohr radius:

$$\left[x \frac{d^2}{dx^2} + (2\ell + 2 - x) \frac{d}{dx} - (\ell + 1 - v) \right] v_\ell = 0 \quad (2.25)$$

This is a Laplace equation and its solution is a confluent hypergeometric series. To find the eigenvalues one now tries a Taylor series

$$v_\ell(x) = 1 + a_1 x + a_2 x^2 + \dots \quad (2.26)$$

for v_ℓ . This satisfies the equation only if the coefficients of each power of x are satisfied, i.e..

$$\begin{aligned} x^0 : \quad & (2\ell + 2)a_1 - (\ell + 1 - v) = 0 \\ x^1 : \quad & 2(2\ell + 3)a_2 - (\ell + 2 - v)a_\ell = 0 \\ x^{p-1} : \quad & p(2\ell + 1 + p)a_p - (\ell + p - v)a_{p-1} = 0 \end{aligned} \quad (2.27)$$

The first line fixes a_1 , the second then determines a_2 , and in general

$$a_p = \frac{(\ell + p - v)}{p(2\ell + 1 + p)} a_{p-1} \quad (2.28)$$

In general Eq. 2.27 will give a p^{th} coefficient on the order of $1/p!$ so

$$v_\ell(x) \sim \sum_{p=0} x^p / p! = e^x \quad (2.29)$$

this spells disaster because it means $y = r^{\ell+1} e^{-x/2} v_\ell(r)$ diverges. The only way in which this can be avoided is if the series truncates, i.e. if v is an integer:

$$v = n = n' + \ell + 1 \quad n = 0, 1, 2, \dots \quad (2.30)$$

so that a_n , will be zero and v_ℓ will have n' nodes. Since $n' \geq 0$, it is clear that you must look at energy level $n \geq \ell + 1$ to find a state with angular momentum ℓ (e.g. the $2d$ configuration does not exist). This gives the eigenvalues of hydrogen (from Eq. 2.24)

$$E_n = -\frac{1}{2} \alpha^2 \mu c^2 / (n' + \ell + 1)^2 = -\frac{1}{2} \frac{\mu(e^2/4\pi\epsilon_0)^2}{\hbar^2} \frac{1}{n^2} = -\frac{hcR_H}{n^2}, \quad (2.31)$$

which agrees with the Bohr formula.

2.2.4 Expectation values of $1/r^k$ for hydrogen wavefunction

The expectation values of $1/r^k$ (k being an integer) for hydrogen wave function ψ_{nlm} are often very useful when calculating the averaged value of physical quantities. They can be evaluated in a systematic way, which will be demonstrated in the following.

For $k = 0$, the expectation value is just 1 because of the normalization of the wavefunction.

For $k = 1$, the expectation value of $1/r$ can be obtained by using the Virial theorem:

$$\left\langle -\frac{e^2}{4\pi\epsilon_0 r} \right\rangle = 2E_n = -\frac{e^2}{4\pi\epsilon_0 a_0 n^2}. \quad (2.32)$$

Thus we have:

$$\left\langle \frac{a_0}{r} \right\rangle = \frac{1}{n^2}. \quad (2.33)$$

For $k = 2$, the expectation value can be evaluated by using the Feynman-Hellmann theorem. As a reminder, the Feynman-Hellmann theorem states that for Hamiltonian $H(\lambda)$ (here λ is an external parameter), we have the following relation:

$$\frac{dE(\lambda)}{d\lambda} = \left\langle \psi(\lambda) \left| \frac{dH(\lambda)}{d\lambda} \right| \psi(\lambda) \right\rangle. \quad (2.34)$$

$E(\lambda)$ and $|\psi(\lambda)\rangle$ are the eigenvalue and eigenstates of $H(\lambda)$, respectively. Here we take the parameter λ to be the angular momentum quantum number ℓ . The energy eigenvalues can be rewritten as $E_\ell \propto 1/(n' + \ell + 1)^2$, with n' being the radial quantum number and remains fixed. The Hamiltonian can also be

rewritten as a function of ℓ :

$$H(\ell) = -\frac{\hbar^2}{2m} \frac{d^2}{dr^2} + \frac{\hbar^2}{2m} \frac{\ell(\ell+1)}{r^2} - \frac{e^2}{4\pi\epsilon_0 r}. \quad (2.35)$$

Using the Feynman-Hellmann theorem, we have

$$\frac{dE_\ell}{d\ell} = \left\langle \psi_\ell \left| \frac{dH(\ell)}{d\ell} \right| \psi_\ell \right\rangle \quad (2.36)$$

Plugging in the expression for E_ℓ ,

$$\frac{\hbar^2}{mn^3 a_0^2} = \frac{\hbar^2}{m} \left(l + \frac{1}{2} \right) \left\langle \frac{1}{r^2} \right\rangle. \quad (2.37)$$

Thus, we finally obtained the expectation value for $1/r^2$:

$$\left\langle \frac{1}{r^2} \right\rangle = \frac{1}{\left(\ell + \frac{1}{2} \right) n^3 a_0^2}. \quad (2.38)$$

For other k values, the expectation values can either be calculated directly from the wavefunction, or by using the recursive relations (called the Kramers-Pasternack relations). Here we will just write down the recursive equations without a proof:

$$\frac{k+1}{n^2} \left\langle r^k \right\rangle - (2k+1)a_0 \left\langle r^{k-1} \right\rangle + \frac{k}{4} \left[(2\ell+1)^2 - k^2 \right] a_0^2 \left\langle r^{k-2} \right\rangle = 0 \quad (2.39)$$

The expectation value for small k values are summarized in the following:

$$\begin{aligned} & \dots \\ \left\langle \frac{a_0^3}{r^3} \right\rangle &= \frac{1}{\ell \left(\ell + \frac{1}{2} \right) (\ell+1)n^3} \\ \left\langle \frac{a_0^2}{r^2} \right\rangle &= \frac{1}{\left(\ell + \frac{1}{2} \right) n^3} \\ \left\langle \frac{a_0}{r} \right\rangle &= \frac{1}{n^2} \\ \left\langle \frac{a_0^0}{r^0} \right\rangle &= 1 \\ & \dots \end{aligned} \quad (2.40)$$

2.3 One Electron Atoms with Cores: The Quantum Defect

2.3.1 Phenomenology

It is observed that the eigenvalues of atoms which have one valence electron have the same density (in energy space) as those of hydrogen, but not the same

positions. If one defines an effective quantum number

$$n^* = (R/E_n)^{1/2} \quad (2.41)$$

such that

$$T_n = -\frac{\text{const}}{n^{*2}} \quad (2.42)$$

then the formula can naturally reproduce the energy levels of a particular term (e.g. D). However, it was found that the n^* values for adjacent levels differ by almost exactly 1.000, especially after the first few terms. Thus the term energy can be written

$$T_n = -\frac{\text{const}}{n^2 - \delta_\ell^2} \quad (2.43)$$

where the *quantum defect* δ_ℓ is defined by

$$\delta_\ell \equiv n - n^*. \quad (2.44)$$

Here, n is the principal quantum number of the valence electron for that term (if one does not know n , using the closest larger integer than n^* still leads to useful results), remains very constant with respect to n , but decreases rapidly with respect to L . A more accurate empirical formula for the term values of a series is the Balmer-Ritz formula

$$T_n = \frac{Z^2 R}{(n - \delta_\ell - \beta_\ell/n^2)^2} \quad (2.45)$$

A somewhat related concept in the study of x-ray lines was proposed by Moseley. When comparing spectra of isoelectronic ions, he found that a useful empirical formula was

$$T_n = \frac{(Z - \Delta Z)^2 R}{n^2} \quad (2.46)$$

in which the charge is adjusted, rather than n . ΔZ may be regarded as the amount of nuclear charge shielded by the core electrons, and an effective nuclear charge, $Z_e = Z - \Delta Z$ can be introduced.

2.3.2 Explanation

It must always be kept in mind that the quantum defect is a phenomenological result. To explain how such a simple result arises is obviously an interesting challenge, but it is not to be expected that the solution of this problem will lead to great new physical insight. The only new results obtained from understanding quantum defects in one electron systems are the connection between the quantum defect and the electron scattering length for the same system (Mott and Massey section III 6.2) which may be used to predict low energy electron scattering cross sections (*ibid.* XVII 9&10), and the simple expressions relating δ_L to the polarizability of the core for larger δ_L [1]. The principal use of the quantum defect is to predict the positions of higher terms in a series for which δ_L is known.

Explanations of the quantum defect range from the elaborate fully quantal explanation of Seaton [2] to the extremely simple treatment of Parsons and Weisskopf [3], who assume that the electron can not penetrate inside the core at all, but use the boundary condition $R(r_c) = 0$ which requires relabeling the lowest ns state 1s since it has no nodes outside the core. This viewpoint has a lot of merit because the exclusion principle and the large kinetic energy of the electron inside the core combine to reduce the amount of time it spends in the core. This is reflected in the true wave function which has n nodes in the core and therefore never has a chance to reach a large amplitude in this region.

To show the physics without much math (or rigor) we turn to the JWKB solution to the radial Schrödinger Equation (see Messiah Ch. VI). Defining the wave number

$$k_\ell(r) = \sqrt{2m[E - V_{\text{eff}}(r)]\hbar} \quad (\text{remember } V_{\text{eff}} \text{ depends on } \ell) \quad (2.47)$$

then the phase accumulated in the classically allowed region is

$$\phi_\ell(E) = \int_{r_i}^{r_o} k_\ell(r) dr \quad (2.48)$$

where r_i and r_o are the inner and outer turning points. Bound state eigenvalues are found by setting

$$[\phi_\ell(E) - \pi/2] = n\pi \quad (2.49)$$

(The $\pi/2$ comes from the connection formulae and would be $1/4$ for $\ell = 0$ state where $r_i = 0$. Fortunately it cancels out.)

To evaluate $\phi(E)$ for hydrogen use the Bohr formula for $n(E)$,

$$\phi_H(E) = \pi(hcR_H/E)^{1/2} + \pi/2 \quad (\text{independent of } \ell) \quad (2.50)$$

In the spirit of the JWKB approximation, we regard the phase as a continuous function of E . Now consider a one-electron atom with a core of inner shell electrons that lies entirely within r_c . Since it has a hydrogenic potential outside of r_c , its phase can be written (where r_{oH}, r_{iH} are the outer and inner turning points for hydrogen at energy E):

$$\begin{aligned} \phi_\ell(E) &= \int_{r_i}^{r_{oH}} k_\ell(r) dr = \int_{r_i}^{r_c} k_\ell(r) dr + \int_{r_c}^{r_{oH}} k_H(r) dr \\ &= \int_{r_i}^{r_c} k(r) dr - \int_{r_{iH}}^{r_c} k_H(r) dr + \int_{r_{iH}}^{r_{oH}} k_H(r) dr \end{aligned}$$

The final integral is the phase for hydrogen at some energy E , and can be written as $\phi_H(E) = (n * + 1/2)\pi$. Designating the sum of the first two integrals by the phase $\delta\phi$, then we have

$$\delta\phi + (n * + 1/2)\pi = (n + 1/2)\pi \quad (2.51)$$

or

$$n* = n - \delta\phi/\pi \quad (2.52)$$

Hence, we can relate the quantum defect to a phase:

$$\delta_\ell = \left[\int_{r_i}^{r_c} k(r) dr - \int_{r_{iH}}^{r_c} k_H(r) dr \right] \frac{1}{\pi} \quad (2.53)$$

since it is clear from Eq. 2.47 and the fact that the turning point is determined by $E + V_{\text{eff}}(r_i) = 0$ that δ_ℓ approaches a constant as $E \rightarrow 0$.

Now we can find the bound state energies for the atom with a core; starting with Eq. 2.49,

$$\begin{aligned} n\pi &= \phi(E) - \pi/2 \\ &= \pi\delta_L(E) + \pi(hcR_H/E)^{1/2} \\ \Rightarrow E &= hcR_H/[n - \delta_L(E)]^2 \equiv \frac{hcR_H}{[n - \delta_\ell^{(0)} + \delta_\ell^{(1)}hcR_H/n]^2} \end{aligned} \quad (2.54)$$

thus we have explained the Balmer-Ritz formula (Eq. 2.45).

If we look at the radial Schrödinger equation for the electron ion core system in the region where $E > 0$ we are dealing with the scattering of an electron by a modified Coulomb potential (Mott & Massey Chapter 3). Intuitively one would expect that there would be an intimate connection between the bound state eigenvalue problem described earlier in this chapter and this scattering problem, especially in the limit $E \rightarrow 0$ (from above and below). Since the quantum defects characterize the bound state problem accurately in this limit one would expect that they should be directly related to the scattering phase shifts $\sigma_\ell(k)$ (k is the momentum of the *free* particle) which obey

$$\lim_{k \rightarrow 0} \cot[\sigma_\ell(k)] = \pi\delta_\ell^o \quad (2.55)$$

This has great intuitive appeal: $\pi\delta_\ell^o$ as discussed above is precisely the phase shift of the wave function with the core present relative to the one with $V = e^2/r$. On second thought Eq. 2.55 might appear puzzling since the scattering phase shift is customarily defined as the shift relative to the one with $V = 0$. The resolution of this paradox lies in the long range nature of the Coulomb interaction; it forces one to redefine the scattering phase shift, $\sigma_\ell(k)$, to be the shift relative to a pure Coulomb potential.

2.3.3 Quantum defects for a model atom

Now we give a calculation of a quantum defect for a potential $V(r)$ which is not physically realistic, but has only the virtue that it is easily solvable. The idea is to put an extra term in the potential which goes as $1/r^2$ so that the radial Schrödinger equation (Eq. 2.15) can be solved simply by adjusting ℓ . The electrostatic potential corresponds to having all the core electrons in a small cloud of size r_n (which is a nuclear size) which decays as an inverse power of r .

$$\phi_\ell(r) = \frac{e}{r} + \frac{(Z-1)er_n}{2r^2}$$

$$\begin{aligned}
\Rightarrow E(r) &= \frac{e}{r^2} + \frac{(Z-1)er_n}{r^3} \\
\Rightarrow Q_{\text{inside}}(r) &= e + \frac{(Z-1)er_n}{r} \\
\Rightarrow \rho(r) &= \frac{dQ/dr}{4\pi r^2} = -(Z-1)er_n/4\pi r^4
\end{aligned} \tag{2.56}$$

At $r = r_n$, $Q_{\text{inside}}(r_n) = Ze$. We presume r_n is the nuclear size and pretend that it is so small we don't have to worry about what happens inside it.

When the potential $V(r) = -e\phi(r)$ is substituted in Eq. 2.17, one has

$$\frac{d^2y}{dr^2} + \left[\frac{2\mu}{\hbar^2} \left[E_n + \frac{e^2}{r} \right] - \frac{\ell(\ell+1) - A}{r^2} \right] y = 0 \tag{2.57}$$

$$\text{where } A = \frac{2\mu(Z-1)e^2r_n}{\hbar^2} \tag{2.58}$$

If one now defines

$$\ell'(\ell'+1) = \ell(\ell+1) - A \tag{2.59}$$

then $\ell' < \ell$ since $A > 0$, and one can write

$$\ell' = \ell - \delta_\ell \tag{2.60}$$

Substituting Eq. 2.59 in Eq. 2.57 gives the radial Schrödinger equation for hydrogen, (Eq. 2.15), except that ℓ' replaces ℓ ; eigenvalues occur when (see Eq. 2.30)

$$v = n' + \ell' + 1 = n^* \quad n' = 0, 1, 2, \dots \tag{2.61}$$

where n' is an integer. Using $n = n' + \ell + 1$ as before, we obtain the corresponding eigenvalues as

$$E_n = hcR_H/n^{*2} = hcR_H/(n - \delta_\ell)^2 \tag{2.62}$$

The quantum defect is independent of E

$$\delta_\ell = \ell - \ell' = n - n^* \tag{2.63}$$

Eq. 2.59 may be solved for δ_ℓ using the standard quadratic form. Retaining the solution which $\rightarrow 0$ as $Z \rightarrow 1$, gives

$$\delta_\ell = (\ell + 1/2) - \sqrt{(\ell + 1/2)^2 - A} \approx \frac{A}{2\ell + 1} \text{ for } A \ll \ell + 1/2 \tag{2.64}$$

This shows that $\delta_\ell \rightarrow 0$ as $\ell \rightarrow \infty$.

In contrast to the predictions of the above simple model, quantum defects for realistic core potentials decrease much more rapidly with increasing ℓ [for example as $(\ell + 1/2)^{-3}$] and generally exhibit $\delta_\ell^{(0)}$ close to zero for all ℓ greater than the largest ℓ value occupied by electrons in the core.

2.4 Metrology and Precision Measurement and Units

As scientists we take the normal human desire to understand the world to quantitative extremes. We demand agreement of theory and experiment to the greatest accuracy possible. We measure quantities way beyond the current level of theoretical understanding in the hope that this measurement will be valuable as a reference point or that the difference between our value and some other nearly equal or simply related quantity will be important. The science of measurement, called metrology, is indispensable to this endeavor because the accuracy of measurement limits the accuracy of experiments and their intercomparison. In fact, the construction, intercomparison, and maintenance of a system of units is really an art (with some, a passion), often dependent on the latest advances in the art of physics (e.g. quantized Hall effect, cold atoms, trapped particle frequency standards).

As a result of this passion, metrological precision typically marches forward a good fraction of an order of magnitude per decade. Importantly, measurements of the same quantity (e.g. α or e/h) in different fields of physics (e.g. atomic structure, QED, and solid state) provide one of the few cross-disciplinary checks available in a world of increasing specialization. Precise null experiments frequently rule out alternative theories, or set limits on present ones. Examples include tests of local Lorentz invariance, and the equivalence principle, searches for atomic lines forbidden by the exclusion principle, searches for electric dipole moments (which violate time reversal invariance), and searches for a “fifth [gravitational] force”.

A big payoff, often involving new physics, sometimes comes from attempts to achieve routine progress. In the past, activities like further splitting of the line, increased precision and trying to understand residual noise have led to the fine and hyperfine structure of H , anomalous Zeeman effect, Lamb shift, and the discovery of the 3K background radiation. One hopes that the future will bring similar surprises. Thus, we see that precision experiments, especially involving fundamental constants or metrology not only solidify the foundations of physical measurement and theories, but occasionally open new frontiers.

This chapter deals briefly with SI units (and its ancestor, the mks system), then systems of units which might appear more natural to a physicist, and then introduces atomic units, which will be used in this course.

2.4.1 Dimensions and Dimensional Analysis

Oldtimers were brought up on the mks system - meter, kilogram, and second. This simple designation emphasized an important fact: three dimensionally independent units are sufficient to span the space of all physical quantities. The dimensions are respectively ℓ - length, m -mass, and t -time. These three dimensions suffice because when a new physical quantity is discovered (e.g. charge, force) it always obeys an equation which permits its definition in terms of m, k , and s . Some might argue that fewer dimensions are necessary (e.g. that time

and distance are the same physical quantity since they transform into each other in moving reference frames); we'll keep them both, noting that the definition of length is now based on the speed of light. Practical systems of units have additional units beyond those for the three dimensions, and often additional "as defined" units for the same dimensional quantity in special regimes (e.g. x-ray wavelengths or atomic masses).

Dimensional analysis consists simply in determining for each quantity its dimension along the three dimensions (seven if you use the SI system rigorously) of the form

$$\text{Dimension } (G) = [G] = m^{-1}l^3t^{-2}, \text{ where } [x] \equiv \text{dimensions of } x.$$

Dimensional analysis yields an estimate for a given unknown quantity by combining the known quantities so that the dimension of the combination equals the dimension of the desired unknown. The art of dimensional analysis consists in knowing whether the wavelength or height of the water wave (both with dimension l) is the length to be combined with the density of water and the local gravitational acceleration to predict the speed of the wave.

2.4.2 SI units

A single measurement of a physical quantity, by itself, never provides information about the physical world, but only about the size of the apparatus or the units used. In order for a single measurement to be significant, some other experiment or experiments must have been done to measure these "calibration" quantities. Often these have been done at an accuracy far exceeding our single measurement so we don't have to think twice about them. For example, if we measure the frequency of a rotational transition in a molecule to six digits, we have hardly to worry about the calibration of the frequency generator if it is a high accuracy model that is good to nine digits. And if we are concerned we can calibrate it with an accuracy of several more orders of magnitude against station WWV or GPS satellites which give time valid to 20 ns or so.

Time/frequency is currently the most accurately measurable physical quantity and it is relatively easy to measure to 10^{-12} . In the SI (System International, the agreed-upon systems of weights and measures) the second is defined as 9,192,631,770 periods of the Cs hyperfine oscillation in zero magnetic field. Superb Cs beam machines at places like NIST-Boulder provide a realization of this definition at about 10^{-14} . Frequency standards based on laser- and evaporatively cooled atoms and ions are doing several orders of magnitude better owing to the longer possible measurement times and the reduction of Doppler frequency shifts. For example, recent experiments on degenerate samples of 100,000 fermionic ^{87}Sr trapped in an optical lattice at 100 nK are reaching fractional frequency uncertainty below 10^{-20} [?]. The expected clock stability is estimated at 3.1×10^{-18} in 1 second. In comparison, the age of the Universe is a few 10^{17} s. To give you an idea of the challenges inherent in reaching this level

of precision, if you connect a 10 meter coaxial cable to a frequency source good at the 10^{-16} level, the frequency coming out the far end in a typical lab will be an order of magnitude less stable - can you figure out why?

The meter was defined at the first General Conference on Weights and Measures in 1889 as the distance between two scratches on a platinum-iridium bar when it was at a particular temperature (and pressure). Later it was defined more democratically, reliably, and reproducibly in terms of the wavelength of a certain orange krypton line. Most recently it has been defined as the distance light travels in $1/299,792,458$ of a second. This effectively defines the speed of light, but highlights the distinction between defining and realizing a particular unit. Must you set up a speed of light experiment any time you want to measure length? No: just measure it in terms of the wavelength of a He-Ne laser stabilized on a particular hfs component of a particular methane line within its tuning range; the frequency of this line has been measured to about a part in 10^{-11} and it may seem that your problem is solved. Unfortunately the reproducibility of the locked frequency and problems with diffraction in your measurement both limit length measurements at about 10^{-11} .

A list of spectral lines whose frequency is known to better than 10^{-9} is available on the homepage of the Bureau International de Poids et Mesures ⁴. Also see the publication on the International System of Units (SI), 2019 by NIST ⁵, and the latest revision of the fundamental physical constants CODATA2018 NIST Reference on Constants, Units, and Uncertainty ⁶.

The third basic unit of the SI system is the kilogram, which until 2018 was the only fundamental SI base unit still defined in terms of an artifact - in this case a platinum-iridium cylinder kept in clean air at the Bureau de Poids et Measures in Sevres, France. The dangers of mass change due to cleaning, contamination, handling, or accident are so perilous that this cylinder has been compared with the dozen secondary standards that reside in the various national measurement laboratories only two times in the last century. Clearly it was one of the major challenges for metrology to replace this artifact kilogram with an atomic definition. One way would have been to simply define Avogadro's number. While atomic mass can be measured to 10^{-11} , there is currently no sufficiently accurate method of realizing such a definition, however. (The unit of atomic mass, designated by m_u is $1/12$ the mass of a ^{12}C atom.) The solution adopted by the 26th General Conference on Weights and Measures (CGPM) in 2018 was instead to define the Planck constant h to be exactly $6.626\ 070\ 15 \times 10^{-34}$ J s. So, h is no longer something we measure, it is defined.

There are four more base units in the SI system - the ampere, kelvin, mole, and the candela - for a total of seven. While three are sufficient (or more than sufficient) to do physics, the other four reflect the current situation that electrical quantities, atomic mass, temperature, and luminous intensity can be and

⁴<https://www.bipm.org/en/publications/mises-en-pratique/standard-frequencies>

⁵<https://doi.org/10.6028/NIST.SP.330-2019>

⁶<https://physics.nist.gov/cuu/Constants/>

are regularly measured with respect to auxiliary standards at levels of accuracy greater than those which can be expressed in terms of the above three base units. Thus measurements of Avogadro's constant, the Boltzmann constant, or the mechanical equivalents of electrical units play a role of interrelating the base units of mole (number of atoms of ^{12}C in 0.012 kg of ^{12}C), kelvin, or the new volt and ohm (defined in terms of Josephson and quantized Hall effects respectively). In fact independent measurement systems exist for other quantities such as x-ray wavelength (using diffraction from calcite or other standard crystals), but these other de facto measurement scales are not formally sanctified by the SI system.

At the 26th CGPM a major step away from artifacts and towards universality was taken, in defining the values of Planck's constant, elementary charge, Boltzmann's constant, the Avogadro constant and to be complete, the luminous efficacy of monochromatic radiation of frequency 540×10^{12} Hz.

We can now simply say, quoting NIST publication 330 (see footnote above) “The International System of Units, the SI, is the system of units in which

- the unperturbed ground state hyperfine transition frequency of the cesium 133 atom $\Delta\nu_{\text{Cs}}$ is $9\ 192\ 631\ 770$ Hz,
- the speed of light in vacuum c is $299\ 792\ 458$ m/s,
- the Planck constant h is $6.626\ 070\ 15 \times 10^{-34}$ J s,
- the elementary charge e is $1.602\ 176\ 634 \times 10^{-19}$ C,
- the Boltzmann constant k is $1.380\ 649 \times 10^{-23}$ J/K,
- the Avogadro constant N_A is $6.022\ 140\ 76 \times 10^{23}$ mol $^{-1}$,
- the luminous efficacy of monochromatic radiation of frequency 540×10^{12} Hz, K_{cd} , is 683 lm/W,

where the hertz, joule, coulomb, lumen, and watt, with unit symbols Hz, J, C, lm, and W, respectively, are related to the units second, meter, kilogram, ampere, kelvin, mole, and candela, with unit symbols s, m, g, A, K, mol, and cd, respectively, according to $\text{Hz} = \text{s}^{-1}$, $\text{J} = \text{kgm}^2\text{s}^{-2}$, $\text{C} = \text{As}$, $\text{lm} = \text{cdm}^2\text{m}^{-2} = \text{cdsr}$, and $\text{W} = \text{kgm}^2\text{s}^{-3}$. The numerical values of the seven defining constants have no uncertainty.”

This definition is nothing short of a paradigm shift, away from particular realizations of the unit, i.e. some metal bar in Paris, and more towards a permanent definition that will remain intact even as measurement capabilities advance. There are still differences in the various constants. The Planck constant h and the speed of light in vacuum c can be seen as fundamental, determining quantum effects and the properties of space-time. Choosing an atomic transition like $\Delta\nu_{\text{Cs}}$ is less fundamental in character, and there are of course difficulties in implementing this definition in the lab, given field fluctuations and density-dependent shifts. At least, it is a universal reference, and for all we know constant in time, compared to e.g. a standard based on the rotation of the earth (which slows down over time).

The electric charge is connected to the strength of the electromagnetic force through the fine-structure constant α . For someone used to the *cgs* system, a hilarious apparent puzzle poses itself, as that person grew up learning $\alpha = \frac{e^2}{\hbar c}$. But now all constants appearing on the right of the equation are defined - so is α now no longer something we should measure, but it's forever fixed by the members of the CGPM 2018? Actually, no, but it gets first more confusing. In the SI system of units, remember that $\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c}$, where ϵ_0 is the vacuum electric permittivity. Does that help? One might complain that it doesn't, since Maxwell taught us that $\epsilon_0\mu_0 = 1/c^2$, which is defined, and even people growing up with the SI system learned that the vacuum magnetic permeability $\mu_0 = 4\pi \times 10^{-7}$. Are we again concluding that α was suddenly defined in 2018? No, the hilarious answer, to put it as a joke, is that "4 π now has an error bar". More seriously, μ_0 is no longer exactly $4\pi \times 10^{-7}$! Performing more and more precise experiments that connect electric charge to length, kilogram and second, for example the "Watt balance" (Kibble balance)⁷, will lead to a more precise value of μ_0 that's just not quite $4\pi \times 10^{-7}$.

2.4.3 Metrology

The preceding discussion gives a rough idea of the definitions and realizations of SI units, and some of the problems that arise in trying to define a unit for some physics quantity (e.g. mass) that will work across many orders of magnitudes. However, it sidesteps questions of the border between metrology and precision measurements. (Here we have used the phrase "precision measurement" colloquially to indicate an accurate absolute measurement; if we were verbally precise, precision would imply only excellent relative accuracy.)

The Boltzmann constant for example gives the ratio of energy scales defined by the first three base units on the one hand and thermal energy that we call one Kelvin. The Kelvin used to be defined by taking the triple point of water as a fundamental fixed point and assigned to it the temperature 273.16 K. That had practical difficulties. With k fixed, the triple point of water is now a thermal energy that needs to be measured and that has an error bar. We are hardly learning any fundamental new physics from such a measurement, which is clearly of metrological character. Similarly, measuring the hfs frequency of Cs would be a metrological experiment in that it would only determine the length of the second.

If we measure the hfs frequency of H with high accuracy, this might seem like a physics experiment since this frequency can be predicted theoretically. Unfortunately theory runs out of gas at about 10^{-6} due to lack of accurate knowledge about the structure of the proton (which causes a 42 ppm shift). Any digits past this are just data collection until one gets to the 14th, at which point one becomes able to use a H maser as a secondary time standard. This has stability advantages over Cs beams for time periods ranging from seconds to days and so might be metrologically useful – in fact, it is widely used in very long baseline

⁷The Kibble balance is a beam balance, with a test mass, coil and magnetic field on one side and a counter balancing mass on the other. A feedback current passing through the coil generates a force that balances the beam.

radio astronomy.

One challenge to using Nature’s “quantized units” such as angular momentum (\hbar) and charge (e), and mass itself (which is quantized, although not so simply), realized through Avogadro’s constant (N_A^{-1} could have been thought of as the mass quantum in grams) was that measurements of these “constants” in the pre-2018 SI units are only accurate to about 10^{-8} , well below the accuracy of the realization of the base units of the SI at the time. In 2018 one turned the challenge around. Instead of defining, for example, the mass of the prototype kg in Paris to be exactly 1 kg, and leaving \hbar to be measured, now \hbar is fixed and the mass of the prototype has to be measured - it now has an error bar.

To illustrate the intricacies one encounters, consider measurement of the Rydberg constant,

$$Ry_\infty = \frac{m_e e^4}{(4\pi)^3 \epsilon_0^2 \hbar^3 c} \quad (2.65)$$

a quantity which is an inverse length (an energy divided by $\hbar c$), closely related to the number of wavelengths per cm of light emitted by a H atom (the units are often given in spectroscopists’ units, cm^{-1} , to the dismay of SI purists. Remember the conversion: $1 \text{ cm}^{-1} \approx 30 \text{ GHz}$). Clearly such a measurement determines a linear combination of the now defined fundamental quantities (and μ_0 , which now has an error bar), which, as we said, were realized pre-2018 only to 10^{-8} accuracy. But Ry_∞ has recently been measured by several labs with results that agree to 10^{-11} . (In fact, the quantity that is measured is the Rydberg *frequency* for hydrogen, $c \times Ry_H$, since there is no way to measure wavelength to such precision.) This example illustrates a fact of life of precision experiments: with care you can trust the metrological realizations of physical quantities at accuracies to 10^{-7} ; beyond that the limit on your measurement may well be partly metrological. In that case, what you measure is not in general clearly related to one single fundamental constant or metrological quantity. The importance of your experiment is determined by the size of its error bar relative to the uncertainty of all other knowledge about the particular linear combination of fundamental and metrological variables that you have measured.

2.5 Universal Units and Fundamental Constants

The sizes of the meter, kilogram and second were originally selected for convenience. They bear no relationship to things which most physicists would regard as universal or fundamental. Given the arbitrary scale of the “convenient units”, physicists generally prefer to use systems of units that are natural for the problem at hand. We will now discuss Planck units, which seem most universal, and Atomic units fulfill this role when discussing atoms and their interactions with light. But let us start at the beginning by considering “are some of the fundamental constants listed by CODATA 2018⁸ more fundamental or universal?”

To begin, we assert that the three most universal constants are:

⁸<https://physics.nist.gov/cuu/Constants/>

- c - the speed of light
- \hbar - the quantum of action
- G - the gravitational constant

These quantities involve light, the quantum and gravitation. Since \hbar can be related to measurements on light, these constants do not depend on the existence of quantized matter (only quantized radiation).

Next, we come to the atomic constants, whose magnitude is determined by the size of the quantized matter which we find all around us. Clearly, the most fundamental of these is:

e - the quantum of charge.

because it is the same (except for the sign) for all particles.

Even though the existence of quantized charge seems independent of the physics which underlies the construction of universal units, this is probably not the case because charge does not have independent units. In fact, the fine structure constant

$$\alpha \equiv \frac{1}{4\pi\epsilon_0}\frac{e^2}{\hbar c} = (137.036\cdots)^{-1} \quad (2.66)$$

is arguably the only fundamental constant truly worthy of that name in atomic physics. High energy physicists regard α as the not so fundamental coupling strength of the electromagnetic interaction – At higher energies and correspondingly smaller distances, it gets larger as we approach the “bare charge” more closely. When we really understand E&M, QED, and the origins of quantized matter, we should be able to predict it. The fact that we have to measure it is a sign of our ignorance, but the good agreement of the many seemingly independent ways of measuring it in different subfields of physics shows that we begin to understand some things.

Other atomic constants like masses (m_e, m_p, m_n, \dots) and magnetic moments (u_e, u_p, u_n) seem to be rather arbitrary at our current level of physical knowledge except that certain relationships are given by QED ($\mu_e = -g_e e \hbar / (2m)$ with g_e a little more than 2) and quark models of the nucleons.

2.6 Atomic Units

Formulae in these notes will be displayed with the factors of \hbar, e, m (m is the electron mass) etc., factored into atomic units to facilitate interpretation; numerical evaluations may be done using either atomic units or Gaussian-esu units. The system of atomic units (a.u.) is defined by setting $\hbar = m = e = 1$. Units for other physical quantities are formed by dimensionally suitable combinations of these units. When their use is frequent, they are given names (analogous to erg, dyne, etc., in cgs). In atomic units, the units of length and energy are the

Quantity	Symbol	Value (to a few digits)	Name
*charge	e	$1.6 \times 10^{-19} C$	
*angular Momentum	\hbar	$1.05 \times 10^{-34} \text{ J s}$	
*mass	m	$0.910 \times 10^{-31} \text{ kg}$	
length	a_o	$\hbar^2/m(e^2/4\pi\epsilon_o)^2 = 0.53 \times 10^{-11} \text{ m}$	bohr
energy	H	$m(e^2/4\pi\epsilon_o)^2/\hbar^2 = 4.36 \times 10^{-18} \text{ J}$ $\Rightarrow 27.2 \text{ eV} \Rightarrow 2 \text{ Ry} = 2.2 \times 10^7 \text{ m}^{-1}$	hartree
time	-	$\hbar^3/m(e^2/4\pi\epsilon_o)^2 = 2.42 \times 10^{-17} \text{ s}$	
velocity	-	$(e^2/4\pi\epsilon_o)/\hbar = \alpha_c = 2.2 \times 10^6 \text{ m/s}$	
magnetic moment	μ_B	$e\hbar/2m = 1.4 \times 10^4 \text{ MHz T}^{-1}$	Bohr magneton
electric field	-	$e/a_o^2 = 5.14 \times 10^{11} \text{ V/m}$	

*Charge, Angular Momentum and Mass are the *Basic Units*

most important - they are called the Bohr (a_o) and Hartree (H) respectively. Expressions for energy and length can generally be expressed in terms of H or a_o and powers of the dimensionless constant α .

2.7 The Fine Structure Constant

$$\alpha = \frac{e^2}{4\pi\epsilon_o\hbar c} = (137.0606 \dots)^{-1} \quad (2.67)$$

is ubiquitous in atomic physics. The name, fine structure, reflects the appearance of this quantity (squared) in the ratio of the hydrogenic fine structure splitting to the Rydberg:

$$\frac{\Delta(\text{fine structure})}{\text{Rydberg}} = \alpha^2 \frac{Z^4}{n^3 \ell(\ell+1)} \quad (2.68)$$

The fine structure constant will often crop up as the ratio between different physical quantities having the same dimensions. An impressive example of this is length: the “hydrogen wavelength” (1/Rydberg), Bohr radius, Compton wavelength, and classical radius of the electron are all related by powers of α .

References

- [1] R.R. Freeman and D. Kleppner, Phys. Rev. A. (1967).
- [2] M.J. Seaton, Proc. Phys. Soc. **88**, 801 (1966).
- [3] R.G. Parsons and V.F. Weisskopf, Z. Phys. **202**, 492 (1967).
- [4] E.R. Cohen and B.N. Taylor, Rev. Mod. Phys. **72**, No. 2 (2000).
- [5] Bothwell, T., et al., Nature **602**, 420 (2022)

Chapter 3

Fine Structure and Lamb Shift

3.1 Fine Structure

Immediately adjacent to Michelson and Morley's announcement of their failure to find the ether in an 1887 issue of the Philosophical Journal is a paper by the same authors reporting that the H _{α} line of hydrogen is actually a doublet, with a separation of 0.33 cm⁻¹. In 1915 Bohr suggested that this "fine structure" of hydrogen is a relativistic effect arising from the variation of mass with velocity. Sommerfeld, in 1916, solved the relativistic Kepler problem and using the old quantum theory, as it was later christened, accounted precisely for the splitting. Sommerfeld's theory gave the lie to Einstein's dictum "The Good Lord is subtle but not malicious", for it gave the right results for the wrong reason: his theory made no provision for electron spin, an essential feature of fine structure. Today, all that is left from Sommerfeld's theory is the fine structure constant $\alpha = e^2/\hbar c$.

The theory for the fine structure in hydrogen was provided by Dirac whose relativistic electron theory (1926) was applied to hydrogen by Darwin and Gordon in 1928. They found the following expression for the energy of an electron bound to a charged nucleus of charge Z and of infinite mass:

$$\frac{E}{mc^2} = \frac{1}{\sqrt{1 + \left(\frac{\alpha Z}{n - k + \sqrt{k^2 - \alpha^2 Z^2}} \right)^2}} \quad (3.1)$$

where n is the principal quantum number, $k = j + 1/2$, and $j = \ell \pm 1/2$. One may say that this equation is "too exact", in the sense that already at low powers in the expansion of α , physics beyond the single-particle Dirac equation comes in - QED corrections such as the Lamb shift, and corrections due to the g -factor of the electron not being exactly 2. Let us therefore rather expand this expression to the lowest "beyond Bohr" correction. The expansion parameter is $Z\alpha$, which indeed gives the scale of velocity of the electron, in units of the speed of light. We find

$$E = mc^2 - \frac{1}{2}mc^2\alpha^2Z^2\frac{1}{n^2} - \frac{1}{2}mc^2\alpha^4Z^4\frac{1}{n^4} \left(\frac{n}{j + 1/2} - \frac{3}{4} \right) + \dots \quad (3.2)$$

The first term is the rest energy of the electron, the second term is Bohr's result, and the third term is the fine structure.

As we see, the fine structure correction lifts the degeneracy of energy levels of differing total angular momentum, and it results, for non s -states, in doublets, according to $j = l + 1/2$ and $j = l - 1/2$.

To understand the origin of the fine structure splitting, it is instructive to consider the Pauli equation, the approximation to the Dirac equation to the

lowest order in v/c . It is a Schrödinger equation for spinors with Hamiltonian

$$H = mc^2 + \frac{p^2}{2m} - \frac{e^2}{r} + H_{FS} \quad (3.3)$$

The first term is the electron's rest energy; the following two terms are the non-relativistic Hamiltonian, and the last term, the fine structure interaction, is given by

$$H_{FS} = -\frac{p^4}{8m^3c^2} + \frac{\hbar^2 e^2}{2m^2 c^2} \frac{1}{r^3} \mathbf{L} \cdot \mathbf{S} - \frac{\hbar^2}{8m^2 c^2} \nabla^2 \frac{e^2}{r} \quad (3.4)$$

The relativistic contributions can be described as the *kinetic*, *spin-orbit*, and *Darwin* terms, H_{kin} , H_{so} , and H_{Dar} , respectively. Each has a straightforward physical interpretation.

3.1.1 Kinetic contribution

Relativistically, the total electron energy is $E = \sqrt{(mc^2)^2 + (pc)^2}$. The kinetic energy is

$$T = E - mc^2 = mc^2 \left(\sqrt{1 + \frac{p^2}{m^2 c^2}} - 1 \right) = \frac{p^2}{2m} - \frac{1}{8} \frac{p^4}{m^3 c^2} + \dots \quad (3.5)$$

Thus

$$H_{\text{kin}} = -\frac{1}{8} \frac{p^4}{m^3 c^2} \quad (3.6)$$

Using first order perturbation theory, the kinetic energy correction is

$$\begin{aligned} \Delta E_{nlm} &= - \int d^3 \mathbf{r} \psi_{nlm}^* \left(\frac{\hbar^4}{8m^3 c^2} \right) \nabla^4 \psi_{nlm} \\ &= -\frac{mc^2}{2} \frac{\alpha^4}{n^4} \left(\frac{n}{l+1/2} - \frac{3}{4} \right). \end{aligned}$$

Note that there is l dependence, which makes sense due to the momentum to the fourth term. This is not yet dependent on j (which we would expect from the Dirac equation solution), but this is fine as we are only considering one term so far.

3.1.2 Spin-Orbit Interaction

According to the Dirac theory the electron has intrinsic angular momentum $\hbar \mathbf{S}$ and a magnetic moment $\boldsymbol{\mu}_e = -g_e \mu_0 \mathbf{S}$. The electron g-factor is $g_e = 2$ in Dirac theory (QED gives corrections to g_e on the order of α). As the electron moves through the electric field of the proton it "sees" a motional magnetic field (see e.g. Purcell, *Electricity and Magnetism*)

$$\mathbf{B}_{\text{mot}} = -\frac{\mathbf{v}}{c} \times \mathbf{E} = -\frac{\mathbf{v}}{c} \times \frac{e}{r^3} \mathbf{r} = \frac{e\hbar}{mc} \frac{1}{r^3} \mathbf{L} \quad (3.7)$$

where $\hbar \mathbf{L} = \mathbf{r} \times m\mathbf{v}$. However, there is another contribution to the effective magnetic field arising from the Thomas precession.

The relativistic transformation of a vector between two moving co-ordinate systems which are moving with different velocities involve not only a dilation, but also a rotation (cf Jackson, *Classical Electrodynamics*). The rate of rotation, the Thomas precession, is

$$\boldsymbol{\Omega}_T = \frac{1}{2} \frac{\mathbf{a} \times \mathbf{v}}{c^2} \quad (3.8)$$

Note that the precession vanishes for co-linear acceleration. However, for a vector fixed in a co-ordinate system moving around a circle, as in the case of the spin vector of the electron as it circles the proton, Thomas precession occurs. From the point of view of an observer fixed to the nucleus, the precession of the electron is identical to the effect of a magnetic field.

$$\mathbf{B}_T = \frac{1}{\gamma_e} \boldsymbol{\Omega}_T. \quad (3.9)$$

Substituting $\gamma_e = e/mc$, and $\mathbf{a} = -e^2 \mathbf{r}/mr^3$ into Eq. 3.8 gives

$$\mathbf{B}_T = -\frac{1}{2} \frac{e\hbar}{mc} \frac{1}{r^3} \mathbf{L} \quad (3.10)$$

Hence the total effective magnetic field is

$$\mathbf{B}' = \frac{1}{2} \frac{e\hbar}{mc} \frac{1}{r^3} \mathbf{L} \quad (3.11)$$

This gives rise to a total spin-orbit interaction

$$H_{so} = -\boldsymbol{\mu} \cdot \mathbf{B}' = \frac{e^2 \hbar^2}{2m^2 c^2} \frac{1}{r^3} \mathbf{S} \cdot \mathbf{L} \quad (3.12)$$

A non-mathematical argument for Thomas precession

This is based on an argument from Purcell.

Consider an object traveling in a circle. Approximate the circle as a polygon with N sides. Let the object be pointing along the side it is traveling on, thus rotating by angle $\frac{2\pi}{N}$ at each corner. In the object's frame, there will be a Lorentz contraction along the direction of motion, which leads to each angle appearing smaller. Let the length of the next segment along the direction of the current segment be L and the transverse length be w . Then the angle in the object's frame is

$$\theta' = \frac{w}{L/\gamma} = \gamma\theta = 2\pi \frac{\gamma}{N}.$$

Therefore, there is a change in the angle in the object's frame of

$$\Delta\theta = 2\pi(\gamma - 1).$$

In the limit $N \rightarrow \infty$, we find an extra precession of the object

$$\frac{\Omega_p}{\omega} = \gamma - 1 \approx \frac{1}{2} \frac{v^2}{c^2},$$

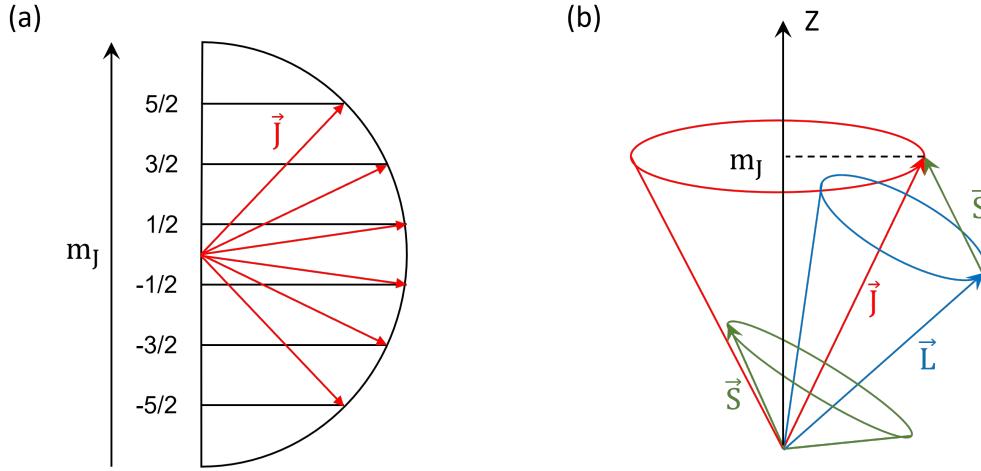


Figure 9. (a) The quantization of m_J for $J = 5/2$. (b) The vectorial representation of coupling \mathbf{L} and \mathbf{S} to form $\mathbf{J} = \mathbf{L} + \mathbf{S}$.

where the final approximation is the non-relativistic limit.

So, the Thomas precession is

$$\boldsymbol{\Omega}_T = \frac{1}{2} \frac{\mathbf{a} \times \mathbf{v}}{c^2},$$

which will appear as a fictitious magnetic field for the electron:

$$\mathbf{B}_T = \frac{1}{\gamma_e} \boldsymbol{\Omega}_T = \frac{1}{\gamma_e} \left(-\frac{1}{2} \frac{e^2 \mathbf{r}}{mr^3} \times \mathbf{v} \frac{1}{c^2} \right) = -\frac{1}{2} \mathbf{B}_{\text{motion}},$$

using that $\mathbf{a} = -\frac{e^2 \mathbf{r}}{mr^3}$ must be the centrifugal acceleration and $\gamma_e = e/mc$.

Thus, the total magnetic field seen by the electron is

$$\mathbf{B}_L = \mathbf{B}_{\text{motion}} + \mathbf{B}_T = \frac{1}{2} \mathbf{B}_{\text{motion}}.$$

The spin-orbit coupling term is thus

$$H_{SO} = -\mu_{\text{electron}} \cdot \mathbf{B}_L = \frac{1}{2} \frac{e^2 \hbar^2}{m^2 c^2} \frac{1}{r^3} \mathbf{S} \cdot \mathbf{L}.$$

This magnetic field has a magnitude of approximately 1 T.

Solving Spin-Orbit Coupling

Recall that for $H_{SO} \propto \mathbf{S} \cdot \mathbf{L}$, \mathbf{S} rotates about \mathbf{L} and \mathbf{L} rotates about \mathbf{S} . This actually results in procession of \mathbf{L} and \mathbf{S} about the total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$. This suggests that these operators are not good quantum numbers.

From evaluating the binomial, we can see that

$$\mathbf{S} \cdot \mathbf{L} = \frac{1}{2} (\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2).$$

Let j be the eigenvalue of \mathbf{J}^2 , which has values $j = |l - s|, \dots, |l + s|$. Since

$s = \frac{1}{2}$, $j = l - \frac{1}{2}$ or $j = l + \frac{1}{2}$. Therefore,

$$\mathbf{S} \cdot \mathbf{L} = \frac{1}{2} (\mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2) = \frac{1}{2} (j(j+1) - l(l+1) - s(s+1)) = \frac{1}{2} \left(j(j+1) - l(l+1) - \frac{3}{4} \right).$$

So, j is a good quantum number. We can write the Hamiltonian term as

$$H_{SO} = \xi(r) \mathbf{L} \cdot \mathbf{S}, \quad \xi(r) = \frac{1}{2} \frac{e^2 \hbar^2}{m^2 c^2} \frac{1}{r^3}.$$

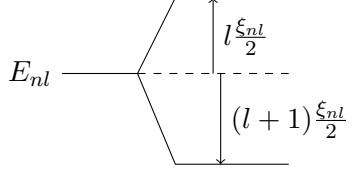
The correction is then

$$\begin{aligned} \Delta E_j &= \langle \mathbf{L} \cdot \mathbf{S} \rangle_j \langle \xi(r) \rangle_{nl} \\ &= \frac{1}{2} \left(j(j+1) - l(l+1) - \frac{3}{4} \right) \xi_{nl}. \end{aligned}$$

This gives us the doublet

$$\Delta E_{l+1/2} = l \frac{\xi_{nl}}{2} \text{ and } \Delta E_{l-1/2} = -(l+1) \frac{\xi_{nl}}{2}.$$

So, we get a splitting of the Bohr level E_{nl} into two terms.



Example | 2p of hydrogen In the 2p state of hydrogen, $l = 1$. So, the $2p_{3/2}$ state is raised by $1\xi_{2,1}/2$ and the $2p_{1/2}$ state is lowered by $-2\xi_{2,1}/2$. There are 4 states in the $3/2$ level (four m_j values) and 2 states in the $1/2$ level (two m_j values). Is the average energy still the same?

$$4 \times (+1) + 2 \times (-2) = 0.$$

So yes, the center of “mass” remains zero.

The average energy of the shifted states should not change. We can see this by considering the average of $\mathbf{L} \cdot \mathbf{S}$: by the symmetry \mathbf{L} and \mathbf{S} , this should be zero.

Finally, we need to compute ξ_{nl} .

$$\begin{aligned} \frac{1}{2} \xi_{nl} &= \frac{e^2 \hbar^2}{4m^2 c^2} \left\langle \frac{1}{r^3} \right\rangle_{nl} \\ &= \frac{e^2 \hbar^2}{4m^2 c^2} \frac{1}{a_0^3} \frac{1}{n^3 l (l+1/2) (l+1)} \end{aligned}$$

where in the second step we recall the expression for $\langle 1/r^3 \rangle_{nl}$ given in the hydrogen section (2.2.4). The angular momentum dependence can be understood intuitively from the centrifugal potential: a larger l implies the electron spends less time near the origin.

Putting the pieces together,

$$E_{SO} = \frac{mc^2}{4} \frac{\alpha^4}{n^4} \frac{n}{l(l+1/2)(l+1)} \left(j(j+1) - l(l+1) - \frac{3}{4} \right).$$

3.1.3 The Darwin Term

Electrons exhibit “Zitterbewegung”, fluctuations in position on the order of the Compton wavelength, \hbar/mc . As a result, the effective Coulombic potential is not $V(\mathbf{r})$, but some suitable average $\bar{V}(\mathbf{r})$, where the average is over the characteristic distance \hbar/mc . To evaluate this, expand $V(\mathbf{r})$ about \mathbf{r} in terms of a displacement \mathbf{s} ,

$$V(\mathbf{r} + \mathbf{s}) = V(\mathbf{r}) + \nabla V \cdot \mathbf{s} + \frac{1}{2} \sum_{ij} \mathbf{s}_{xi} \mathbf{s}_{xj} \frac{\partial^2 V}{\partial x_i \partial x_j} + \dots \quad (3.13)$$

Assume that the fluctuations are isotropic, so that the average $\langle s_x^2 \rangle = \frac{1}{3} \langle \mathbf{s}^2 \rangle = \frac{1}{3} \left(\frac{\hbar}{mc} \right)^2$ and $\langle s_x s_y \rangle = 0$ etc. Then the time average of $V(\mathbf{r} + \mathbf{s}) - V(\mathbf{r})$ is

$$\Delta V \sim \frac{1}{2} \left[\frac{1}{3} \left(\frac{\hbar}{mc} \right)^2 \right] \nabla^2 V = -\frac{1}{6} \frac{e^2 \hbar^2}{m^2 c^2} \nabla^2 \left(\frac{1}{r} \right) \quad (3.14)$$

The precise expression for the Darwin term is

$$H_{\text{Dar}} = -\frac{1}{8} \frac{e^2 \hbar^2}{m^2 c^2} \nabla^2 \left(\frac{1}{r} \right) \quad (3.15)$$

So we see that the exact prefactor is $1/8$ rather than $1/6$, fairly close to our rough estimate. Evaluating the Laplacian $\nabla^2 \frac{1}{r} = -4\pi\delta(\mathbf{r})$, the Darwin term becomes

$$H_{\text{Dar}} = \frac{\pi}{2} \frac{e^2 \hbar^2}{m^2 c^2} \delta(\mathbf{r}).$$

For the $l = 0$ states,

$$\langle H_{\text{Darwin}} \rangle_{n00} = \frac{\pi}{2} \frac{e^2 \hbar^2}{m^2 c^2} |\psi_{n00}(0)|^2 = \frac{1}{2} \frac{e^2 \hbar^2}{m^2 c^2} \frac{1}{n^3 a_0^3} = \frac{1}{2} mc^2 \frac{\alpha^4}{n^3}.$$

Note that the Darwin term acts to make ${}^2S_{1/2}$ and ${}^2P_{1/2}$ degenerate.

3.1.4 Evaluation of the fine structure interaction

The spin orbit-interaction is not diagonal in \mathbf{L} or \mathbf{S} due to the term $\mathbf{L} \cdot \mathbf{S}$. However, it is diagonal in $\mathbf{J} = \mathbf{L} + \mathbf{S}$. H_{so} and H_{Dar} are likewise diagonal in \mathbf{J} . Hence, finding the energy level structure due to the fine structure interaction involves evaluating $\langle n, l, S, j, m_j | H_{\text{FS}} | n, l, S, j, m_j \rangle$. Note that $\langle H_{\text{so}} \rangle$ vanishes in an S state, and that $\langle H_{\text{Dar}} \rangle$ vanishes in all states but an S state. It is left as an exercise to show that

$$E_{\text{FS}}(n, j) = (\alpha^2 mc^2) \left(-\frac{\alpha^2}{2n^4} \right) \left(\frac{n}{j+1/2} - \frac{3}{4} \right) \quad (3.16)$$

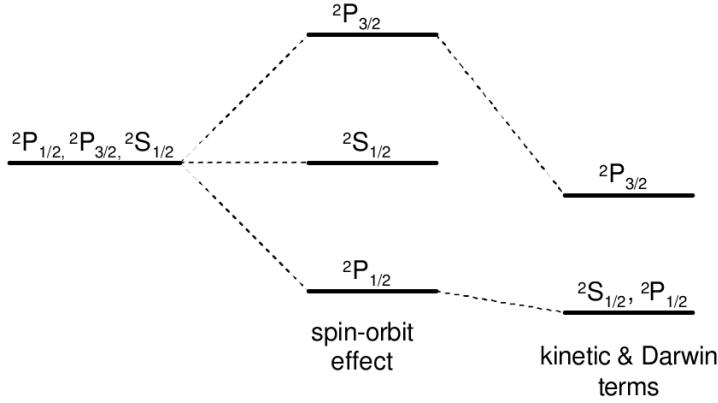


Figure 10. Fine structure of $n = 2$ levels in hydrogen. The degeneracy between the $^2S_{1/2}$ and $^2P_{1/2}$ levels, which looks accidental in non-relativistic quantum mechanics, is really deeply rooted in the relativistic nature of the system. The degeneracy is ultimately broken in QED by the Lamb Shift.

Note that states of a given n and j are degenerate. This degeneracy is a crucial feature of the Dirac theory.

3.2 The Lamb Shift

According to the Dirac theory, states of the hydrogen atom with the same values of n and j are degenerate. Hence, in a given term, $(^2S_{1/2}, ^2P_{1/2})$, $(^2P_{3/2}, ^2D_{3/2})$, $(^2D_{5/2}, ^2F_{5/2})$, etc. form degenerate doublets. However, as described in Chapter 1, this is not exactly the case. Because of vacuum interactions, not taken into account, in the Dirac theory, the degeneracy is broken. The largest effect is in the $n = 2$ state. The energy splitting between the $^2S_{1/2}$ and $^2P_{1/2}$ states is called the Lamb Shift. A simple physical model due to Welton and Weisskopf demonstrate its origin.

Because of zero point fluctuation in the vacuum, empty space is not truly empty. The electromagnetic modes of free space behave like harmonic oscillators, each with zero-point energy $h\nu/2$. The density of modes per unit frequency interval and per volume is given by the well known expression

$$\rho(\nu)d\nu = 8\pi \frac{\nu^2}{c^3} d\nu \quad (3.17)$$

Consequently, the zero-point energy density is

$$W_\nu = \frac{1}{2}h\nu\rho(\nu) = 4\pi \frac{h\nu^3}{c^3} \quad (3.18)$$

With this energy we can associate a spectral density of radiation

$$W_\nu = \frac{1}{8\pi} (\overline{E_\nu^2} + \overline{B_\nu^2}) = \frac{1}{8\pi} E_\nu^2 \quad (3.19)$$

The bar denotes a time average and E_ν and B_ν are the field amplitudes. Hence,

$$\bar{E}_\nu^2 = \frac{32\pi^2 h \nu^3}{c^3} \quad (3.20)$$

For the moment we shall treat the electron as if it were free. Its motion is given by

$$m\ddot{s}_\nu = eE_\nu \cos 2\pi\nu t \quad (3.21)$$

$$\overline{s_\nu^2} = \frac{e^2}{32\pi^4 m^2 \nu^4} \bar{E}_\nu^2 = \frac{e^2 h}{\pi^2 m^2 c^3} \frac{1}{\nu} \quad (3.22)$$

The effect of the fluctuation \mathbf{s}_ν is to cause a change ΔV in the average potential

$$\Delta V = \overline{V(r + s_\nu)} - V(r) \quad (3.23)$$

$V(\mathbf{r} + \mathbf{s}_\nu)$ can be found by a Taylor's expansion:

$$V(\mathbf{r} + \mathbf{s}_\nu) = V(\mathbf{r}) + \Delta V \cdot \mathbf{s}_\nu + \frac{1}{2} \sum_{ij} \frac{\partial^2 V}{\partial s_{\nu,i} \partial s_{\nu,j}} s_{\nu,i} s_{\nu,j} + \dots \quad (3.24)$$

When we average this in time, the second term vanishes because \mathbf{s} averages to zero. For the same reason, in the final term, only contributions with $i = j$ remain. We have, taking the average,

$$\overline{V(\mathbf{r} + \mathbf{s}_\nu)} = V(\mathbf{r}) + \frac{1}{2} \sum_i \frac{\partial^2 V_i}{\partial s_{\nu,i}^2} \overline{s_{\nu,i}^2} \quad (3.25)$$

Since $\overline{s_{\nu,i}^2} = \overline{s_\nu^2}/3$ we obtain finally

$$\overline{V(r + s_\nu)} = \frac{s_\nu^2}{6} \nabla^2 V(r) \quad (3.26)$$

Since $\nabla^2 V(\mathbf{r}) = 4\pi Z e^2 \delta(\mathbf{r})$, we obtain the following expression for the change in energy

$$\delta W_\nu = \frac{2\pi}{3} Z e^2 s_\nu^2 \langle n, \ell, m | \delta(\mathbf{r}) | n, \ell, m \rangle \quad (3.27)$$

The matrix element gives contributions only for S states, where its value is

$$|\Psi_{n,0,0}(0)|^2 = \frac{Z^3}{\pi n^3 a_0^3} \quad (3.28)$$

Combining Eqs. 3.22, 3.28 into Eq. 3.27 yields

$$\delta W_\nu = \frac{2}{3} e^2 s_\nu^2 \frac{Z^4}{n^3 a_0^3} = \frac{2}{3} \frac{e^4 Z^4}{m^2 c^3 \pi^2} \frac{1}{n^3 a_0^3} \frac{h}{\nu} \quad (3.29)$$

Integrating over some yet to be specified frequency limits, we obtain

$$\delta W = \frac{2}{3} \frac{e^4}{m^2 c^3} \frac{Z^4}{\pi^2} \frac{h}{n^3 a_0^3} \ln\left(\frac{\nu_{\max}}{\nu_{\min}}\right) \quad (3.30)$$

At this point, atomic units come in handy. Converting by the usual prescription, we obtain

$$\delta W = \frac{4}{3\pi} \alpha^3 \frac{Z^4}{n^3} \ln\left(\frac{\nu_{\max}}{\nu_{\min}}\right) \text{hartree} = \frac{4}{3\pi} \alpha^5 m c^2 \frac{Z^4}{n^3} \ln\left(\frac{\nu_{\max}}{\nu_{\min}}\right) \quad (3.31)$$

The question remaining is how to choose the cut-off frequencies for the integration. It is reasonable that ν_{\min} is approximately the frequency of an orbiting electron, Z^2/n^3 in atomic units. At lower energies, the electron could not respond. For the upper limit, a plausible guess is the rest energy of the electron, mc^2 . Hence, $\nu_{\max}/\nu_{\min} \sim Z^2/(n^3\alpha^2)$.

For the $2S$ state, this gives

$$\delta W = \frac{1}{6\pi} \alpha^3 \ln \frac{8}{\alpha^2} = 2.46 \times 10^{-7} \text{ atomic units} = 1,600 \text{ MHz} \quad (3.32)$$

The actual value is 1,058 MHz.

Bibliography

- [1] John M. Blatt, Journal of Comp. Phys. **1**, 382 (1967).

Chapter 4

Effects of the Nucleus on Atomic Structure

Until now we have discussed atoms as if the nuclei were point charges with no structure and infinite mass. Real nuclei have finite mass, possibly non-zero angular momentum, \mathbf{I} , and a charge which is spread out over a finite volume. As a result, they possess magnetic dipole moments and electric quadrupole moments coupled to the angular momentum, and possibly higher moments as well. All of these properties affect the atomic energy levels at a level about 10^{-5} rydberg. Here's a catalog of these effects:

Cause	Result	How Observed
Finite Mass	Mass Shift	Isotope shifts
Finite Volume of Charge	Volume Shift	
Magnetic Dipole ($I \geq 1/2$)	Hyperfine structure	Energy splittings
Electric Quadrupole ($I \geq 1$)	Hyperfine structure	

The first two effects produce only a small shift of the spectral line, and thus the only quantity accessible to measurement is the variation of the line position between different isotopes of the same element. (In atoms with only one or two electrons it may be possible to predict the position of a line with significant accuracy to deduce the isotope shift absolutely.) Laser spectroscopy makes it possible to measure isotope shifts to at least 10^{-9} Ry or $10^{-3} - 10^{-4}$ of the shift.

The moments of the nucleus couple to its spin which interacts with the angular momentum of the rest of the atom. This splits the energy levels of the

atom according to the magnitude $|\mathbf{F}|$, where $\mathbf{F} = \mathbf{I} + \mathbf{J}$. The resulting hyperfine structure can be measured with almost limitless precision (certainly $< 10^{-18}$ Rydberg) using the techniques of RF spectroscopy. Hyperfine transitions in Cs and H are currently the best available time and frequency standards. Generally speaking, magnetic dipole interactions predominate in atoms and electric quadrupole interactions in molecules.

With the exception of the mass shift, the manifestations of nuclear structure in atomic spectra provide important information on the static properties of nuclei which are among the most precise information about nuclei. Unfortunately the great precision of the atomic measurements is generally lost in deducing information about nuclear structure because the core electrons affect the magnetic and electric interactions of the valence electrons with the nucleus.

4.1 Hyperfine interaction

4.1.1 Introduction

Hyperfine structure, as its name suggests, is extremely small on the scale of atomic interactions. To give an idea of its size, note that fine structure, which arises from relativistic effects and the spin-orbit interaction, is $O(\alpha^2)$ (i.e. of order α^2 on the scale of atomic interactions). The magnetic hyperfine interaction, which arises from the interaction of the nuclear magnetic moment with the surrounding electrons, is $O(\alpha^2 m_e/M_p)$, which is approximately 1000 times smaller. Nevertheless, studies of hyperfine structure have played an important role in the determination of nuclear properties. Perhaps more relevant today is the role of hyperfine structure in many laser-atomic experiments, particularly those that involve manipulating atoms with light. This is because hyperfine structure affects the optical selection rules and the transfer of momentum and angular momentum to atoms by light. The magnetic hyperfine interaction is most important for atoms with an unpaired electron. Consequently, the alkali-metal atoms, which are the workhorses for laser-atomic physics, all display prominent hyperfine structure, sometimes to the despair of the experimenter.

The fact that the nucleus is a charge cloud with angular momentum suggests the possibility that it might possess magnetic and electric moments. Time reversal and parity invariance restrict the possible magnetic moments to dipole, octopole, ... and the possible electric moments to monopole ($Q = Ze$), quadrupole... The magnetic dipole and electric quadrupole interactions are dominant in the hyperfine interaction.

The magnetic dipole moment can be measured only if the nucleus has $I \geq 1/2$, and it splits only those levels for which $J \geq 1/2$. Similarly, the electric quadrupole interaction is observable only when I and J are both ≥ 1 .

Note that the units used in this section are ESU.

4.1.2 Classical analysis of the magnetic hyperfine interaction

The magnetic moment of the nucleus is generally expressed in terms of the nuclear magneton,

$$\mu_N = \frac{e\hbar}{2M_P c} \quad (4.1)$$

and the nuclear g-factor, g_I

$$\mu = g_I \mu_N I. \quad (4.2)$$

In some cases it is convenient to express the nuclear moment in terms of the Bohr magneton. This is done by defining the g-factor g'_I by

$$\mu = g'_I \mu_B I. \quad (4.3)$$

To emphasize the fact that the nuclei are complex particles we note that the g-factors of the neutron and proton are

$$g_P = +5.586 \quad g_N = -3.826, \quad (4.4)$$

neither one of which is close to a simple integer.

The magnetic moment of the nucleus couples to the magnetic field produced at the nucleus by the electrons in the atom. As a result \mathbf{J} and \mathbf{I} are coupled together to form \mathbf{F} , the total angular momentum of the entire atom

$$\mathbf{F} = \mathbf{J} + \mathbf{I} \quad (4.5)$$

The magnetic coupling between \mathbf{J} and \mathbf{I} adds a term to the Hamiltonian for the magnetic dipole hyperfine structure which is the interaction between a nucleus with magnetic moment $\boldsymbol{\mu}_I$, and the magnetic field \mathbf{B}_J due to a single valence electron.

$$H = -\boldsymbol{\mu}_I \cdot \mathbf{B}_J \quad (4.6)$$

The electron's magnetic field is proportional to its angular momentum \mathbf{J} , and so we can write

$$H = ah \mathbf{I} \cdot \mathbf{J} \quad (4.7)$$

Writing $\mathbf{I} \cdot \mathbf{B}_J$ as $(\mathbf{I} \cdot \mathbf{J}) (\mathbf{J} \cdot \mathbf{B}_J) / J^2$, we have

$$ah = -\frac{\mu_I}{I} \left\langle \frac{\mathbf{J} \cdot \mathbf{B}_J}{J^2} \right\rangle \quad (4.8)$$

The constant a is called the hyperfine coupling constant. By convention, it is written in units of frequency.

There are two contributions to \mathbf{B}_J , orbital and spin: $\mathbf{B}_J = \mathbf{B}_L + \mathbf{B}_S$. We shall first evaluate the fields classically. The magnetic moment of the electron is given by

$$\boldsymbol{\mu}_e = -g_e \mu_B \mathbf{S} \quad (4.9)$$

where μ_B is the Bohr magneton. (The negative sign is taken by convention, so that $g_e \sim 2$ is a positive number.) Then,

$$\mathbf{B}_L(0) = \int \frac{Id\mathbf{s} \times \hat{\mathbf{r}}}{r^2} \rightarrow -e\mathbf{v} \times \left(\frac{-\mathbf{r}}{r^3} \right) = -2\mu_B \frac{\mathbf{L}}{r^3} \quad (4.10)$$

$$\mathbf{B}_S(0) = -\frac{1}{r^3} [\boldsymbol{\mu}_e - 3(\boldsymbol{\mu}_e \cdot \hat{\mathbf{r}})\hat{\mathbf{r}}] = +g_e \frac{\mu_B}{r^3} [\mathbf{S} - 3(\mathbf{S} \cdot \hat{\mathbf{r}})\hat{\mathbf{r}}] \quad (4.11)$$

$$\mathbf{B}_J(0) = -\frac{2\mu_B}{r^3} \left[\mathbf{L} - \frac{g_e}{2} (\mathbf{S} - 3(\mathbf{S} \cdot \hat{\mathbf{r}})\hat{\mathbf{r}}) \right]. \quad (4.12)$$

Note that the expression of $\mathbf{B}_S(0)$ is only correct for $r \neq 0$.

We want to evaluate

$$\langle \mathbf{J} \cdot \mathbf{B}_J \rangle = -\frac{2\mu_B}{r^3} [\ell(\ell+1) - S(S+1) + \langle 3(\mathbf{S} \cdot \hat{\mathbf{r}})\hat{\mathbf{r}} \cdot (\mathbf{L} + \mathbf{S}) \rangle] \quad (4.13)$$

Using $\langle (\mathbf{S} \cdot \hat{\mathbf{r}})^2 \rangle = \langle (\hat{\sigma} \cdot \hat{n}/2)^2 \rangle = 1/4$, $\hat{\mathbf{r}} \cdot \mathbf{L} = 0$, and $S(S+1) = 3/4$, we obtain

$$\langle \mathbf{J} \cdot \mathbf{B}_J \rangle = -2\mu_B \langle \frac{1}{r^3} \rangle \ell(\ell+1) \quad (4.14)$$

So after all this work, the field turns out to depend only on ℓ . From Eq. 4.8, we obtain

$$ah = +\frac{g_e \mu_I \mu_B}{I} \langle \frac{1}{r^3} \rangle \frac{\ell(\ell+1)}{J(J+1)} \quad (4.15)$$

For a hydrogenic atom

$$\langle \frac{1}{r^3} \rangle = \frac{Z^3}{n^3} \frac{1}{\ell(\ell+1)(\ell+1/2)a_0^3} \quad (4.16)$$

Writing $\mu_I = g_I \mu_N I = g_I \mu_B (m_e/M_p) I$, we finally obtain

$$ah = \frac{g_e \mu_B Z^3 g_I \frac{m_e}{M_p}}{n^3 a_0^3} \frac{1}{(\ell+1/2)J(J+1)} \quad (4.17)$$

For hydrogen in the ground state, Eq. 4.17 gives

$$ah = \frac{4g_e}{3} hc R \alpha^2 g_I \frac{m_e}{M_p} \quad (4.18)$$

However, the argument has a flaw: $\langle 1/r^3 \rangle$ diverges for s -states. We must treat these states as a special case.

The orbital magnetic field is absent in s -states. However, the electron has finite probability of being at the origin and it must be regarded as a magnetic “cloud” with magnetization

$$\mathbf{M}(r) = \boldsymbol{\mu}_e |\Psi(\mathbf{r})|^2 \quad (4.19)$$

The magnetization gives rise to a field at the origin

$$\mathbf{B}(0) = \mathbf{H}(0) + 4\pi \mathbf{M}(0) \quad (4.20)$$

The magnetization can be viewed as the sum of a small uniform sphere at the origin, plus a hollow sphere containing the remainder of the magnetization. It is easily shown that the field due to the hollow sphere vanishes. However, the uniform sphere give rise to a finite value of $\mathbf{H}(0)$ due to an equivalent surface

magnetic charge density

$$\sigma_m = \mathbf{M}(0) \cdot \hat{\mathbf{n}} = \mathbf{M}(0) \cos \theta \quad (4.21)$$

which acts as the source of \mathbf{H} .

$$H_z(0) = - \int \frac{\sigma \cos \theta dS}{r^2} = - \frac{4\pi}{3} M(0). \quad (4.22)$$

Since $\mathbf{M}(0) = \mu_e |\psi(0)|^2 = -g_e \mu_B \mathbf{S} |\psi(0)|^2$,

$$\mathbf{B}(0) = - \frac{8\pi}{3} g_e \mu_B \mathbf{S} |\Psi(0)|^2 \quad (4.23)$$

which leads to the same result as Eq. 4.17. This s -state interaction is often called the “contact” (in the sense of touch) term.

We can summarize these results by combining Eqs. 4.12 and Eq. 4.23 taking $g_e = 2$:

$$\mathbf{B}(0) = -2\mu_B \left[\frac{\mathbf{L}}{r^3} - \frac{\mathbf{S}}{r^3} + \frac{3(\mathbf{S} \cdot \hat{r})\hat{r}}{r^3} + \frac{8}{3}\pi\delta(\mathbf{r})\mathbf{S} \right] \quad (4.24)$$

The first three terms in the bracket average to zero in an s -state; the last term contributes only in an s -state.

In the calculations above, we have treated the δ -function part of the dipolar interaction separately. However, the δ -function contribution have the same origin as the “usual” dipolar field, as long as we enforce the vector potential \mathbf{A} to be divergence-free even at $r = 0$. We will give a derivation below.

The vector potential produced by the magnetic moment of the electron reads:

$$\mathbf{A}_S = \frac{\mu_e \times \mathbf{r}}{r^3} \quad (4.25)$$

The corresponding magnetic field can be calculated by taking the curl:

$$\mathbf{B}_S = \nabla \times \left[\frac{\mu_e \times \mathbf{r}}{r^3} \right] = -\nabla \times \left[\mu_e \times \nabla \frac{1}{r} \right] \quad (4.26)$$

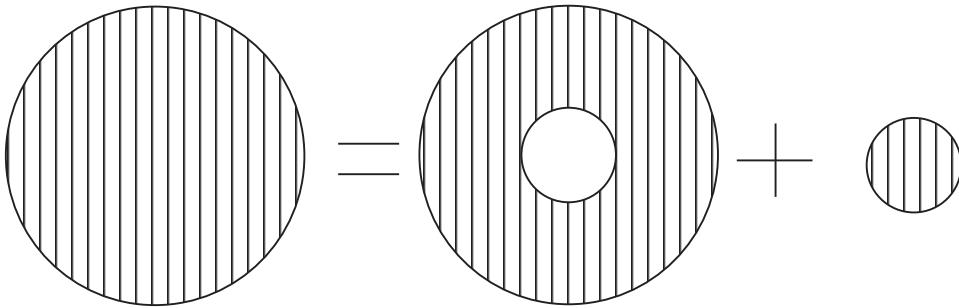


Figure 11. Decomposition of a spherically symmetric cloud of magnetization for finding the field at its center.

By using identities in vector calculus, it reduces to:

$$B_S = -g_e\mu_B \left[\nabla(\mathbf{S} \cdot \nabla \frac{1}{r}) - \mathbf{S} \cdot \nabla^2 \frac{1}{r} \right] \quad (4.27)$$

The above expression can be further simplified by noticing:

$$\partial_i \partial_j \frac{1}{r} = \frac{3\hat{\mathbf{r}}_i \hat{\mathbf{r}}_j - \delta_{ij}}{r^3} - \frac{4\pi}{3} \delta_{ij} \delta^3(\mathbf{r}) \quad (4.28)$$

The intuitive understanding of 4.28 is the following: the left-hand side reduces to the first term at $r \neq 0$. However, the first term is traceless while the left-hand side is not: $\text{Tr}[\partial_i \partial_j \frac{1}{r}] = -4\pi \delta^3(\mathbf{r})$. Therefore, we need the second term $-\frac{4\pi}{3} \delta_{ij} \delta^3(\mathbf{r})$ to carry its trace.

By combining 4.28 with 4.27, we get:

$$B_S = -g_e\mu_B \left[-\frac{\mathbf{S}}{r^3} + \frac{3(\mathbf{S} \cdot \hat{\mathbf{r}})\hat{\mathbf{r}}}{r^3} \right] - \frac{8\pi}{3} g_e \mu_B \mathbf{S} \delta(\mathbf{r}) \quad (4.29)$$

After taking $g_e = 2$, equation 4.29 is exactly the same as our previous result (the spin part in 4.24).

4.1.3 Hyperfine structure at zero magnetic field

The Hamiltonian at zero magnetic field is

$$H = ah \mathbf{I} \cdot \mathbf{J} \quad (4.30)$$

The total angular momentum is $\mathbf{F} = \mathbf{I} + \mathbf{J}$. In zero or low field, F and m_F are good quantum numbers. A “good” quantum number is the eigenvalue of an operator. At zero field, for instance, eigenfunctions of H are eigenfunctions of $\mathbf{F} = \mathbf{I} + \mathbf{J}$.

Physically, \mathbf{I} and \mathbf{J} are tightly coupled by the $ah\mathbf{I} \cdot \mathbf{J}$ interaction: they precess about each other, and about \mathbf{F} . Using $\mathbf{F}^2 = (\mathbf{I} + \mathbf{J})^2$, we obtain

$$\langle \mathbf{I} \cdot \mathbf{J} \rangle = [F(F+1) - J(J+1) - I(I+1)]/2 \quad (4.31)$$

and for the energy levels

$$W(F, m) = \frac{ah}{2} [F(F+1) - J(J+1) - I(I+1)]. \quad (4.32)$$

Note that $m = \langle F_z \rangle$ remains a good quantum number at all fields. (The symbols F and I are used both for operators and eigenvalues: the meaning is clear from the context). F has values $|I - J|, |I - J + 1| \dots |I + J|$. The interval between adjacent terms is

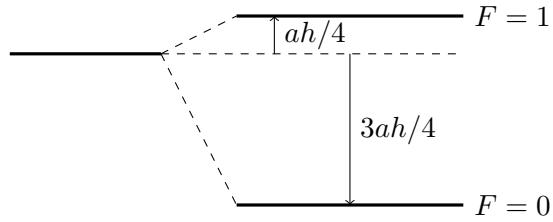
$$W(F) - W(F-1) = haF \quad (4.33)$$

This result, known as the Landé Interval Rule, played an important role in the early development of the theory of hyperfine structure. Furthermore, it is easy to show that

$$\sum_F (2F+1)W(F) = 0 \quad (4.34)$$

so that the “center of gravity” of a hyperfine multiplet is zero.

Example For the ground state of hydrogen, $I = J = S = \frac{1}{2}$. Therefore, we get two states, $F = 1$ which is increased by $\frac{1}{4}ah$ and $F = 0$ which is decreased by $\frac{3}{4}ah$. Note that these obey a sum rule.



4.1.4 Electric quadrupole interaction

If the nucleus does not have a spherically symmetric charge distribution, it probably has a non-zero electric quadrupole moment

$$Q = \frac{1}{e} \int d^3r \rho(\mathbf{r}) [3z^2 - r^2] \quad (4.35)$$

which is < 0 for an oblate charge distribution. In contrast to the nuclear magnetic dipole, which is predominantly determined by the unpaired nucleons, Q is sensitive to collective deformations of the nucleus. Some nuclei are observed with 30% differences between polar and equatorial axes, so Q can be comparable to $\langle r^2 \rangle$, i.e. $\approx 10^{-24}$ cm².

The interaction energy of the quadrupole moment Q with the electron can be found by expanding the term $|\mathbf{r}_e - \mathbf{r}_N|^{-1}$ in spherical harmonics and evaluating the resulting expressions in terms of Clebsch-Gordon coefficients. The resulting energy shifts are then

$$E_{hf}^Q = BC(C + 1) \quad (4.36)$$

where

$$B = \frac{3(Q/a_o^2)}{8I(2I-1)J(J+1)} \langle \frac{1}{r^3} \rangle R_\infty \quad (4.37)$$

and

$$C = [F(F+1) - J(J+1) - I(I+1)] \quad (4.38)$$

[Note that $\mathbf{I} \cdot \mathbf{J}/\hbar^2$, which was involved in H_{mag}^{hf} , is equal to $C/2$].

The preceding expressions, like the corresponding ones for the magnetic interactions, have several significant omissions. The most important are relativistic corrections and core shielding corrections. Calculations of core shielding have been made by Sternheimer [2], and the quadrupole shielding by the core is sometimes prefixed by his name.

The following is a detailed explanation of the form of the electric quadrupole interaction.

4.1.5 Multipole expansion

The electrostatic interaction between a charged nucleus and the charged electrons (and other nuclei in molecules) is

$$\begin{aligned} H_E &= \int d^3r_n \rho_n(\mathbf{r}_n) \varphi(\mathbf{r}_n) & \rho_n - \text{charge density of nucleus} \\ &= \int d^3r_n \int d^3r_e \frac{\rho_n(\mathbf{r}_n)\rho_e(\mathbf{r}_e)}{|\mathbf{r}_n - \mathbf{r}_e|} & \varphi - \text{el. potential at location } \mathbf{r}_n \end{aligned}$$

Consider charges exterior to nucleus, so restrict $|\mathbf{r}_e| > R_N \geq |\mathbf{r}_n|$, where R_N is the radius of the nucleus. With $\mathbf{r} = \mathbf{r}_e - \mathbf{r}_n$ we can expand

$$\begin{aligned} \frac{1}{r} &= \frac{1}{\sqrt{r_e^2 + r_n^2 - 2r_e r_n \cos \theta}} \\ &= \frac{1}{r_e} + \frac{r_n}{r_e^2} P_1 + \frac{r_n^2}{r_e^3} P_2 + \frac{r_n^3}{r_e^4} P_3 + \dots \end{aligned}$$

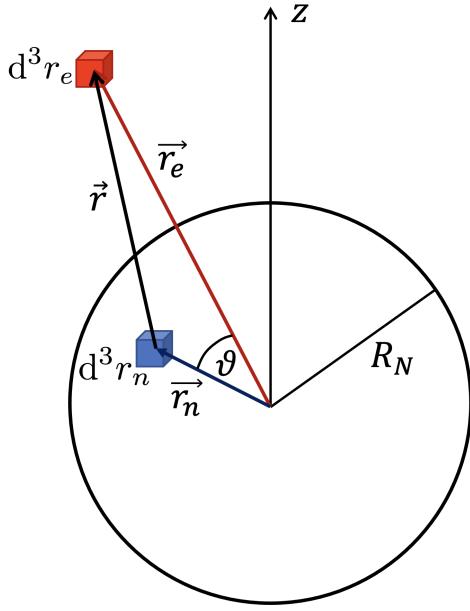


Figure 12. Setup for calculation of electrostatic interaction between electrons and nucleus.

where θ is the angle between \vec{r}_e and \vec{r}_n (see Fig. 12). The P_l are the Legendre polynomials of $\cos \vartheta$. The first few, and the general form, are

$$\begin{aligned} P_0 &= 1 \\ P_1 &= \cos \vartheta \\ P_2 &= \frac{1}{2} (3 \cos^2 \vartheta - 1) \\ P_3 &= \frac{1}{2} (5 \cos^3 \vartheta - 3 \cos \vartheta) \\ P_l &= \frac{1}{2^l l!} \frac{d^l}{(d \cos \vartheta)^l} (\cos^2 \vartheta - 1)^l \end{aligned}$$

This gives us the multipole expansion of H_E :

$$\begin{aligned} H_E &= \sum_l H_{E,l} \\ H_{E,l} &= \int d^3r_n \int d^3r_e \frac{\rho_e^e(\vec{r}_e)\rho_n(\vec{r}_n)}{r_e} \left(\frac{r_n}{r_e}\right)^l P_l(\cos \vartheta) \end{aligned}$$

where ρ_e^e is the charge density external to R_N . $H_{E,l}$ represents the interaction energy from the multipole moment of order 2^l .

It seems difficult to do the integral as ϑ depends on \vec{r}_n and \vec{r}_e . But we can actually write $P_l(\cos \vartheta)$ in terms of products of spherical harmonics involving separately the angles (ϑ_n, φ_n) and (ϑ_e, φ_e) that describe the orientation of \vec{r}_n and \vec{r}_e . This must be so, as $\cos \vartheta$, as well as any function of it, is, for fixed orientation of \vec{r}_n , a function of the angles describing the orientation of \vec{r}_e (so ϑ_e

and φ_e), i.e. $\cos \vartheta$ is a function on the unit sphere described by $\hat{\mathbf{r}}_e$. Similarly, keeping \mathbf{r}_e fixed, $\cos \vartheta$ is a function on the unit sphere described by the unit vector $\hat{\mathbf{r}}_n$. I can decompose any function on the unit sphere using spherical harmonics. One can show (see e.g. Jackson, Electrodynamics)

$$P_l(\cos \vartheta) = \frac{4\pi}{2l+1} \sum_{m=-l}^l (-1)^m Y_{l,-m}(\vartheta_n, \varphi_n) Y_{lm}(\vartheta_e, \varphi_e)$$

We thus get the decomposition

$$\begin{aligned} H_{E,l} &= \mathcal{Q}^{(l)} \cdot \mathcal{F}^{(l)} \\ &= \sum_{m=-l}^l (-1)^m Q_m^{(l)} F_{-m}^{(l)} \\ Q_m^{(l)} &= \sqrt{\frac{4\pi}{2l+1}} \int d^3 r_n \rho_n(\mathbf{r}_n) r_n^l Y_{lm}(\vartheta_n, \varphi_n) \\ F_m^{(l)} &= \sqrt{\frac{4\pi}{2l+1}} \int d^3 r_e \rho_e(\mathbf{r}_e) r_e^{-l-1} Y_{lm}(\vartheta_e, \varphi_e) \end{aligned}$$

$l = 0$: Monopole interaction We have

$$H_{E,0} = Ze \phi^e$$

with

$$Ze = Q_0^{(0)} = \int d^3 r_n \rho_n(\mathbf{r}_n)$$

the total nuclear charge and

$$\phi^e = F_0^{(0)} = \int d^3 r_e \frac{\rho_e^e(\mathbf{r}_e)}{r_e}$$

the electrostatic potential from external charges ($r_e > R_N$) at the nucleus.

$l = 1$ Dipolar interaction (will turn out to be zero)

$$H_{E,1} = -p_z E_z^e - \frac{1}{2} p_+ E_-^e - \frac{1}{2} p_- E_+^e$$

with

$$p_z = Q_0^{(1)} = \int d^3 r_n \rho_n(\mathbf{r}_n) \mathbf{r}_n \cos \vartheta_n = \int d^3 r \rho_n(\mathbf{r}_n) z_n$$

which is the z -component of the nuclear electrical dipole moment, and

$$p_{\pm 1} \equiv p_x \pm i p_y = \pm \sqrt{2} Q_{\pm 1}^{(1)} = \int d^3 r_n \rho_n(\mathbf{r}_n) (x_n \pm i y_n)$$

E_z^e is the electric field at the nucleus due to external charges, given by

$$E_z^e = -F_0^{(1)} = - \int d^3 r_e \frac{\rho_e^e(\mathbf{r}_e)}{r_e^2} \cos \vartheta_e = - \int d^3 r_e \frac{\rho_e^e(\mathbf{r}_e)}{r_e^3} z_e = - \frac{\partial \varphi_e^e}{\partial z}$$

$$E_{\pm}^e = E_x^e \pm iE_y^e = \pm F_{\pm}^{(1)} = - \int d^3r_e \frac{\rho_e^e}{r_e^3} (x_e \pm iy_e) = - \frac{\varphi_e^e}{\partial x} \mp i \frac{\partial \varphi_e^e}{\partial y}$$

$l = 2$: Electric quadrupole interaction

$$H_{E,2} = \sum_{m=-2}^2 (-1)^m Q_m (\nabla E^e)_{-m}$$

Q_m is the nuclear electric quadrupole tensor, $(\nabla E^e)_m$ the electric field gradient tensor. The various components of Q_m are

$$\begin{aligned} Q_0 &= Q_0^{(2)} = \frac{1}{2} \int d^3r_n \rho_n(\mathbf{r}_n) r_n^2 (3 \cos^2 \vartheta_n - 1) = \frac{1}{2} \int d^3r_n \rho_n (3z_n^2 - r_n^2) \\ Q_{\pm 1} &= Q_{\pm 1}^{(2)} = \mp \sqrt{\frac{3}{2}} \int d^3r_n \rho_n z_n (x_n \pm iy_n) \\ Q_{\pm 2} &= Q_{\pm 2}^{(2)} = \sqrt{\frac{3}{8}} \int d^3r_n \rho_n (x_n \pm iy_n)^2 \end{aligned}$$

For the electric field gradient tensor, we have

$$\begin{aligned} (\nabla E^e)_0 &= F_0^{(2)} = \frac{1}{2} \int d^3r_e \frac{\rho_e}{r_e^3} (3 \cos^2 \vartheta - 1) = -\frac{1}{2} \frac{\partial E_z^e}{\partial z} \\ (\nabla E^e)_{\pm 1} &= \pm \frac{1}{\sqrt{6}} \frac{\partial E_{\pm}^e}{\partial z} \\ (\nabla E^e)_{\pm 2} &= -\frac{\sqrt{6}}{12} \left(\frac{\partial}{\partial x} \pm i \frac{\partial}{\partial y} \right) E_{\pm}^e \end{aligned}$$

Using the above representation of the nuclear electric quadrupole tensor and the electric field gradient tensor is efficient, given by five numbers, without redundancies. Alternatively, we could also use cartesian components Q_{ij} and $(\nabla E^e)_{ij}$ with i and $j = x, y, z$. These are 3×3 matrices. When demanding them to be traceless and symmetric, they each again are described by five numbers. One has

$$\begin{aligned} H_{E,2} &= -\frac{1}{6} \sum_{i=x,y,z} \sum_{j=x,y,z} Q_{ij} (\nabla E^e)_{ij} \\ Q_{ij} &= \int d^3r_n \rho_n (3x_{ni}x_{nj} - \delta_{ij}r_n^2) \\ (\nabla E^e)_{ij} &= - \int d^3r_e \frac{\rho_e^e}{r_e^5} (3x_{ei}x_{ej} - \delta_{ij}r_e^2) = -\frac{\partial^2 \varphi_e^e}{\partial x_i \partial x_j} \end{aligned}$$

4.1.6 Theoretical restrictions on multipole orders

The nucleus, having definite spin I , has its orientation fully specified by the orientation of the spin angular momentum \mathbf{I} .

Parity consideration If

- all nuclear electrical effects arise from electrical charges

- there is no degeneracy of nuclear states with different parity
- the nuclear Hamiltonian is unaltered by an inversion of coordinates ($\mathbf{r} \rightarrow -\mathbf{r}$)

then no odd (l odd) electrical multipole can exist. So there is no electric dipole or octupole moment.

Sketch of proof: The wave function of the nucleus must obey

$$\begin{aligned}\psi(\mathbf{r}_1, \dots, \mathbf{r}_A) &= \pm \psi(-\mathbf{r}_1, \dots, -\mathbf{r}_A) \\ \Rightarrow |\psi(\mathbf{r}_1, \dots, \mathbf{r}_A)|^2 &= |\psi(-\mathbf{r}_1, \dots, -\mathbf{r}_A)|^2\end{aligned}$$

Now for even l , the Y_{lm} is unchanged by inversion, but for odd l , Y_{lm} reverses sign.

$$\Rightarrow \int d^3 r_n \rho_n(\mathbf{r}_n) Y_{lm}(\vartheta_n, \varphi_n) = 0 \quad \text{for odd } l.$$

See Purcell and Ramsey, Phys. Rev. 78, 699 (1950). Search for neutron EDM. Outcome:

$$d_{\text{neutron}} < e \cdot 5 \cdot 10^{-20} \text{ cm}$$

Constraint on observable multiple order For nuclear spin I it is impossible to observe a nuclear multipole moment of order 2^l for $l > 2I$.

Proof: If $\rho_n = \psi_n^* \psi_n$ we have

$$Q_m^{(l)} = \int d^3 r_n \psi_n^* r_n^l Y_{lm} \psi_n.$$

Now Y_{lm} is an orbital wavefunction for orbital angular momentum l . ψ_n is the wavefunction of angular momentum I . So $Y_{lm} \psi_n$ is the wavefunction of a system with angular momentum between $|l-I|$ and $l+I$. ψ_n^* and $\psi_n^* r_n^e$ represent wavefunctions of angular momentum I .

For the integral to be non-zero, we therefore have to have that I lies between $|l-I|$ and $l+I$. Therefore $l \leq 2I$. In other words, I , I , and l must satisfy the triangle rule.

An analogous result holds for the electric field gradient tensor and the total angular momentum J of the electronic wavefunction: Atoms (or molecules) in an angular momentum state J , the field tensor $F_m^{(l)}$ is zero unless $l \leq 2J$.

So for example, even a nucleus with a large I and a non-zero nuclear quadrupole moment cannot have an electric quadrupole interaction energy with an atom whose $J = \frac{1}{2}$.

More intuitively, a $J = \frac{1}{2}$ angular momentum state is a “sensor” allowing us to probe the nuclear charge distribution only using two alignments (say aligned and anti-aligned with the nuclear spin), giving us at most two different energy levels. That would be enough to see a nuclear electric dipole (but that is zero), and to sense the magnetic field created by the magnetic moment of the nucleus.

However, it is not enough information to distinguish a sphere from a cigar or a pancake-shaped charge distribution (they all would give the same energy for

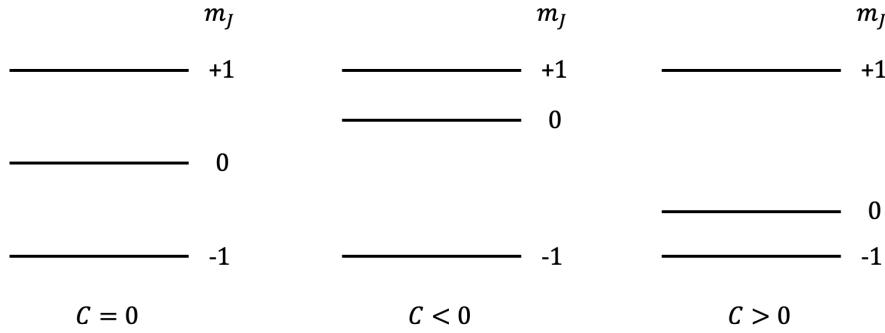


Figure 13. Three outcomes when probing a nucleus with $I \geq 1$ using a $J = 1$ electronic state. (Left) Without quadrupole moment all one can see is a splitting from an effective magnetic field (e.g. the actual magnetic field due to the magnetic moment of the nucleus). With quadrupole moment, one sees the central $m_J = 0$ level shifted up ($C < 0$, middle) or down ($C > 0$, right) with respect to the average of the two $m_J = \pm 1$ levels.

the two orientations of \mathbf{J}). The two orientations, i.e. the result for the two m_J values, which correspond semi-classically to two values of $\cos \vartheta$, are not enough to tell me about the nuclear shape along directions transverse to \mathbf{I} . But take instead $J = 1$. Now there are three alignments of our “sensor”, e.g. aligned, anti-aligned, and transverse to \mathbf{I} . Correspondingly, there are three energy levels (see Fig. 13). Generally, all possible outcomes can be modelled by an effective Hamiltonian $m_J g \mu_B B + C m_J^2$. The magnetic field will cause symmetric splitting of $m_J = \pm 1$ about $m_J = 0$, while the second term will shift $m_J = \pm 1$ together up ($C > 0$) or down ($C < 0$) with respect to $m_J = 0$. But m_J^2 comes from a term like $(\mathbf{I} \cdot \mathbf{J})^2 \propto \cos^2 \vartheta$, which can be expressed as a quadrupole interaction plus a constant.

4.1.7 Simple examples for a non-zero quadrupole moment

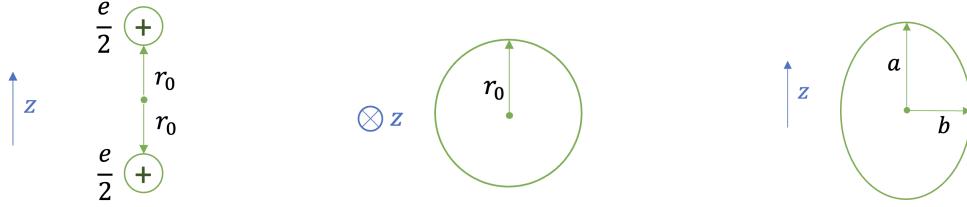


Figure 14. Left: two positive charges; center: a ring of uniform charge density, with total charge e ; right: an ellipsoid of uniform charge density, with total charge e .

Two positive charges (Fig. 14, left):

$$Q \equiv \frac{1}{e} \int d^3r_n \rho_n (3z^2 - r^2)$$

gives

$$Q = 2r_0^2 > 0.$$

A ring of uniform charge density, with total charge e (Fig. 14, center):

$$Q = -r_0^2 < 0.$$

An ellipsoid of uniform charge density, with total charge e (Fig. 14, right):

$$\frac{x^2 + y^2}{b^2} + \frac{z^2}{a^2} = 1$$

gives

$$\begin{aligned} Q &= \rho_n \int dz \int d^2r_\perp (2z^2 - r_\perp^2) \\ &= \rho_n b^2 a \int d\tilde{z} d^2\tilde{r}_\perp (2a^2 \tilde{z}^2 - b^2 \tilde{r}_\perp^2) \\ &= \rho_n b^2 a \left(2a^2 \int d^3\tilde{r} \tilde{z}^2 - b^2 \int d^3\tilde{r} \tilde{r}_\perp^2 \right) \end{aligned}$$

With

$$\begin{aligned} \int d^3\tilde{r} \tilde{z}^2 &= \frac{1}{3} \int d^3\tilde{r} \tilde{r}^2 = \frac{4\pi}{3} \int d\tilde{r} \tilde{r}^4 = \frac{4\pi}{15}, \\ \int d^3\tilde{r} \tilde{r}_\perp^2 &= \frac{2}{3} \int d^3\tilde{r} \tilde{r}^2 = \frac{8\pi}{15}, \end{aligned}$$

We have

$$Q = \frac{1}{e} \frac{8\pi}{15} \rho_n b^2 a (a^2 - b^2).$$

Substituting $e = \frac{4\pi}{3} b^2 a \rho_n$, we get

$$Q = \frac{2}{5} (a^2 - b^2) > 0.$$

Note that if $a < b$, then $Q < 0$.

4.1.8 Calculating quadrupole moment tensor

To calculate the quadrupole moment tensor, we need to rotate the tensor in the nuclear frame into the lab frame, where the orientation of the nucleus is given by I and m_I . Now conveniently, the components of the coordinate vector \mathbf{r}_n transform just like the components of I , and products like $(x_n \pm iy_n) \cdot z_n$ transform like the properly symmetrized $\frac{1}{2}(I_{\pm}I_z + I_zI_{\pm})$.

So we have

$$\begin{aligned} Q_0 &= \frac{1}{2} \int d^3 r_n \rho_n (3z_n^2 - r_n^2) = C \cdot \frac{1}{2} (3I_z^2 - \mathbf{I}^2), \\ Q_{\pm 1} &= \mp \sqrt{\frac{3}{2}} C \cdot \frac{1}{2} (I_{\pm}I_z + I_zI_{\pm}), \\ Q_{\pm 2} &= \sqrt{\frac{3}{8}} C I_{\pm}^2. \end{aligned}$$

Definition of the quadrupole moment

$$\begin{aligned} Q &\equiv \frac{1}{e} \int d^3 r_n \rho_{nII} (3z_n^2 - r_n^2) \\ &= \frac{1}{e} \int d^3 r_n \rho_{nII} r_n^2 (3 \cos^2 \vartheta - 1) \\ &= \langle r_n^2 (3 \cos^2 \vartheta - 1) \rangle_{II} \end{aligned}$$

where ρ_{nII} is the charge density when nucleus is in the orientation state with $m_I = I$, and $\langle \dots \rangle_{II}$ is an average in $|II\rangle$.

$$\begin{aligned} Q &= \frac{2}{e} \langle II | Q_0 | II \rangle \\ &= \frac{C}{e} (3I^2 - I(I+1)) \\ &= \frac{C}{e} I(2I-1) \end{aligned}$$

which gives

$$\begin{aligned} Q_0 &= \frac{eQ}{2I(2I-1)} (3I_z^2 - I(I+1)), \\ Q_{\pm 1} &= \mp \frac{\sqrt{6}}{2} \frac{eQ}{2I(2I-1)} (I_{\pm}I_z + I_zI_{\pm}), \\ Q_{\pm 2} &= \frac{\sqrt{6}}{2} \frac{eQ}{2I(2I-1)} I_{\pm}^2. \end{aligned}$$

Now analogously, the orientation of the atom or molecule is specified by the

orientation of its angular momentum \mathbf{J} . One has

$$\begin{aligned} (\nabla E^e)_0 &= \frac{e q_J}{2J(2J-1)} (3J_z^2 - J(J+1)), \\ (\nabla E^e)_{\pm 1} &= \mp \frac{\sqrt{6}}{2} \frac{e q_J}{2J(2J-1)} (J_z J_{\pm} + J_{\pm} J_z), \\ (\nabla E^e)_{\pm 2} &= \frac{\sqrt{6}}{2} \frac{e q_J}{2J(2J-1)} J_{\pm}^2, \end{aligned}$$

with

$$q_J \equiv \frac{1}{e} \int d^3 r_e \rho_{eJJ}^e \frac{3 \cos^2 \vartheta - 1}{r_e^3} = \frac{1}{e} \left\langle \frac{\partial^2 \phi_e^e}{\partial z^2} \right\rangle_{JJ} = - \left\langle \frac{3 \cos^2 \vartheta - 1}{r_e^3} \right\rangle_{JJ},$$

where $\langle \dots \rangle_{JJ}$ is the average when atom or molecule is in state with $m_J = J$.

Alternatively, we can express Q and ∇E in cartesian coordinates:

$$\begin{aligned} Q_{ij} &= \frac{e Q}{I(2I-1)} \left(3 \frac{I_i I_j + I_j I_i}{2} - \delta_{ij} I(I+1) \right) \\ (\nabla E)_{ij} &= - \frac{e q_J}{J(2J-1)} \left(3 \frac{J_i J_j + J_j J_i}{2} - \delta_{ij} J(J+1) \right). \end{aligned}$$

Now we get

$$\begin{aligned} H_{E2} &= -\frac{1}{6} \sum_{i=x,y,z} \sum_{j=x,y,z} Q_{ij} (\nabla E^e)_{ij} \\ &= \frac{e^2 q_J Q}{2I(2I-1)J(2J-1)} \left(3(\mathbf{I} \cdot \mathbf{J})^2 + \frac{3}{2} \mathbf{I} \cdot \mathbf{J} - I(I+1)J(J+1) \right). \end{aligned}$$

See Appendix C in N. Ramsey, Molecular Beams.

4.1.9 Diagonalization of the electric quadrupole interaction

We need to diagonalize $3(\mathbf{I} \cdot \mathbf{J})^2 + \frac{3}{2} \mathbf{I} \cdot \mathbf{J} - I(I+1)J(J+1)$. Using $\mathbf{F} = \mathbf{J} + \mathbf{I}$,

$$\begin{aligned} \mathbf{J} \cdot \mathbf{I} &= \frac{1}{2} (F(F+1) - I(I+1) - J(J+1)) \equiv \frac{1}{2} C, \\ \Rightarrow 3(\mathbf{I} \cdot \mathbf{J})^2 + \frac{3}{2} \mathbf{I} \cdot \mathbf{J} - \mathbf{I}^2 \mathbf{J}^2 &= \frac{3}{4} C(C+1) - I(I+1)J(J+1), \\ \Rightarrow H_{E2} &= \frac{e^2 q_J Q}{2I(2I-1)J(2J-1)} \left(\frac{3}{4} C(C+1) - I(I+1)J(J+1) \right). \end{aligned}$$

Note that for $F = I + J$, $H_{E2} = \frac{e^2 q_J Q}{4}$.

The last term with $\mathbf{I}^2 \mathbf{J}^2$ is a constant and often omitted or included in the reference energy.

$$\Rightarrow H'_{E2} = h b 2 \mathbf{I} \cdot \mathbf{J} (2 \mathbf{I} \cdot \mathbf{J} + 1),$$

which in the $|F, m_F\rangle$ basis is just

$$\langle F, m_F | H_{E2} | F, m_F \rangle = h b C(C+1).$$

Together with the magnetic dipole moment interaction $a h \mathbf{I} \cdot \mathbf{J}$, this gives

$$\langle F, m_F | H_{HF} | F, m_F \rangle = h \frac{a}{2} C + h b C(C+1),$$

with $C = F(F+1) - I(I+1) - J(J+1)$.

Example: Nuclear quadrupole interaction in ^{23}Na Consider a ^{23}Na atom, which has $I = \frac{3}{2}$, $S = \frac{1}{2}$.

Its ground states $F = 1$ and $F = 2$ have a hyperfine splitting of $1.77\text{GHz} = 2a_{\text{hf}} \Rightarrow a_{\text{hf}} = 885\text{MHz}$.

Its excited states are:

$$3^2P_{1/2} \quad J = \frac{1}{2} \Rightarrow F = 1, 2$$

Magnetic dipole interaction: $a_{1/2} = 94.4\text{MHz}$

Electric quadrupole interaction: None! ($J = \frac{1}{2}$)

$$3^2P_{3/2} \quad J = \frac{3}{2} \Rightarrow F = 0, 1, 2, 3$$

Magnetic dipole interaction: $a_{3/2} = \frac{1}{5}a_{1/2}$

Electric quadrupole interaction: $b = 2.72\text{MHz}$.

Excited states feature 1: $a_{3/2} = \frac{1}{5}a_{1/2}$. Why is this so? Let's calculate their respective magnetic hyperfine interactions:

$$\begin{aligned} h a &= \frac{\mu_I}{I} \frac{\mathbf{B}_z \cdot \mathbf{J}}{|\mathbf{J}|^2} \\ &= \frac{\mu_I}{I} \frac{\mathbf{J}}{J(J+1)} \cdot 2\mu_0 \left(\frac{\mathbf{L}}{r^3} - \frac{\mathbf{S}}{r^3} + \frac{3(\mathbf{S} \cdot \hat{r}) \hat{r}}{r^3} \right) \\ &= \frac{\mu_I}{I} \frac{2\mu_0}{J(J+1)} \underbrace{\left\langle \frac{1}{r^3} \right\rangle \left((\mathbf{L} - \mathbf{S}) \cdot (\mathbf{L} + \mathbf{S}) + 3(\mathbf{S} \cdot \hat{r})^2 \right)}_{\mathbf{L}^2 - \mathbf{S}^2 + 3 \cdot \frac{1}{4}} \\ &= 2\mu_0 \frac{\mu_I}{I} \left\langle \frac{1}{r^3} \right\rangle \frac{L(L+1)}{J(J+1)} \quad \Rightarrow \text{Note } \mathbf{B}_{\mathbf{J}} \cdot \mathbf{J} \text{ only depends on } L! \\ &= 2\mu_0 \frac{\mu_I}{I} \frac{Z^3}{n^3 a_0^3} \frac{1}{(L + \frac{1}{2})J(J+1)} \\ &\Rightarrow \frac{a_{1/2}}{a_{3/2}} = \frac{\frac{3}{2} \left(\frac{3}{2} + 1 \right)}{\frac{1}{2} \left(\frac{1}{2} + 1 \right)} = 5. \end{aligned}$$

Excited states feature 2: no electric quadrupole interaction in $^2S_{1/2}$ and $^2P_{1/2}$. Both of them have spherical charge distributions (see Fig. 15) so $q_{J=1/2} = 0$. In

particular, consider the angular part of ${}^2P_{1/2}$:

$$\begin{aligned} |J = \frac{1}{2}, m_J = \frac{1}{2}\rangle &= \sqrt{\frac{2}{3}}|L = 1, m_L = 1, S = \frac{1}{2}, m_S = -\frac{1}{2}\rangle \\ &\quad + \sqrt{\frac{1}{3}}|L = 1, m_L = 0, S = \frac{1}{2}, m_S = +\frac{1}{2}\rangle \\ \Rightarrow \langle \vartheta, \phi | J = \frac{1}{2}, m_J = \frac{1}{2} \rangle &= \sqrt{\frac{2}{3}}Y_{11}(\vartheta, \phi)|S = \frac{1}{2}, m_S = -\frac{1}{2}\rangle \\ &\quad + \sqrt{\frac{1}{3}}Y_{10}(\vartheta, \phi)|S = \frac{1}{2}, m_S = +\frac{1}{2}\rangle. \end{aligned}$$

And

$$\begin{aligned} |\langle \vartheta, \phi | J = \frac{1}{2}, m_J = \frac{1}{2} \rangle|^2 &= \frac{2}{3}|Y_{11}(\vartheta, \phi)|^2 + \frac{1}{3}|Y_{10}(\vartheta, \phi)|^2 \\ &= \frac{2}{3} \frac{3}{8\pi} \sin^2 \vartheta + \frac{1}{3} \frac{3}{4\pi} \cos^2 \vartheta \\ &= 1 = \text{constant}. \end{aligned}$$

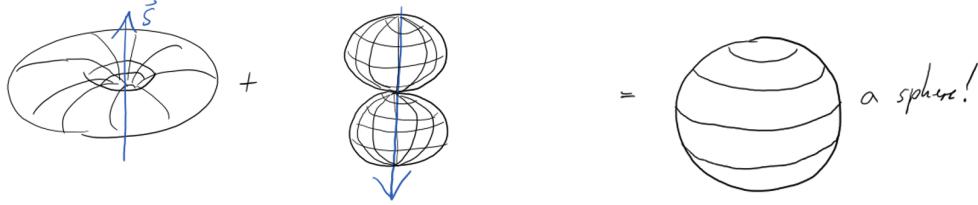


Figure 15. Angular contributions sum to zero in ${}^2P_{1/2}$.

Of course, we shouldn't be surprised that a $J = \frac{1}{2}$ state doesn't have an electric field gradient at the nucleus. An electric field gradient is a tensor of rank 2, and I cannot combine $\frac{1}{2}$ and $\frac{1}{2}$ to make total angular momentum 2. Or again more colloquially, $J = \frac{1}{2}$ gives only two orientations, up and down, and these two will not allow me to distinguish a gradient in the electric field. An observed energetic splitting could always be interpreted as an effective magnetic dipole interaction. See Fig. 16 for an illustration.

With $J = \frac{1}{2}$, I could in principle detect an electric field, i.e., a gradient of the potential. However, this does not occur for electronic wavefunctions of definite parity, as then

$$|\Psi_e(\hat{r})|^2 = |\Psi_e(-\hat{r})|^2 \text{ and } \int d^3r_e \frac{\rho_e(\hat{r})}{r_e^2} \cos \vartheta_e = 0$$

Excited states feature 3: non-zero quadrupole interaction in ${}^2P_{3/2}$. For a ${}^2P_{3/2}$ ${}^{23}\text{Na}$ atom, $J = \frac{3}{2}$, which can give an electric field gradient; and $I = \frac{3}{2}$, which

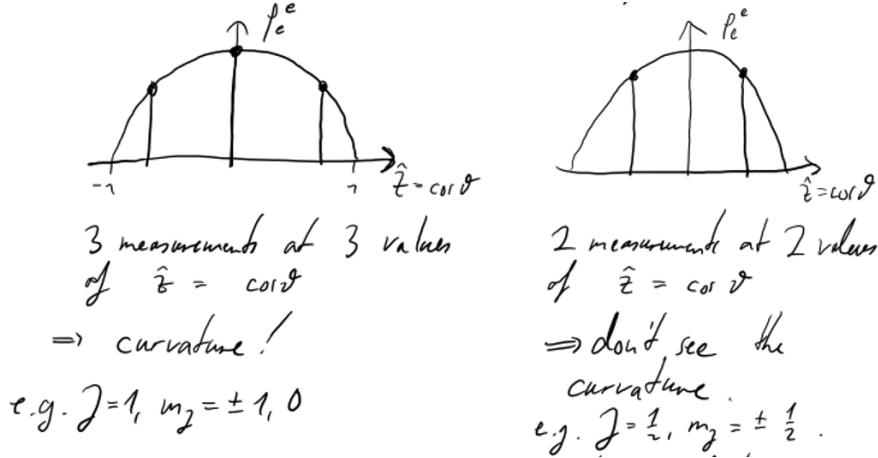


Figure 16. To see a gradient in E , i.e., a curvature in ϕ_e^e , I need to measure in at least 3 places.

can possess an electronic quadrupole moment.

$$\begin{aligned}
q_J &= - \left\langle \frac{3 \cos^2 \vartheta_e - 1}{r_e^3} \right\rangle_{JJ} \quad \text{due to single valence electron} \\
&= - \left\langle \frac{1}{r_e^3} \right\rangle \left\langle 3 \cos^2 \vartheta_e - 1 \right\rangle_{JJ} \\
\left\langle 3 \cos^2 \vartheta_e - 1 \right\rangle_{JJ} &= 0 \quad \text{for } J = \frac{1}{2} \quad (\text{as it should}) \\
&= \int d\Omega |Y_{11}|^2 \cdot (3 \cos^2 \vartheta - 1) \\
&= \frac{3}{8\pi} \cdot 2\pi \int d\vartheta \sin \vartheta \sin^2 \vartheta (3 \cos^2 \vartheta - 1) \\
&= \frac{3}{4} \int_{-1}^1 d(\cos \vartheta) (1 - \cos^2 \vartheta) (3 \cos^2 \vartheta - 1) \\
&= \frac{3}{4} \int_{-1}^1 dz (-3z^4 + 4z^2 - 1) \\
&= -\frac{2}{5} \\
\left\langle \frac{1}{r_e^3} \right\rangle &= \frac{Z^3}{n^3 a_0^3} \frac{1}{l(l+1)(l+\frac{1}{2})} \quad \text{with } Z \text{ and } n \text{ effective values.}
\end{aligned}$$

Note that fine structure splitting is $\delta = \mu_0^2 \langle \frac{1}{r^3} \rangle (2l+1)Z$, so $\langle \frac{1}{r_e^3} \rangle = \frac{\delta}{\mu_B^2 (2l+1)Z} \Rightarrow q_J = \frac{2}{15} \frac{\delta}{\mu_B^2 Z}$.

A better value of $\langle \frac{1}{r^3} \rangle$ comes directly from the measured hyperfine structure:

$$\begin{aligned}
\langle \frac{1}{r^3} \rangle &= \frac{\hbar a}{\mu_B (\frac{\mu_I}{I})} \frac{J(J+1)}{2L(L+1)} \quad \text{with } \mu_I = g_I \mu_N I \\
\Rightarrow q_{J=\frac{3}{2}} &= \frac{2}{5} \frac{\hbar a_{J=3/2} \frac{3}{2} \cdot \frac{5}{2}}{\mu_B g_I \mu_N 2 \cdot 2} \\
&= \frac{3}{8} \frac{\hbar a_{J=3/2}}{\mu_B g_I \mu_N} = \frac{3}{40} \frac{\hbar a_{J=1/2}}{\mu_B g_I \mu_N} \\
h b &= \frac{3}{8} \frac{e^2 q_J Q}{I(2I-1)J(2J-1)} \\
&= \frac{1}{24} e^2 q_J Q \\
&= \frac{1}{320} h a_{J=1/2} \frac{e^2 Q}{g_I \mu_B^2 \left(\frac{m}{m_P}\right)} \quad \text{with } \mu_B = \frac{e\hbar}{2mc} \Rightarrow \frac{\mu_B}{e} = \frac{\hbar}{2mc} = \lambda_c^{\frac{1}{2}} \\
&= \frac{4}{320 h a_{J=1/2}} \frac{Q}{\lambda_c^2} \frac{m_P}{m g_I} \\
\Rightarrow \frac{Q}{\lambda_c^2} &= 80 \frac{b}{a_{J=1/2}} g_I \frac{m}{m_P}
\end{aligned}$$

where $g_{\pm} \frac{m}{m_P} = 0.0008$, $24 b = 2.72 \text{ MHz}$, $a_{J=1/2} = 94.4 \text{ MHz}$

$$\Rightarrow \frac{Q}{\lambda_c^2} = \frac{10}{3} \frac{e q Q}{h a_{J=1/2}} g_I \frac{m}{m_P} = 7.7 \cdot 10^{-5}$$

$$Q = 0.11 \cdot 10^{-24} \text{ cm}^2 = 0.11 \text{ barn} > 0.$$

This was first measured by Perl, Rabi, Senitzky, Phys. Rev. 97, 838 (1954).

Structure of ^{23}Na $^2P_{3/2}$ We now use our previous calculations to derive the energy structure of ^{23}Na $^2P_{3/2}$.

The contribution from electric quadropole interaction can be calculated from:

$$24 b = e^2 q_J Q / \hbar = 2.72 \text{ MHz}$$

$$J = \frac{3}{2}, I = \frac{3}{2}; a_{3/2} = 18.5 \text{ MHz} (\approx \frac{1}{5} a_{1/2})$$

For each hyperfine level, we have a corresponding $C = F(F+1) - I(I+1) - J(J+1)$:

$$\begin{aligned}
F = 0 \Rightarrow C &= -\frac{3}{2} \cdot \frac{5}{2} \cdot 2 = -\frac{15}{2} \\
F = 1 \Rightarrow C &= 2 - \frac{15}{2} = -\frac{11}{2} \\
F = 2 \Rightarrow C &= 6 - \frac{15}{2} = -\frac{3}{2} \\
F = 3 \Rightarrow C &= 12 - \frac{15}{2} = \frac{9}{2}
\end{aligned}$$

Notice how the center of gravity stays put, i.e., $\sum_{F=0,1,2,3} (2F + 1) \cdot C_F = 0$.

Using $H_{HF} = \frac{a\hbar}{2} \cdot C + b h(C(C+1) - \frac{4}{3}I(I+1)J(J+1))$, we can arrive at the energy structure of ${}^2P_{3/2}$ in ${}^{23}\text{Na}$ (Fig. 17).

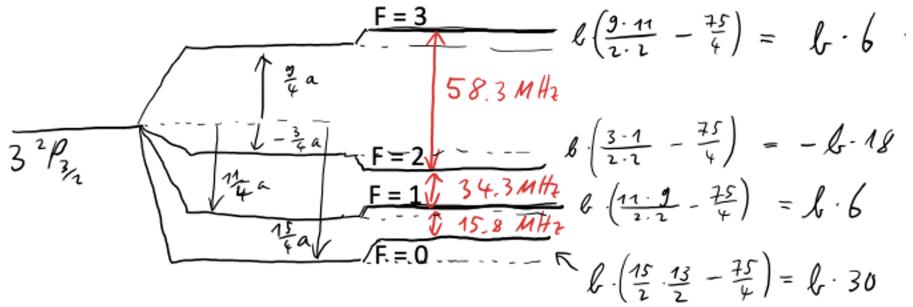


Figure 17. Hyperfine splitting in ${}^{23}\text{Na}$ ${}^2P_{3/2}$ due to magnetic dipole interaction (left) and electric quadropole interaction (right).

Figure 18 sketches the hyperfine splittings in ${}^2P_{3/2}$, ${}^2P_{1/2}$ and ${}^2S_{1/2}$ for comparison.

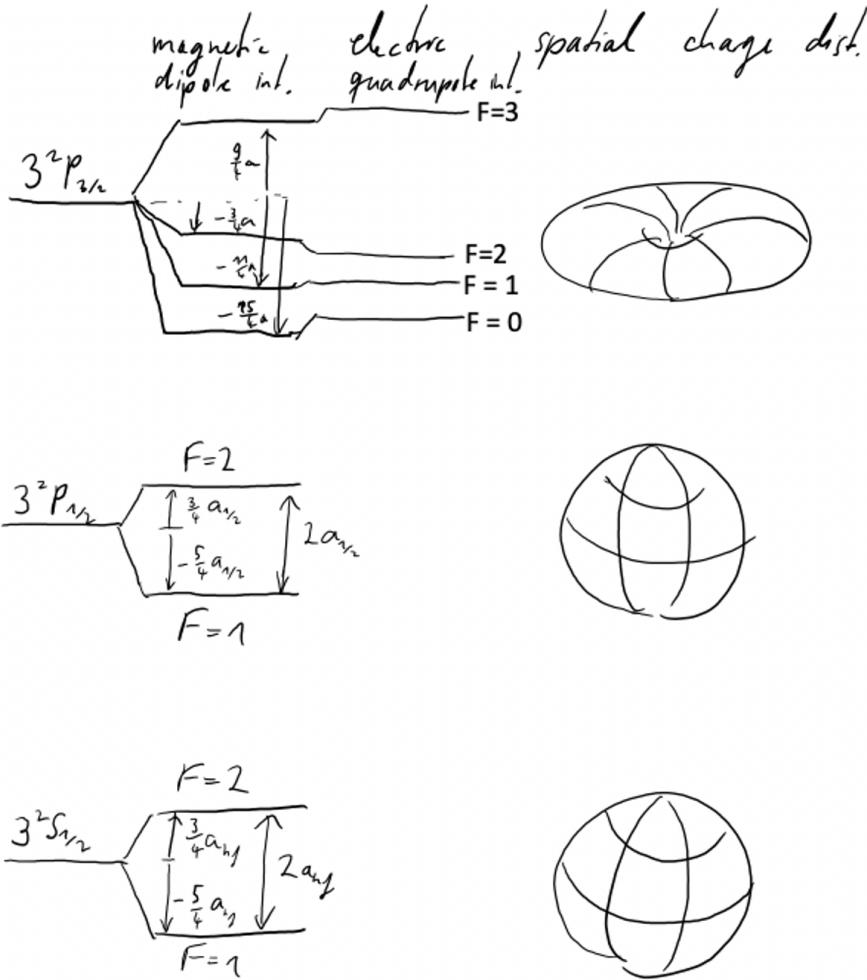


Figure 18. Hyperfine splitting in ^{23}Na due to the magnetic dipole interaction and the electric quadrupole interaction.

4.1.10 Order of magnitude of hyperfine structure

If one examines the magnetic hyperfine structure in Eq. 4.15, it is obviously quite similar to the expression for the fine structure expression in Chapter 3. The ratio is

$$\frac{E^{\text{mag hfs}}}{E^{\text{fs}}} \approx g_I \left(\frac{m}{M_p} \right) \frac{1}{Z} \quad (4.39)$$

which is typically 10^{-3} to 10^{-4} . For a neutral atom, one can estimate

$$\frac{E_{\text{mag}}^{\text{hfs}}}{\hbar} = (\ell + 3/4)^{-3} \text{ GHz} \quad (4.40)$$

with a factor of 10 spread in either direction.

The quadrupole interaction is generally considerably smaller. An estimate is

$$B/\hbar \approx 0.01 Z(\ell + 1/2)^{-3} \text{ GHz} \quad (4.41)$$

Thus one generally expects that magnetic hyperfine structure dominates electric hyperfine structure in atoms. The opposite is generally true in molecules for two reasons: unpaired electrons are relatively rare, and the molecular binding mechanism can create large electric field gradients at the sites of the nuclei.

In concluding this discussion of hyperfine structure in atoms I would like to point out that the preceding formulae are, except in hydrogenic atoms, only approximations and never permit one to extract the nuclear dipole or quadrupole moment with the full accuracy of laser spectroscopy experiment—let alone R.F. spectroscopy experiments. Thus A and B in the combined hyperfine energy formula

$$E^{\text{hfs}} = E_{\text{mag}}^{\text{hfs}} + E_{\text{el}}^{\text{hfs}} = \frac{1}{2}AC + BC(C + 1)$$

(with C from Eq. 4.38) should be regarded primarily as empirical constants from the standpoint of atomic physics. Even if the problems of connecting A and B with the nuclear moments could be solved, the principal result would be better measurements of nuclear properties.

4.2 Isotope Effects

When comparing the spectral lines originating from atoms whose nuclei differ only in the number of neutrons (eg. different isotopes of the same element), effects due to the finite mass and volume of the nucleus become apparent. Even neglecting hyperfine structure (by taking the center of gravity of the observed splitting), the spectral lines of the different isotopes vary slightly in position—generally at the many parts per million level. The difference between the lines of the various isotopes is referred to as the isotope shift: it is observed to have both positive (heavier isotope has higher energy spacing) and negative values.

General speaking, light ($A < 40$) elements have positive frequency shift whereas heavy elements ($A > 60$) have negative shifts. This reflects the contribution of two distinct physical processes to the shift; the finite mass shift (almost always positive), and the nuclear volume shift (almost always negative). These will be discussed separately.

4.2.1 Mass effect

The origin of the mass effect is obvious from the Bohr energy level formula

$$E_n = E_n^\circ \left(\frac{M}{m+M} \right) \approx E_n^\circ \left(1 - \frac{m}{M} \right) \quad (4.42)$$

where the term involving m/M comes from solving the two body electron-nucleus (of mass M) system using the relative coordinate and associated reduced mass. Obviously increasing M increases E_n .

In two (or more) electron atoms the situation becomes more complicated due to the relative motion of the electrons. It would, for example, be possible to arrange the electrons symmetrically on opposite sides of the nucleus in which case there would be zero isotope effect. The virial theorem assures us that the mean value of the kinetic energy equals the negative of the total energy, so if we

treat the nuclear motion as a perturbation on a fixed nucleus solution, the mass effect will be:

$$\Delta E_{n,M} = \frac{-p^2}{2M} = -\frac{1}{2M} \left[\sum_i \mathbf{p}_i \right]^2 = -\frac{m}{M} \left[\underbrace{\frac{1}{2m} \sum p_i^2}_{\text{Normal Shift}} + \underbrace{\frac{1}{2m} \sum_{i \neq j} \mathbf{p}_i \cdot \mathbf{p}_j}_{\text{Specific Shift}} \right] \quad (4.43)$$

The first term is called the normal shift since (using the virial theorem again) it is

$$\Delta E_{n,M}^{\text{Normal}} = -\frac{m}{M} E_n^0 \quad (4.44)$$

The second term is called specific because it depends on the atom's quantum state. A discussion can be found in Sobel'man (pp. 224-6). $\Delta E^{\text{Specific}} = 0$ unless there are two or more valence electrons. For electronic configuration specified by quantum numbers n, s, ℓ , Sobel'man finds:

$$\Delta E_{ns,n'\ell}^{\text{specific}} = (1 - 2S) \frac{m}{M} \frac{3f_{ns,n'\ell}}{2} \hbar\omega_{n's',n\ell} \quad (4.45)$$

where $3f_{ns,n'\ell}$ is the oscillator strength (see Chapter 6). Thus the specific shift has opposite signs for $S = 1$ and $S = 0$ states—a reflection of the fact that the specific isotope shift is closely related to the exchange interaction. Eq. 4.45 also reflects the general result that $\Delta E^{\text{Specific}} = 0$ unless the two electrons are connected by an allowed dipole transition (otherwise f will vanish). Furthermore the specific isotope shift is of the same order of magnitude as the normal isotope shift: for $f > 2/3$, in fact, it can be larger (reversing the sign of the mass dependence of the isotope effect.)

The preceding discussion shows that the fractional energy shift of a level due to the mass of the nucleus decreases rapidly with increasing mass of the nucleus. The normal part of this shift has a variation in the fractional magnitude due to a change ΔM in the mass of the isotope of

$$\frac{\Delta E_{n,M+\Delta M}^{\text{Normal}} - \Delta E_{n,M}^{\text{Normal}}}{E_n} = \left(-\frac{m}{M + \Delta M} + \frac{m}{M} \right) = \frac{m}{M} \left(\frac{\Delta M}{M} \right) \quad (4.46)$$

which decreases as M^{-2} , reaching 10 parts per million for a nucleus with $A = 54$ (assuming $\Delta M = 1$).

4.2.2 Volume effect

Inside the nucleus, the electrostatic potential no longer behaves like Ze/r , but is reduced from this value. If the valence electron(s) penetrate significantly into this region (eg. for s electrons) then its energy will rise, relative to the value for a point nucleus, because of this reduced potential. Adding neutrons to the nucleus generally spreads out the charge distribution, causing a further rise in its energy. This reduction in the binding energy results in a decrease of the transition energy and therefore to a negative mass shift (assuming that the s state is the lower energy state involved in the transition).

For an s state, the density of the electron probability distribution at the nucleus is given by the semi-empirical Fermi-Segré formula [1]:

$$|\Psi_s(0)|^2 = \frac{Z_a^2 Z}{\pi a_o^2 n^{*3}} \left(1 + \left| \frac{\partial \delta_s}{\partial n} \right| \right) \quad (4.47)$$

where δ_s is the quantum defect and $Z_a e$ the charge of the atomic core. Combining this with a model of the nuclear charge cloud results (Sobel'man p. 229) in the nuclear volume correction to the energy (of an s electron):

$$\Delta E_n^V = Z_a^2 \frac{R_\infty}{n^{*3}} \left(1 + \left| \frac{\partial \delta_s}{\partial n} \right| \right) C \quad (4.48)$$

with

$$C = \frac{4(\gamma + 1)}{[\Gamma(2\gamma + 1)]^2} B(\gamma) \left(\frac{2Zr_o}{a_o} \right)^{2\gamma} \frac{\delta r_o}{r_o} \quad (4.49)$$

where

$$\gamma = [1 - \alpha^2 Z^2]^{1/2} \quad (4.50)$$

Γ is the gamma function $\Gamma(N + 1) = N!$, $B(\gamma)$ is a factor which depends on the nuclear charge distribution. For a charged shell

$$B(\gamma) = (2\gamma + 1)^{-1} \quad (4.51)$$

and for a uniform charge

$$B(\gamma) = (2\gamma + 1)^{-1} \left(\frac{3}{2\gamma + 3} \right) \quad (4.52)$$

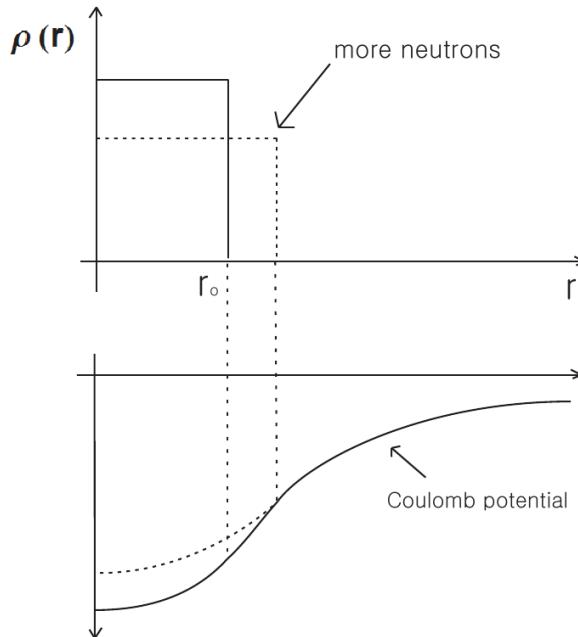


Figure 19. Simplified example of volume effect.

The nuclear radius is taken as (for atomic number A)

$$r_o = 1.15 \times 10^{-15} A^{1/3} m. \quad (4.53)$$

so that

$$\frac{\delta r_o}{r_o} = \frac{\delta A}{3A} \quad (4.54)$$

There are obviously a number of assumptions in these equations, and they should not be expected to work as well as expressions for the nuclear mass shift. Sobelman states that the observed shift is generally 1/2 to 3/4 of the one given above except for non-spherical nuclei (eg. rare-earth nuclei) which have anomalously large shifts.

References

- [1] E.Fermi and E. Segré, Memorie dell'Accademia d'Italia **4**(Fisica), 131. An English translation is in Hindmarsh, Pergamon Press 1967, p. 259.
- [2] R.M. Sternheimer, Phys. Rev. **164**, 10 (1967).

Chapter 5

Atoms in Magnetic Fields

5.1 The Landé g -factor

In this section we treat the interaction of the electron's orbital and spin angular momentum with external static magnetic fields. Previously, in the chapter on fine structure, we have considered the spin-orbit interaction: the coupling of electron spin to the magnetic field generated by the nucleus (which appears to move about the electron in the electron's rest frame). The spin orbit interaction causes the orbital and spin angular momenta of the electron to couple together to produce a total spin which then couples to the external field; the magnitude of this coupling is calculated here for weak external fields.

5.1.1 Magnetic moment of circulating charge (classical)

The energy of interaction of a classical *magnetic moment* μ with a magnetic field \mathbf{B} is

$$U = -\mu \cdot \mathbf{B} \quad (5.1)$$

indicating that the torque tends to align the moment along the field. In classical electrodynamics the magnetic moment of a moving point particle about some point in space is independent of the path which it takes, but depends only on the product of the ratio of its charge to mass m , and angular momentum ℓ . This result follows from the definitions of angular momentum

$$\mathbf{L} \equiv \mathbf{r} \times \mathbf{p} = \mathbf{m}[\mathbf{r} \times \mathbf{v}] \quad (5.2)$$

and magnetic moment

$$\boldsymbol{\mu} \equiv \frac{1}{2}\mathbf{r} \times \mathbf{i} = \frac{q}{2}[\mathbf{r} \times \mathbf{v}] \quad (5.3)$$

where \mathbf{i} is the current and \mathbf{v} the velocity (see Jackson Ch.5). The equality of the bracketed terms implies

$$\boldsymbol{\mu} = \frac{q}{2m}\mathbf{L} \equiv \gamma_\ell \mathbf{L} \quad (5.4)$$

where γ_ℓ is referred to as the *gyromagnetic ratio*. This is a general result for any turbulently rotating blob provided only that it has a constant ratio of charge to mass throughout.

For an electron with orbital angular momentum ℓ

$$\boldsymbol{\mu}_\ell = -\frac{e}{2m}\mathbf{L} \equiv -\mu_B\mathbf{L}/\hbar \quad (5.5)$$

which is the classical result, and μ_B is the *Bohr magneton*:

$$\mu_B = \frac{e\hbar}{2m} = 9.27408(4) \times 10^{-24} J T^{-1} \rightarrow 1.39983 \times 10^4 \text{ MHz} \times B / (\text{Tesla}) \quad (5.6)$$

5.1.2 Intrinsic electron spin and magnetic moment

When Uhlenbeck and Goudsmit suggested [1] that the electron had an intrinsic spin $S = \frac{1}{2}$, it soon became apparent that it had a magnetic moment twice as large as would be expected on the basis of Eq. 5.4. (This implies that the electron cannot be made out of material with a uniform ratio of charge to mass.) This is accounted for by writing for the *intrinsic electron moment*

$$\boldsymbol{\mu}_s = -g_e\mu_B\mathbf{S}/\hbar \quad (5.7)$$

where the quantity $g_e = 2$ is called the electron *g-factor*. (The negative sign permits treating g_e as a positive quantity, which is the convention.) This factor was predicted by the Dirac theory of the electron, probably its greatest triumph. Later, experiments by Kusch, followed by Crane et al., and then by Dehmelt and coworkers, have shown (for both electrons and positrons).

$$\frac{g_e}{2} = 1.0011596521869(41) \quad (5.8)$$

This result has been calculated from quantum electrodynamics, which gives

$$\frac{g_e}{2} = 1 + \frac{1}{2}\left(\frac{\alpha}{\pi}\right) - 0.3258\left(\frac{\alpha}{\pi}\right)^2 + 0.13\left(\frac{\alpha}{\pi}\right)^3 + \dots \quad (5.9)$$

The agreement between the prediction of quantum electrodynamics and experiment on the electron g-factor is often cited as the most precise test of theory in all of physics.

5.1.3 Vector model of the Landé *g*-factor

In zero or weak magnetic field the spin orbit interaction couples \mathbf{S} and \mathbf{L} together to form $\mathbf{J} = \mathbf{L} + \mathbf{S}$, and this resultant angular momentum interacts with the

applied magnetic field with an energy

$$U = -g_j \mu_B \mathbf{B} \cdot \mathbf{J}/\hbar \quad (5.10)$$

which defines g_j .

The interaction of the field is actually with $\boldsymbol{\mu}_s$ and $\boldsymbol{\mu}_\ell$, however g_j is not simply related to these quantities because $\boldsymbol{\mu}_s$ and $\boldsymbol{\mu}_\ell$ precess about \mathbf{J} instead of the field. As Landé showed in investigations of angular momentum coupling of different electrons [2], it is a simple matter to find g_j by calculating the sum of the projections of $\boldsymbol{\mu}_s$ and $\boldsymbol{\mu}_\ell$ onto \mathbf{J} .

The projection of $\boldsymbol{\mu}_\ell$ on \mathbf{J} is

$$\mu_{\ell j} = \frac{-\mu_B |\mathbf{L}|}{\hbar} \frac{\mathbf{L} \cdot \mathbf{J}}{|\mathbf{L}| |\mathbf{J}|} \quad (5.11)$$

The projection of $\boldsymbol{\mu}_s$ on \mathbf{J} is

$$\mu_{sj} = -g_e \mu_B \frac{|\mathbf{S}|}{\hbar} \frac{\mathbf{S} \cdot \mathbf{J}}{|\mathbf{S}| |\mathbf{J}|} \quad (5.12)$$

The definition of g_j gives

$$g_j = -\frac{(\mu_{\ell j} + \mu_{sj})}{|\mathbf{J}| \mu_B / \hbar} \quad (5.13)$$

Taking $g_e = 2$

$$g_j = \frac{\mathbf{L} \cdot (\mathbf{L} + \mathbf{S}) + 2\mathbf{S} \cdot (\mathbf{L} + \mathbf{S})}{|\mathbf{J}|^2} \quad (5.14)$$

$$= 1 + \frac{j(j+1) + s(s+1) - \ell(\ell+1)}{2j(j+1)} \quad (5.15)$$

using $2\mathbf{L} \cdot \mathbf{S} = \mathbf{J}^2 - \mathbf{L}^2 - \mathbf{S}^2$.

If a transition from a level with angular momentum j' is to a level with j'' takes place in a magnetic field, the resulting spectral line will be split into three or more components—a phenomenon known as the *Zeeman effect*. For transitions with a particular Δm , say $\Delta m = -1$, the components will have shifts

$$\Delta E_{z,m,-1} = [g_{j'} m - g_{j''} (m-1)] \mu_B B = [(g_{j'} - g_{j''}) m - g_{j''}] \mu_B B \quad (5.16)$$

If $g_{j'} = g_{j''}$ (or if j' or $j'' = 0$) then $\Delta E_{z,m,-1}$ will not depend on m (or there will be only one transition with $\Delta m = -1$) and there will be only 3 components of the line ($\Delta m = +1, 0, -1$); this is called the normal Zeeman splitting. If neither of these conditions holds, the line will be split into more than 3 components and the Zeeman structure is termed “anomalous”—it can’t be explained with classical atomic models.

5.2 Hyperfine structure in an applied field

The Hamiltonian in an applied field \mathbf{B}_0 is

$$H = ah\mathbf{I} \cdot \mathbf{J} - \boldsymbol{\mu}_J \cdot \mathbf{B}_0 - \boldsymbol{\mu}_I \cdot \mathbf{B}_0 = ah\mathbf{I} \cdot \mathbf{J} + g_j\mu_B\mathbf{J} \cdot \mathbf{B}_0 - g_I\mu_B\mathbf{I} \cdot \mathbf{B}_0 \quad (5.17)$$

By convention, we take $\boldsymbol{\mu}_J = -g_j\mu_B\mathbf{J}$. Note that we are expressing the nuclear moment in terms of the Bohr magneton, and that $g_I \ll g_j$. (The nuclear moment is often expressed in terms of the nuclear magneton, in which case $\boldsymbol{\mu}_I = g'_I\mu_N\mathbf{I}$, where μ_N is the nuclear magneton.) What are the quantum numbers and energies? Before discussing the general solution, let us look at the limiting cases.

5.2.1 Low field

The total angular momentum is $\mathbf{F} = \mathbf{I} + \mathbf{J}$. In low field, F and m_F are good quantum numbers. Each level F contains $(2F+1)$ degenerate states. In a weak field \mathbf{B}_0 the $(2F+1)$ fold degeneracy is lifted. We can treat the terms

$$H_z = -(\boldsymbol{\mu}_j + \boldsymbol{\mu}_I) \cdot \mathbf{B}_0 \quad (5.18)$$

as a perturbation. \mathbf{J} and \mathbf{I} are not good quantum numbers, only their components parallel to \mathbf{F} are important. Thus

$$\langle \mathbf{J} \cdot \mathbf{B}_0 \rangle = \frac{\langle \mathbf{J} \cdot \mathbf{F} \rangle \mathbf{F} \cdot \mathbf{B}_0}{F^2} \quad (5.19)$$

$$H_z = -\mu_B[-g_j(\mathbf{J} \cdot \mathbf{F}) + g_I(\mathbf{I} \cdot \mathbf{F})] \frac{\mathbf{F} \cdot \mathbf{B}_0}{F^2} \quad (5.20)$$

Since $g_I \ll g_j$, we can usually neglect it. We can rewrite this result as

$$H_z = g_F\mu_B m_B \quad (5.21)$$

$$g_F = \frac{\langle \mathbf{J} \cdot \mathbf{F} \rangle}{F^2} g_j = \frac{g_j}{2} \frac{F(F+1) + j(j+1) - I(I+1)}{F(F+1)} \quad (5.22)$$

For example, let $I = 3/2, j = 1/2; F = 2, 1$. Then

$$F = 2; \quad W(2) = (3/4)ah; \quad g_F = g_j/4 \quad (5.23)$$

$$F = 1; \quad W(1) = -(5/4)ah; \quad g_F = -g_j/4 \quad (5.24)$$

5.2.2 High field

If $\boldsymbol{\mu}_j \cdot \mathbf{B}_0 \gg ah\mathbf{I} \cdot \mathbf{J}$, then \mathbf{J} is quantized along \mathbf{B}_0 . Although $\boldsymbol{\mu}_I \cdot \mathbf{B}_0$ is not necessarily large compared to the hyperfine interaction, the $\mathbf{I} \cdot \mathbf{J}$ coupling assures that \mathbf{I} is also quantized along \mathbf{B}_0 . Thus m_I and m_j are good quantum numbers. In this case, Eq. 5.17 can be written

$$H = ahm_i m_j + g_j\mu_B m_j B_0 - g_I\mu_B m_I B_0 \quad (5.25)$$

The second term on the right is largest. Usually the first term is next largest, and the nuclear terms is smallest. The diagram below shows low and high field behavior for hyperfine structure for $I = 3/2, j = 1/2$.

5.2.3 General solution

Finding eigenfunctions and eigenvalues of the hyperfine Hamiltonian for arbitrary field requires diagonalizing the energy matrix in some suitable representation. To obtain a rough idea of the expected results, one can smoothly connect the energy levels at low and high field, bearing in mind that $m = m_I + m_j$ is a good quantum number at all fields.

For $J = 1/2$, the eigenvalues of (Eq. 5.17) can be found exactly. The energies are given by the Breit-Rabi formula

$$W(m) = -\frac{1}{2} \frac{\Delta W}{2I+1} - g_I \mu_B B_0 m \pm \frac{\Delta W}{2} \sqrt{1 + \frac{4mx}{2I+1} + x^2}, \quad (5.26)$$

where the + sign is for $F = I + 1/2$, and the – sign is for $F = I - 1/2$. ΔW is the zero field energy separation.

$$\Delta W = W(F = I + 1/2) - W(F = I - 1/2) = ah \left(\frac{2I+1}{2} \right) \quad (5.27)$$

The parameter x is given by

$$x = \frac{(g_e + g_I)\mu_B B_0}{\Delta W} \quad (5.28)$$

Physically, x is the ratio of the paramagnetic interaction (the “Zeeman energy”) to the hyperfine separation. The Breit-Rabi energy level diagram for hydrogen and deuterium are shown below. The units reflect current interest in atom trapping. Low-field quantum numbers are shown. It is left as an exercise to identify the high field quantum numbers.

References

- [1] G.E. Uhlenbeck and S. Goudsmit, Nature **117**, 264 (1926).
- [2] A. Landé, Zeitschrift für Physik **15**, 189 (1923), English translation on p. 186 of Hindmarsh.

Chapter 6

Atoms in Electric Fields

This section deals with how atoms behave in static electric fields. The method is straightforward, involving second order Rayleigh-Schrödinger perturbation theory. The treatment describes the effects of symmetry on the basic interaction, polarizability, and the concept of oscillator strength.

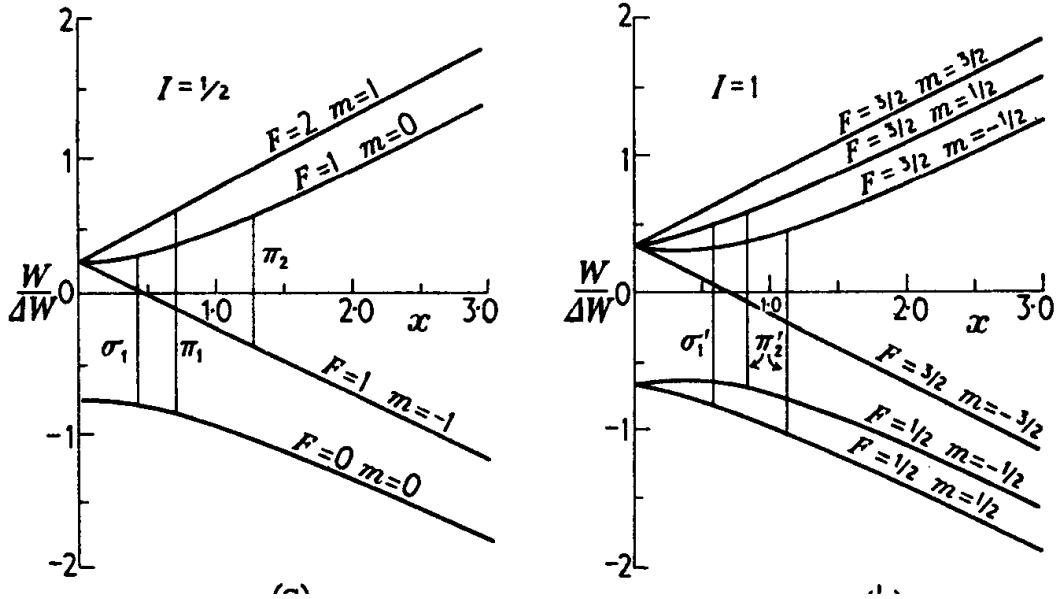


Figure 20. Energy level structure for a single-electron atom with nucleaar spin $I = 1/2$, such as hydrogen (left), and $I = 1$, such as deuterium (right). From *Molecular Beams* by N.F. Ramsey.

Let us review the concept of parity. Parity is a consequence of space inversion.

$$x \rightarrow -x \quad (6.1)$$

$$y \rightarrow -y \quad (6.2)$$

$$z \rightarrow -z \quad (6.3)$$

$$\text{or } \mathbf{r} \rightarrow -\mathbf{r} \quad (6.4)$$

We propose an operator that (in the spirit of the rotation operator introduced earlier) takes an initial ket and returns a ket with the above inversion operation performed.

$$|\alpha\rangle \rightarrow \pi|\alpha\rangle \quad (6.5)$$

We require that this operator is unitary and that it has the following key property (or, perhaps more precisely we define the operator through)

$$\pi^\dagger \mathbf{x} \pi = -\mathbf{x} \quad (6.6)$$

which implies

$$\mathbf{x}\pi|\mathbf{x}_1\rangle = -\pi\mathbf{x}|\mathbf{x}_1\rangle = -\pi\mathbf{x}_1|\mathbf{x}_1\rangle = -\mathbf{x}_1\pi|\mathbf{x}_1\rangle \quad (6.7)$$

To put the finest point on it, $\pi|\mathbf{x}'\rangle$ is an eigenket of \mathbf{x} with eigenvalue of $-\mathbf{x}'$.

Finally, the eigenvalues of π are ± 1 and

$$\pi^{-1} = \pi = \pi^\dagger \quad (6.8)$$

Position is "odd" under space inversion or "odd under the parity operator". Angular momentum, on the other hand is even.

$$\pi^\dagger \mathbf{L}\pi = \pi^\dagger \mathbf{x} \times \mathbf{p}\pi = \mathbf{x} \times \mathbf{p} = \mathbf{L} \quad (6.9)$$

Because of this property position and momentum are called vectors or polar vectors and angular momentum is called an axial or psuedo vector. What about wavefunction? What does the parity operator do to wavefunctions? Well it depends on the wavefunction. For example, consider the spherical harmonics (the angular part of the hydrogen atom eigenstates). Some of the wavefunctions are odd under parity and some are even. (In one dimension a cosine wave is "even" whereas a sine wave is "odd".)

$$\pi|Y_l^m\rangle = (-1)^l|Y_l^m\rangle \quad (6.10)$$

Now, consider the case where a state is an energy eigenket and the parity operator commutes with Hamiltonian. Such a ket is not necessarily an eigenket of the parity operator. Consider, for example, the case of the hydrogen atom for $n = 2$. Neglecting higher order pertubations to the hamiltonian, $|n = 2\rangle$ can be made up of a combination of two eigenkets with different parities,

$$|n = 2\rangle = a_1|n = 2, l = 0, m\rangle + a_2|n = 2, l = 1, m'\rangle \quad (6.11)$$

Without any degeneracies eigenstates of the hamiltonian are indeed eigenstates of the parity operator if the hamiltonian and π commute. This idea of parity gives rise to what is called a selection rule. Selection rules, in general, are nothing more than the statement that certain operators connect certain states ($\langle i|A|j\rangle \neq 0$ for certain i, j) and do not connect other states (that is, $\langle i'|A|j'\rangle = 0$ for certain i', j'). Consider, for example, the \mathbf{x} operator and two different parity eigenstates,

$$\pi|\alpha\rangle = p_\alpha|\alpha\rangle \quad \pi|\beta\rangle = p_\beta|\beta\rangle \quad p_{\alpha,\beta} = \pm 1 \quad (6.12)$$

then

$$\langle\beta|\mathbf{x}|\alpha\rangle = 0 \text{ unless } p_\alpha = -p_\beta \quad (6.13)$$

One can see this in the following way

$$\langle\beta|\mathbf{x}|\alpha\rangle = \langle\beta|\pi^{-1}\pi\mathbf{x}\pi^{-1}\pi|\alpha\rangle = p_\alpha p_\beta (-1)\langle\beta|\mathbf{x}|\alpha\rangle \quad (6.14)$$

which can be true only if $p_\alpha p_\beta = -1$. \mathbf{x} is an "parity odd operator" and it connects states of opposite parity. "Even operators" connect states of the same parity.

6.1 Atoms in a Static Electric Field

We can use this basic idea in understanding the problem of an atom subjected to an electric field. We begin by writing down the potential due to a collection of charges,

$$\Phi(\mathbf{x}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l q_{lm} Y_{lm}^m(\theta, \phi) \frac{C}{r^{l+1}} \quad (6.15)$$

where

$$q_{lm} \equiv \int Y_l^{m*} r^l \rho(\mathbf{x}) d^3x \quad (6.16)$$

where $\rho(x)$ is the charge distribution. q_{00} is the total charge, q_{1x} are the dipole moments, q_{2x} are the quadrupole moments, etc. The energy U of an overall neutral collection of charges in an electric field $\mathcal{E} = \mathcal{E}\hat{z}$ can similarly be expanded as

$$U = -\mathbf{d} \cdot \hat{z}\mathcal{E} - \frac{1}{2}\alpha\mathcal{E}^2 + O(\mathcal{E}^3) + \dots \quad (6.17)$$

where d is the dipole and α is the polarizability.

Now we are in a better position to solve the problem of the hydrogen atom in a static electric field, $\mathcal{E} = \mathcal{E}\hat{z}$, just about the simplest example.

The hamiltonian for this problem can be written

$$H = H_0 + H' = H_0 + e\mathcal{E}z \quad (6.18)$$

where H_0 is the "unperturbed" hamiltonian for the hydrogen atom.

We chose to solve this via matrix methods. The first step is to write down the matrix elements for the hamiltonian in a basis of our choosing. Let's try with the $|nlm\rangle$ basis kets, the eigenkets of H_0 . So, H_0 only contributes diagonal elements to the matrix, E_n . As e and E are scalars, not operators, we need only consider the effect of z . First, z is a parity odd operator, connecting only states of different parity. Thus H' contributes nothing to the diagonal entries nor to any entries with the same angular momentum, l . States of the same parity but whose angular momentum differ by more than $\Delta l = \pm 1$ also result in zero because ... Finally, H' also only connects states of the same m . One can see this by noting that

$$z = Y_1^0 r \sqrt{\frac{4\pi}{3}} = r \cos \theta \quad (6.19)$$

which is an even function in θ . Any states differing by $\Delta m = \pm 1$ would then result in an integral of two even functions (one of those being the $\cos \theta$ originating from the z) and an odd function in θ which is zero. This resulta can also be seen directly by noting a result of the Wigner-Eckhart theorem that $\langle njm|z|njm'\rangle = \alpha(n, j)\langle njm|J_z|njm'\rangle$ where α is just a number. Thus, we produce the "selection rules" for the $\mathbf{E} = E\hat{z}$ operator,

$$\Delta m = 0 \quad l = \pm 1 \quad (6.20)$$

NOTE that this strictly applies only the this specific operator. If E were pointing in some other direction then things might (and do) change.

The matrix for the our hamiltonian reads then

$$\begin{pmatrix} E_1 & 0(e/o) & 0(e/o) & eE\langle z \rangle & 0(p) & \dots \\ & E_2 & 0(p) & 0(p) & 0(p) & \\ & & E_2 & 0(p) & 0(e/o) & \\ & & & E_2 & eE\langle z \rangle & \\ & & & & E_2 & \end{pmatrix} \quad (6.21)$$

where the entries arranged in $|100\rangle, |211\rangle, |21 - 1\rangle, |210\rangle, |200\rangle$ order. The 0's are designated with an indication of "why" those particular entries in the matrix are zero, e/o meaning even/odd ($\Delta m = 0$ selection rule) and p meaning parity ($\Delta l = \pm 1$ selection rule). As mentioned above, the H' contribution to the diagonal elements is zero due to parity. Because the $n = 2$ states are degenerate, degenerate perturbation theory must be used to solve the problem. Of course we know that in reality the problem is more complex than this. Both fine, hyperfine and the Lamb shift have been neglected. Solving the problem taking this into account would indicate the use of second order perturbation theory.

To see how this all shakes out, let's go ahead and apply perturbation theory directly.

$$E_n^{(1)} = \langle n | H' | n \rangle = 0 \text{ by parity} \quad (6.22)$$

$$E_n^{(2)} = \sum_{m,m \neq n} \frac{|\langle n | H' | m \rangle|^2}{E_n - E_m} = e^2 \mathcal{E}^2 \sum_{m,m \neq n} \frac{|\langle n | z | m \rangle|^2}{E_n - E_m} \quad (6.23)$$

$$|n^{(1)}\rangle = \sum_{m,m \neq n} |\langle n | H' | m \rangle|^2 \quad (6.24)$$

If one is in the case where this simple perturbation theory does not work because of degenerate states (leading to $E_m - E_n = 0$ in the denominator then it is best just to diagonalize the Hamiltonian in relation to H' . If you do that for the case of $n = 2$ you find that the eigenstates are

$$|n = 2, l = 1, m = 1\rangle \quad E_2 \quad (6.25)$$

$$|2, 1, -1\rangle \quad E_2 \quad (6.26)$$

$$\frac{1}{\sqrt{2}}(|2, 1, 0\rangle + |2, 0, 0\rangle) \quad E_2 + cE \quad (6.27)$$

$$\frac{1}{\sqrt{2}}(|2, 1, 0\rangle - |2, 0, 0\rangle) \quad E_2 - cE \quad (6.28)$$

where c is a constant. The last two states have a linear response to the electric field, or a linear Stark effect. Even is there were a small splitting between the different states in the $n = 2$ manifold, if the field interaction were higher than the splitting then there would be also be a linear Stark effect. At lower fields the interaction would be second order (second order perturbation theory would be called for) and the response would be quadratic in the applied electric field. Notice the the new eigenstates are a mixture of states of different parity. This mixture allows for a *dipole* to be formed and it is the interaction of the electric field with this dipole that gives rise to a linear response to the field. It is this dipole that is talked about by chemists when they say that a molecule "has

a dipole moment". Molecules "have dipole moments" because they have closely lying states of opposite parity so small fields put them in the linear Stark regime. But make no mistake, at low enough fields, the response would be quadratic, just as it is in the case of atoms.

Now, all of this has been talked about under the (essentially correct) assumption that $[H, \pi] = 0$ and, therefore, that the eigenstates of the H atom are also parity eigenstates. But what if $[H, \pi] \neq 0$? This occurs when the weak force is involved and will likely be present in nature and described, eventually, by extensions to the Standard Model. Such mechanisms can lead to the presence of permanent electric dipole moments of elementary particles.

6.2 Some Results of Stationary Perturbation Theory

For reference, we recapitulate some elementary results from perturbation theory. Assume that the Hamiltonian of a system may be written as the sum of two parts

$$H = H_0 + H' \quad (6.29)$$

and that the eigenstates and eigenvalues of H_0 are known:

$$H_0|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle \quad (6.30)$$

If it is not possible to find the eigenvalues of H exactly, it is possible to write power series expressions for them that converge over some interval. If H' is time independent, the problem is stationary and the appropriate perturbation theory is Rayleigh-Schrödinger stationary state perturbation theory, described in most texts in quantum mechanics. We write

$$E_n = E_n^{(0)} + E_n^{(1)} + E_n^{(2)} + \dots \quad (6.31)$$

$$|n\rangle = |n^{(0)}\rangle + |n^{(1)}\rangle + \dots \quad (6.32)$$

and express the $(i+1)^{\text{th}}$ order perturbation in terms of $E^{(i)}$ and $|n^{(i)}\rangle$. The energies are given by

$$E_n^{(m)} = \langle n^{(0)} | H' | n^{(m-1)} \rangle \quad (6.33)$$

We shall only use the lowest two orders here. The first order results are

$$E_n^{(1)} = \langle n | H | n \rangle \quad (6.34)$$

$$|n^{(1)}\rangle = \sum_m' \frac{|m\rangle \langle m| H' |n\rangle}{E_n - E_m} \quad (6.35)$$

The symbol \sum' indicates that the term $m = n$ is excluded. It is understood that the sum extends over continuum states. Note that the state function is not properly normalized, but that the error is quadratic in H' .

The second order results are

$$E_n^{(2)} = \sum_n' \frac{|\langle m|H'|n \rangle|^2}{E_n - E_m} \quad (6.36)$$

$$|n^{(2)}\rangle = \sum_m' |m\rangle \left[\frac{\langle m|H'|n \rangle}{E_n - E_m} \left[1 - \frac{\langle n|H'|n \rangle}{E_n - E_m} \right] + \sum_p' \frac{\langle m|H'|p \rangle \langle p|H'|n \rangle}{(E_n - E_m)(E_n - E_p)} \right] \quad (6.37)$$

In second order perturbation theory the effect of a coupling of n and m by H' is to push the levels apart, independent of the value of H'_{nm} . Consequently, states coupled by H' always repel each other.

6.3 Perturbation Theory of Polarizability

We turn now to finding the energy and polarizability of an atom in a static field along the $+z$ direction. We apply perturbation theory taking H_0 to describe the unperturbed atomic system and

$$H' = -\mathbf{d} \cdot \hat{\mathbf{z}}\mathcal{E} = ez\mathcal{E} \quad (6.38)$$

As discussed in Sect. ??, parity requires that $H'_{mm} = 0$ so the first order perturbation vanishes. To second order, the energy is given by

$$E_n = E_n^{(0)} - e^2\mathcal{E}^2 \sum_m' \frac{|\langle m|z|n \rangle|^2}{E_m - E_n} \quad (6.39)$$

If we compare this results with the potential energy of a charge distribution interacting with an electric field, (Eq. ??), we can identify the polarizability interaction with the second term in this equation. As a result the polarizability in state n is given by

$$\alpha_n = 2e^2 \sum_m' \frac{|\langle m|z|n \rangle|^2}{E_m - E_n} \quad (6.40)$$

Note that this has the dimensions of length³, i.e. volume.

The induced dipole moment can be found from the polarization.

$$\mathbf{d} = \alpha\mathcal{E}\hat{\mathbf{z}} = 2e^2\mathcal{E}\hat{\mathbf{z}} \sum_m' \frac{|\langle m|z|n \rangle|^2}{E_m - E_n} \quad (6.41)$$

An alternative way to calculate the dipole moment is to calculate the expectation value of the dipole operator, Eq. ??, using the first order perturbed state vector.

$$\mathbf{d}_{nm} = (\langle n^{(0)} | + \langle n^{(1)} |) \mathbf{d} (|n^{(0)}\rangle + |n^{(1)}\rangle) \quad (6.42)$$

$$= 2\text{Re}[\langle n^{(0)} | \mathbf{d} | n^{(1)} \rangle] = 2e^2\text{Re} \left[\sum_{s,m} \frac{\langle n^{(0)} | s|m \rangle \langle m|z|n^{(0)} \rangle}{E_m - E_n} \right] \hat{\mathbf{s}} \cdot \hat{\mathbf{z}}\mathcal{E} \quad (6.43)$$

where the sum is over $s = x, y, z$. Only the term $s = z$ will contribute, and it will yield an interaction energy in agreement with Eq. 6.39.

As an example, for the ground state of hydrogen we can obtain a lower limit for the polarizability by considering only the contribution to the sum of the $2P$ state. Values for the various moments in hydrogen are given in Bethe and Salpeter, Section 63. Using $| < 2P | r | 1S > |^2 = 1.666$, and $E_{2p} - E_{1S} = 3/8$, we obtain $\alpha = 2.96$ atomic units (i.e. $2.96 a_0^3$).

The polarizability of the ground state of hydrogen can be calculated exactly. It turns out that the $2P$ state makes the major contribution, and that the higher bound states contribute relatively little. However, the continuum makes a significant contributions. The exact value is 4.5.

To put this polarizability in perspective, note that the potential of a conducting sphere of radius R in a uniform electric field \mathcal{E} is given by

$$V(r, \theta) = -\mathcal{E} \cos \theta \left(r - \frac{R^3}{r^2} \right) \quad (r \geq R) \quad (6.44)$$

The induced dipole moment is $R^3 \mathcal{E}$, so that the polarizability is R^3 . For the ground state of hydrogen, $\bar{r}^3 = 2.75$, so to a crude approximation, in an electric field hydrogen behaves like a conducting sphere.

Polarizability may be approximated easily, though not accurately, using Unsöld's approximation in which the energy term in the denominator of Eq. 6.43 is replaced by an average energy interval $\overline{E_m} - E_n$. The sum can then be evaluated using the closure rule $\sum_m |m><m| = 1$. (Note that the term $m = n$ does not need to be excluded from the sum, since $< n | z | n > = 0$). With this approximation,

$$a_n = \frac{2e^2}{\overline{E_m} - E_n} \sum_m < n | z | m > < m | z | n > = \frac{2e^2 < n | z^2 | n >}{\overline{E_m} - E_n} \quad (6.45)$$

For hydrogen in the ground state, $\overline{z^2} = 1$. If we take the average excitation energy to be $\overline{E_m} = 0$, the result is $\alpha = 4$.

6.4 Beyond the quadratic Stark effect

It should be obvious from the previous discussion that the Stark effect for a state of g is quadratic only when

$$\epsilon << \frac{E_i - E_g}{e |\langle i | \mathbf{r} | g \rangle|} \quad (6.46)$$

when i is the nearest state of opposite parity to g .

If g is the ground state, we can expect $E_i = E_g \sim 0.5 E_g$, $E_g \sim 0.3$ Hartree and $| < r > |^{-1} \approx | < r^{-1} > | = 2E_g/e^2$ (virial theorem). Hence the Stark shift should be quadratic if the field is well below the critical value

$$\epsilon_{crit} = \frac{0.5 \times 2(0.3)^2 m^2 e^8}{e^3 \hbar^4} \approx 0.1 \frac{e}{a_0^2} \quad (6.47)$$

$[e/a_0^2 \text{ is atomic unit of field}] \approx 5 \times 10^8 \text{ V/cm}$ —a field three orders of magnitude in excess of what can be produced in a laboratory except in a vanishingly small volume.

If g is an excited state, say $|g\rangle = |n\ell\rangle$, this situation changes *dramatically*. In general, the matrix element $\langle n, \ell + 1 | \mathbf{r} | n, \ell \rangle \sim n^2 a_0^2$ and ΔE to the next level of opposite parity depends on the quantum defect:

$$\Delta E = E_{n, \ell+1} - E_{n, \ell} = \frac{-R_H}{(n - \delta_{\ell+1})^2} - \frac{-R_H}{(n - \delta_\ell)^2} \approx 2R_H(\delta_{\ell+1} - \delta_\ell)/n^{*3} \quad (6.48)$$

Thus the critical field is lowered to

$$\begin{aligned} \epsilon_{\text{crit}} &= \frac{\Delta E}{e \langle |\mathbf{r}| \rangle} = \frac{me^4}{\hbar^2} \frac{1}{ea_0} \frac{\delta_{\ell+1} - \delta_\ell}{n^{*5}} \\ &= \frac{e}{a_0^2} \frac{\delta_{\ell+1} - \delta_\ell}{n^5} \approx 5 \times 10^9 \frac{\delta_{\ell+1} - \delta_\ell}{n^{*5}} \frac{\text{volts}}{\text{cm}} \end{aligned} \quad (6.49)$$

Considering that quantum defects are typically $\leq 10^{-5}$ when $\ell \geq \ell_{\text{core}} + 2$ (ℓ_{core} is the largest ℓ of an electron in the core), it is clear that even 1 V/cm fields will exceed ϵ_{crit} for higher ℓ levels if $n > 7$. Large laboratory fields (10^5 V/cm) can exceed ϵ_{crit} even for S states if $n^* \geq 5$.

When the electric field exceeds ϵ_{crit} states with different ℓ but the same n are degenerate to the extent that their quantum defects are small. Once ℓ exceeds the number of core electrons, these states will easily become completely mixed by the field and they must be diagonalized exactly. The result is eigenstates possessing apparently permanent electric dipoles with a resulting linear Stark shift (see following figure). As the field increases, these states spread out in energy. First they run into states with the same n but different quantum defects; then the groups of states with different n begin to overlap. At this point a matrix containing all n, ℓ states with ℓ greater or equal to m_ℓ must be diagonalized. The only saving grace is that the lowest n states do not partake in this strong mixing; however, the n states near the continuum always do if there is an e -field present.

The situation described above differs qualitatively for hydrogen since it has no quantum defects and the energies are degenerate. In this case the zero-field problem may be solved using a basis which diagonalizes the Hamiltonian both for the atom above and also in the presence of an electric field. This approach corresponds to solving the H atom in parabolic–ellipsoidal coordinates and results in the presence of an integral quantum number which replaces ℓ . The resulting states possess permanent dipole moments which vary with this quantum number and therefore have linear Stark effects even in infinitesimal fields. Moreover the matrix elements which mix states from different n manifolds vanish at all fields, so the upper energy levels from one manifold cross the lower energy levels from the manifold above without interacting with them.

The following example shows the high field stark effect for Li. Only the S

term in Li has an appreciable quantum defect, and it has been suppressed by selecting final states with $m_\ell = 1$.

The dramatic difference between the physical properties of atoms with $n > 10$ and the properties of the same atoms in their ground state, coupled with the fact that these properties are largely independent of the type of atom which is excited, justifies the application of the name Rydberg atoms to highly excited atoms in general.

6.5 Field ionization

If an atom is placed in a sufficiently high electric field it will be ionized, a process called *field ionization*. An excellent order of magnitude estimate of the field ϵ_{ion} , required to ionize an atom which is initially in a level bound by energy E can be obtained by the following purely classical argument: the presence of the field adds the term $U(z) = e\epsilon z$ to the potential energy of the atom. This produces a potential with a maximum $U_{\text{max}} < 0$ and the atom will ionize if $U_{\text{max}} < -E$.

The figure shows the combined potential as well as U_{atom} and U_{field}

$$U_{\text{total}}(z) = U_{\text{atom}}(z) + U_{\text{field}}(z) = \frac{-e^2}{|z|} + e\epsilon z \quad (6.50)$$

The appropriate maximum occurs at

$$z_{\text{max}} = -\left(\frac{e}{\epsilon}\right)^{1/2} \quad (6.51)$$

as determined from $dU/dz = 0$. Equating U_{max} and $-E$ gives

$$\epsilon_{\text{ion}} = \frac{E^2}{4e^2} = \frac{1}{16n^{*4}} = 3.2 \times 10^8(n^{*})^{-4} \text{ V/cm} \quad (6.52)$$

for level with energy $-E$ and quantum number n^* .

The predictions of this formula for E_{ion} is usually accurate within 20% in spite of its neglect of both quantum tunneling and the change in E produced by the field. [This latter deficiency is remedied in the comparison with Li data shown in the preceding part of this section because the eye naturally uses the ionization field appropriate to the perturbed energy of the state rather than its zero-field energy.] Tunneling manifests itself as a finite decay rate for states which classically lie lower than the barrier. The increase of the ionization rate with field is so dramatic, however, that the details of the experiment do not influence the field at which ionization occurs very much: calculations [?] show the ionization rate increasing from $10^5/\text{sec}$ to $10^{10}/\text{sec}$ for a 30% increase in the field.

Oddly enough the classical prediction works worse for H than for any other atom. This is a reflection of the fact that certain matrix elements necessary to mix the n, ℓ states (so the wave function samples the region near U_{min}) are rigorously zero in H, as discussed in the preceding part of this section. Hence the orbital ellipse of the electron does not precess and can remain on the side of

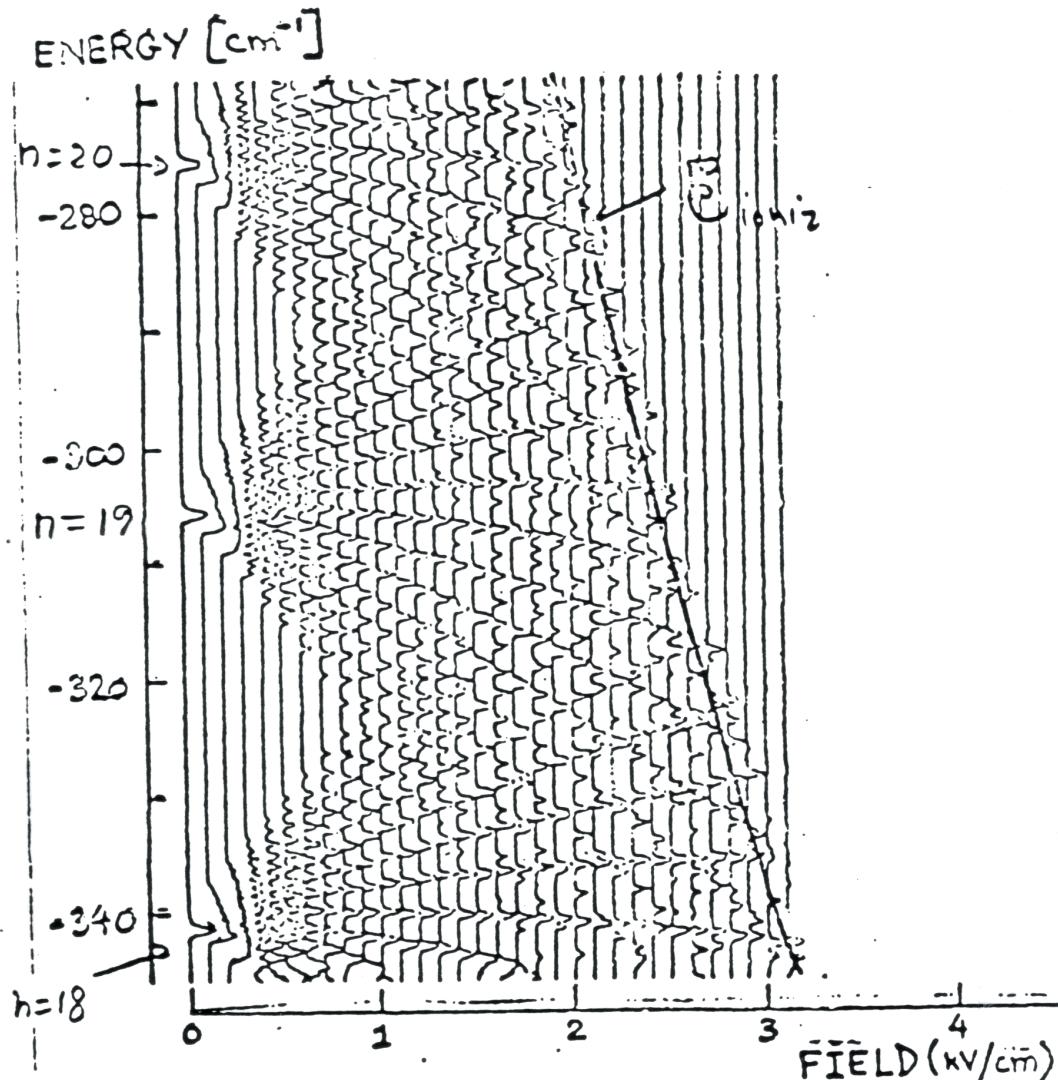


Figure 21. Stark effect and field ionization in Li for levels with $m = 1$. Each vertical line represents a measurement at that field of the number of atoms excited (from the $3s$ state) by radiation whose energy falls the indicated amount below the ionization limit. Thus the patterns made by absorption peaks at successive field strengths represent the behavior of the energy levels with increasing field. At zero field the levels group according to the principal quantum number n ; at intermediate field the levels display a roughly linear Stark effect, and at high fields they disappear owing to field ionization. The solid line is the classically predicted ionization field (see next section). Figure taken from Littman, Kash and Kleppner.

the nucleus. There its energy will increase with ϵ , but it will not spill over the lip of the potential and ionize.

6.6 Atoms in an Oscillating Electric Field

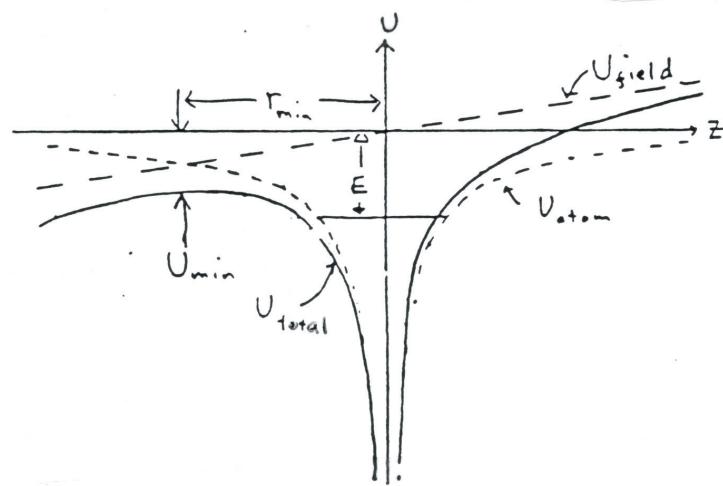


Figure 22. Potential diagram for field ionization.

There is a close connection between the behavior of an atom in a static electric field and its response to an oscillating field, i.e. a connection between the Stark effect and radiation processes. In the former case, the field induces a static dipole moment; in the latter case, it induces an oscillating moment. An oscillating moment creates an oscillating macroscopic polarization and leads to the absorption and emission of radiation. We shall calculate the response of an atom to an oscillating field

$$\mathcal{E}(\omega, t)\hat{\mathbf{e}} = \mathcal{E}\hat{\mathbf{e}} \cos \omega t \quad (6.53)$$

where $\hat{\mathbf{e}}$ is the polarization vector for the field. For a weak field the time varying state of this system can be found from first order time dependent perturbation theory. We shall write the electric dipole operator as $\mathbf{D} = -\mathbf{er}$. (This is a change of notation. Previously the symbol was \mathbf{d} .) The Hamiltonian naturally separates into two parts, $H = H_0 + H'(t)$, where H_0 is the unperturbed Hamiltonian and

$$H' = -\mathbf{D} \cdot \hat{\mathbf{e}} \mathcal{E} \cos \omega t = -\frac{1}{2}(e^{i\omega t} + e^{-i\omega t})\mathcal{E}\hat{\mathbf{e}} \cdot \mathbf{D} \quad (6.54)$$

We shall express the solution of the time dependent Schroedinger equation in terms of the eigenstates of H_0 , $|n\rangle$.

$$|\psi\rangle = \sum_n a_n e^{-i\omega_n t} |n\rangle \quad H_0 |\psi\rangle = \hbar \sum_n a_n \omega_n e^{-i\omega_n t} |n\rangle \quad (6.55)$$

where $\omega_n = E_n/\hbar$. Because of the perturbation $H'(t)$, the a_n 's become time dependent, and we have

$$i\hbar \frac{\partial \psi}{\partial t} = (H_0 + H') \sum_n a_n e^{-i\omega_n t} |n\rangle = \sum_n \hbar(a_n \omega_n + \dot{a}_n) |n\rangle e^{-i\omega_n t} \quad (6.56)$$

Left multiplying the final two expressions by $\langle k|$ to project out the k -th terms yields

$$\dot{a}_k = (i\hbar)^{-1} \sum_n \langle k | H'(t) | n \rangle a_n e^{i\omega_{kn} t} \quad (6.57)$$

where $\omega_{kn} = \omega_k - \omega_n$. In perturbation theory, this set of equations is solved by a set of approximations to a_k labeled $a_k^{(i)}(t)$. Starting with

$$a_n^{(0)}(t) = a_n(0) \quad (6.58)$$

one sets

$$\dot{a}_k^{(i+1)}(t) = (i\hbar)^{-1} \sum_n \langle k | H'(t) | n \rangle a_n^{(i)}(t) e^{i\omega_{kn} t} \quad (6.59)$$

and solves for the successive approximations by integration.

We now apply this to the problem of an atom which is in its ground state g at $t = 0$, and which is subject to the interaction of Eq. 6.54. Consequently $a_g(0) = 1$, $a_{n \neq g}(0) = 0$. Substituting in Eq. 6.59 and integrating from $t' = 0$ to t gives

$$\begin{aligned} a_k^{(1)}(t) &= (i\hbar)^{-1} \int_0^t dt' \langle k | H'(t') | g \rangle e^{i\omega_{kg} t'} \\ &= -(i\hbar)^{-1} \langle k | \hat{\mathbf{e}} \cdot \mathbf{D} | g \rangle \frac{\mathcal{E}}{2} \int_0^t dt' \left[e^{i(\omega_{kg} + \omega)t'} + e^{i(\omega_{kg} - \omega)t'} \right] \\ &= \frac{\mathcal{E}}{2\hbar} \langle k | \hat{\mathbf{e}} \cdot \mathbf{D} | g \rangle \left[\frac{e^{i(\omega_{kg} + \omega)t} - 1}{\omega_{kg} + \omega} + \frac{e^{i(\omega_{kg} - \omega)t} - 1}{\omega_{kg} - \omega} \right] \end{aligned} \quad (6.60)$$

The -1 terms in the square bracketed term arises because it is assumed that the field was turned on instantaneously at $t = 0$. They represent transients that rapidly damp and can be neglected.

The term with $\omega_{kg} + \omega$, in the denominator is the counter-rotating term. It can be neglected if one is considering cases where $\omega \approx \omega_{kg}$ (i.e. near resonance), but we shall retain both terms and calculate the expectation value of the first order time dependent dipole operator $\langle \mathbf{D}(\omega, t) \rangle$

$$\begin{aligned} \langle \mathbf{D}(\omega, t) \rangle &= 2\text{Re} \left\{ \langle g | \mathbf{D} | \sum_k a_k^{(1)}(t) e^{-i\omega_{kg} t} | k \rangle \right\} \\ &= \mathcal{E} \text{Re} \left[\sum_k \frac{\langle g | \mathbf{D} | k \rangle \langle k | \hat{\mathbf{e}} \cdot \mathbf{D} | g \rangle}{\hbar} \left\{ \frac{e^{i\omega t}}{\omega_{kg} + \omega} + \frac{e^{-i\omega t}}{\omega_{kg} - \omega} \right\} \right] \end{aligned} \quad (6.61)$$

If we consider the case of linearly polarized light ($\hat{\mathbf{e}} = \hat{\mathbf{z}}$), then

$$d_z(\omega, t) = \frac{2e^2}{\hbar} \sum_k \frac{\omega_{kg} |\langle k | z | g \rangle|^2}{\omega_{kg}^2 - \omega^2} \mathcal{E} \cos \omega t \quad (6.62)$$

We can write d_z in terms of a polarizability $\alpha(\omega)$:

$$\alpha(\omega) = \frac{2e^2}{\hbar} \sum_k \frac{\omega_{kg} |\langle k | z | g \rangle|^2}{\omega_{kg}^2 - \omega^2} \quad (6.63)$$

This result diverges if $\omega \rightarrow \omega_{kg}$. Later, when we introduce radiative damping, the divergence will be avoided in the usual way.

6.7 Oscillator Strength

Eq. 6.63 resemble the oscillating dipole moment of a system of classical oscillators. Consider a set of oscillators having charge q_k , mass m , and natural frequency ω_k , driven by the field $\mathcal{E} \cos \omega t$. The amplitude of the motion is given by

$$z_k = \frac{q_k}{m(\omega_k^2 - \omega^2)} \mathcal{E} \cos \omega t \quad (6.64)$$

If we have a set of such oscillators, then the total oscillating moment is given by

$$d_z(\omega, t) = \frac{1}{m} \sum_k \frac{q_k^2}{(\omega_k^2 - \omega^2)} \mathcal{E} \cos \omega t \quad (6.65)$$

This is strongly reminiscent of Eq. 6.62. It is useful to introduce the concept of oscillator strength, a dimensionless quantity defined as

$$f_{kj} = \frac{2m}{\hbar} \omega_{kj} |\langle k | z | j \rangle|^2 \quad (6.66)$$

where k and j are any two eigenstates. Note that f_{kj} is positive if $E_k > E_j$, i.e. for absorption, and negative if $E_k < E_j$. Then, Eq. 6.62 becomes

$$d_z(t) = \sum_k f_{kg} \frac{e^2}{m(\omega_{kg}^2 - \omega^2)} \mathcal{E} \cos \omega t \quad (6.67)$$

Comparing this with Eq. 6.65, we see that the behavior of an atom in an oscillating field mimics a set of classical oscillators with the same frequencies as the eigenfrequencies of the atom, but having effective charge strengths $q_k^2 = f_{kg} e^2$.

The oscillator strength is useful for characterizing radiative interactions and also the susceptibility of atoms. It satisfies an important sum rule, the Thomas-Reiche-Kuhn sum rule:

$$\sum_k f_{kg} = 1 \quad (6.68)$$

We prove by considering the general Hamiltonian

$$H = \frac{1}{2} \sum_j \mathbf{p}_j^2 + V(\mathbf{r}_1, \mathbf{r}_2, \dots). \quad (6.69)$$

Using the commutator relation

$$[A, B^2] = [A, B]B + B[A, B], \quad (6.70)$$

and the relation $[r_j, p_k] = i\hbar\delta_{jk}$, we have

$$[\mathbf{r}, H] = \frac{i\hbar}{m}\mathbf{p} \quad (6.71)$$

where $\mathbf{r} = \sum \mathbf{r}_j$, and $\mathbf{p} = \sum \mathbf{p}_j$. However,

$$\langle j | [\mathbf{r}, H] | k \rangle = (E_k - E_n) \langle j | \mathbf{r} | k \rangle \quad (6.72)$$

Consequently,

$$\langle j | \mathbf{r} | k \rangle = \frac{i}{m} \frac{\langle j | \mathbf{p} | k \rangle}{\omega_{kj}} \quad (6.73)$$

where $\omega_{kj} = (E_k - E_j)/\hbar$. Thus, we can write Eq. 6.66 in either of two forms:

$$f_{kj} = \frac{2i}{\hbar} \langle j | p_z | k \rangle \langle k | z | j \rangle = -\frac{2i}{\hbar} \langle k | p_z | j \rangle \langle j | z | k \rangle \quad (6.74)$$

Taking half the sum of these equations and using the closure relation $\sum_k |k\rangle\langle k| = 1$, we have

$$\sum_k f_{kj} = \frac{i}{\hbar} [\langle j | p_z z - z p_z | j \rangle] = 1 \quad (6.75)$$

We have calculated this for a one-electron atom, but the application to a Z-electron atom is straightforward because the Hamiltonian in Eq. 6.69 is quite general. In this case

$$\sum_k f_{kj} = Z. \quad (6.76)$$

Here j is some eigenstate of the system, and the index k describes all the eigenstates of all the electrons – including continuum states. In cases where only a single electron will be excited, however, for instance in the optical regime of a “single-electron” atom where the inner core electrons are essentially unaffected by the radiation, the atom behaves as if it were a single electron system with $Z = 1$.

Note that f_{kj} is positive if $\omega_{kj} > 0$, i.e. if the final state lies above the initial state. Such a transition corresponds to absorption of a photon. Since $f_{jk} = -f_{kj}$, the oscillator strength for emission of a photon is negative.

Our definition of oscillator strength, Eq. 6.66, singles out a particular axis, the $\hat{\mathbf{z}}$ -axis, fixed by the polarization of the light. Consequently, it depends on the orientation of the atom in the initial state and final states. It is convenient to introduce the average oscillator strength (often simply called the oscillator strength), by letting $|z_{kj}|^2 \rightarrow |r_{kj}|^2/3$, summing over the initial m state and averaging over the final state.

$$\overline{f_{kj}} = \frac{2}{3} \frac{m}{\hbar} \omega_{kj} \frac{1}{2J_j + 1} \sum_{m,m'} |\langle j, J_j, m' | \mathbf{r} | k, J_k, m \rangle|^2 \quad (6.77)$$

(This is the conversion followed by Sobelman.) It is evident that

$$\overline{f_{jk}} = -\frac{2J_j + 1}{2J_k + 1} \overline{f_{kj}} = -\frac{g_j}{g_k} \overline{f_{kj}}, \quad (6.78)$$

where g_j is the multiplicity factor for state j . An extensive discussion of the sum rules and their applications to oscillator strengths and transition momenta can be found in Bethe and Salpeter, section 6.1. Among the interesting features they point out is that transitions from an initial state $|n, \ell\rangle$ to a final state $|n', \ell'\rangle$ on the average have stronger oscillator strengths for absorption if $\ell' > \ell$, and stronger oscillator strengths for emission if $\ell' < \ell$. In other words, atoms “like” to increase their angular momentum on absorption of a photon, and decrease it on emission. The following page gives a table of oscillator strengths for hydrogen in which this tendency can be readily identified. (Taken from *The Quantum Mechanics of One- and Two-Electron Atoms*, H.A. Bethe and E.E. Salpeter, Academic Press (1957).)

Table 14. Oscillator strengths for hydrogen.

Initial	$1s$	$2s$	$2p$		$3s$	$3p$		$3d$	
Final	$n p$	$n p$	$n s$	$n d$	$n p$	$n s$	$n d$	$n p$	$n f$
$n = 1$	—	—	-0.139	—	—	-0.026	—	—	—
2	0.4162	—	—	—	-0.041	-0.145	—	-0.417	—
3	0.0791	0.4349	0.014	0.696	—	—	—	—	—
4	0.0290	0.1028	0.0031	0.122	0.484	0.032	0.619	0.011	1.016
5	0.0139	0.0419	0.0012	0.044	0.121	0.007	0.139	0.0022	0.156
6	0.0078	0.0216	0.0006	0.022	0.052	0.003	0.056	0.0009	0.053
7	0.0048	0.0127	0.0003	0.012	0.027	0.002	0.028	0.0004	0.025
8	0.0032	0.0081	0.0002	0.008	0.016	0.001	0.017	0.0002	0.015
$n = 9 \text{ to } \infty \text{ together}$	0.0109	0.0268	0.0007	0.023	0.048	0.002	0.045	0.0007	0.037
asymptotic	$1.6 n^{-3}$	$3.7 n^{-3}$	$0.1 n^{-3}$	$3.3 n^{-3}$	$6.2 n^{-3}$	$0.3 n^{-3}$	$6.1 n^{-3}$	$0.07 n^{-3}$	$4.4 n^{-3}$
Discrete spectrum	0.5650	0.6489	-0.119	0.928	0.707	-0.121	0.904	-0.402	1.302
Continuous spectrum	0.4350	0.3511	0.008	0.183	0.293	0.010	0.207	0.002	0.098
Total	1.000	1.000	-0.111	1.111	1.000	-0.111	1.111	-0.400	1.400
\bar{E}	0.54	0.61	0.6	0.42	0.78	0.47	—	0.39	—

Initial	$4s$	$4p$		$4d$		$4f$	
Final	$n p$	$n s$	$n d$	$n p$	$n f$	$n d$	$n g$
$n = 1$	—	-0.010	—	—	—	—	—
2	-0.009	-0.034	—	-0.073	—	—	—
3	-0.097	-0.161	-0.018	-0.371	—	-0.727	—
4	—	—	—	—	—	—	—
5	0.545	0.053	0.610	0.028	0.890	0.009	1.345
6	0.138	0.012	0.149	0.006	0.187	0.0016	0.183
7	0.060	0.006	0.063	0.002	0.072	0.0005	0.058
8	0.033	0.003	0.033	0.001	0.037	0.0003	0.027
$n = 9 \text{ to } \infty \text{ together}$	0.082	0.006	0.075	0.002	0.081	0.0006	0.045
asymptotic	$9.3 n^{-3}$	$0.7 n^{-3}$	$9.1 n^{-3}$	$0.3 n^{-3}$	$8.6 n^{-3}$	$0.05 n^{-3}$	$3.5 n^{-3}$
Discrete spectrum	0.752	-0.126	0.912	-0.406	1.267	-0.715	1.658
Continuous spectrum	0.248	0.015	0.199	0.006	0.133	0.001	0.056
Total	1.000	-0.111	1.111	-0.400	1.400	-0.714	1.714
\bar{E}	1.25	0.72	—	0.45	—	0.32	—

Figure 23. Oscillator strengths for hydrogen. From Bethe and Salpeter *The Quantum Mechanics of One- and Two-Electron Atoms*.

6.8 DC Stark Effect of a diatomic molecule

In the molecule's own frame, it has a definite dipole moment d . In the lab frame, however, it is not aligned and has $\langle d \rangle = 0$. In the absence of degeneracies, every eigenstate has definite parity.

Proof: Let P be the parity operator, H the Hamiltonian, and $|\Psi\rangle$ the eigenstate. If $[H, P] = 0$, then from $H|\Psi\rangle = E|\Psi\rangle$, we get $HP|\Psi\rangle = PH|\Psi\rangle = EP|\Psi\rangle$. Therefore, $P|\Psi\rangle$ is also an eigenstate of H with the same eigenenergy.

If there are no degeneracies, we can immediately conclude that $P|\Psi\rangle = \alpha|\Psi\rangle$ with the some number α . And since $P^2|\Psi\rangle = |\Psi\rangle$, we must have $\alpha^2 = 1 \Rightarrow \alpha = \pm 1$.

The dipole operator $\hat{\mathbf{d}} = d\hat{\mathbf{r}}$ has odd parity. Therefore, $\langle\Psi|\hat{\mathbf{d}}|\Psi\rangle = 0$. Formally,

$$\begin{aligned} P\hat{\mathbf{d}}P &= -\hat{\mathbf{d}} \\ \Rightarrow \langle\Psi|\hat{\mathbf{d}}|\Psi\rangle &= \langle\Psi|P\hat{\mathbf{d}}P|\Psi\rangle = -\langle\Psi|\hat{\mathbf{d}}|\Psi\rangle \\ \Rightarrow \langle\Psi|\hat{\mathbf{d}}|\Psi\rangle &= 0 \end{aligned}$$

Colloquially, no direction is singled out in the lab frame, so $\langle\hat{\mathbf{d}}\rangle = 0$.

Hamiltonian of a simple molecule, a "symmetric top" molecule: The Hamiltonian is given by

$$H = \frac{\mathbf{J}^2}{2\Theta},$$

where Θ is the moment of inertia (more generally a tensor), and \mathbf{J} is the angular momentum operator.

We can verify that $[H, P] = 0$ indeed.

The eigenstates of this Hamiltonian are spherical harmonics $Y_{JM}(\vartheta, \phi)$, and the eigenenergies are $E = \frac{\hbar^2 J(J+1)}{2\Theta} \equiv B J(J+1)$, where B is a rotational constant.

The spherical harmonics $Y_{JM}(\vartheta, \phi)$ have definite parity:

$$PY_{JM}(\vartheta, \phi) = Y_{JM}(\pi - \vartheta, \phi + \pi) = (-1)^J Y_{JM}(\vartheta, \phi).$$

And so

$$\begin{aligned} \langle JM|\hat{\mathbf{d}}|JM\rangle &= d \int d\Omega Y_{JM}^* \hat{\mathbf{r}} Y_{JM} = 0 \\ \text{e.g. } \langle JM|\hat{\mathbf{z}}|JM\rangle &= d \int d\Omega Y_{JM}^* \cos \vartheta Y_{JM} = 0 \end{aligned}$$

Indeed, as expected on general grounds, the molecule has no "permanent" electric dipole moment in the lab frame. The orientation of the molecule (described by ϑ and ϕ) is arbitrary, and any (ϑ, ϕ) pair is as likely as $(\pi - \vartheta, \pi + \phi)$. This is also clear when plotting a few $|Y_{JM}(\vartheta, \phi)|^2$ (see Fig. 24). For none of these you could say, "they point this way".

Applying an electric field Now let's apply an electric field in the (say) z-direction, $\mathbf{E} = E\hat{\mathbf{z}}$. The interaction is then given by $V = -\hat{\mathbf{d}} \cdot \mathbf{E} = -dE \cos \vartheta$.

We immediately know that there won't be a first-order shift: $\langle JM|V|JM\rangle = 0$ as $\langle JM|\cos \vartheta|JM\rangle = 0$.

As a side note: obviously, there are degeneracies here, as the eigenenergies for a given J are $(2J+1)$ -fold degenerate. However, all eigenstates $|JM\rangle$ for a

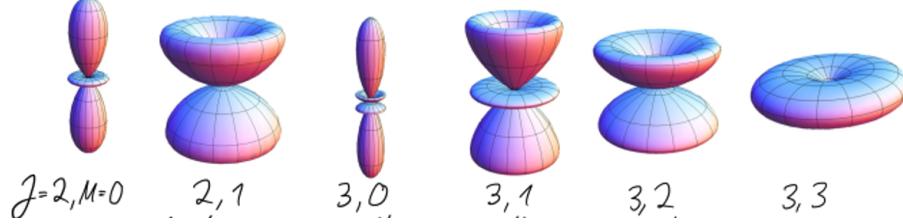


Figure 24. A few spherical harmonics.

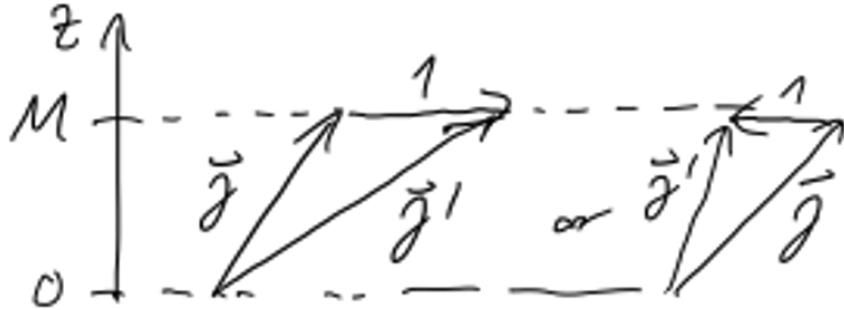


Figure 25. The triangle rule.

given J have the same parity $(-1)^J$, so even any superposition of the $|JM\rangle$ for fixed J are parity eigenstates.

In order to orient our molecule, the E-field will have to mix in states with different angular momentum J .

For second-order perturbation theory, we need matrix elements

$$\langle J'M' | \cos \vartheta | JM \rangle = \sqrt{\frac{4\pi}{3}} \int d\Omega Y_{J'M'}^*(\vartheta, \phi) Y_{10}(\vartheta, \phi) Y_{JM}(\vartheta, \phi).$$

Now this gives selection rules $M' = M$, $J' = J \pm 1$, also called the triangle rule, as shown in Fig. 25.

One has indeed

$$Y_{10} Y_{JM} = b_{JM} Y_{J+1,M} + b_{J-1,M} Y_{J-1,M}$$

$$\text{with } b_{JM} = \sqrt{\frac{(J+M+1)(J-M+1)}{(2J+1)(2J+3)}} \sqrt{\frac{3}{4\pi}}.$$

So $\langle J-1, M | \cos \vartheta | J, M \rangle = \sqrt{\frac{J^2 - M^2}{(2J-1)(2J+1)}}$ are the only non-zero matrix elements. Therefore, the state $|J, M\rangle$ is directly connected only to its "neighbors", $|J-1, M\rangle$ and $|J+1, M\rangle$. These two states are, in turn, mixed with their own neighboring states, $|J-2, M\rangle$ and $|J+2, M\rangle$. Therefore, we have in principle an infinite matrix. But we could truncate it.

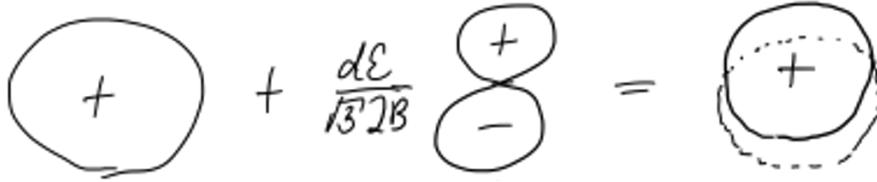


Figure 26. The new eigenstate has an orientation along the electric field \mathcal{E} .

Example of a dipole moment through mixing in different J states. The ground state $|0, 0\rangle$ is only connected to $|1, 0\rangle$. The matrix element is

$$\begin{aligned}\langle 0, 0 | \cos \vartheta | 1, 0 \rangle &= \sqrt{\frac{4\pi}{3}} \int d\Omega Y_{00}^* Y_{10}^2 \\ &= \frac{1}{\sqrt{3}} \int d\Omega Y_{10}^2 \\ &= \frac{1}{\sqrt{3}}\end{aligned}$$

The energy shift is

$$\begin{aligned}\Delta E_{00} &= \frac{|\langle 1, 0 | \cos \vartheta | 0, 0 \rangle|^2}{E_{0,0} - E_{1,0}} d^2 \mathcal{E}^2 \\ &= -\frac{1}{3} \frac{d^2 \mathcal{E}^2}{2B} \quad \text{where } E_{1,0} = B \cdot 1 \cdot 2 = 2B \\ &= -\frac{1}{6} \frac{d^2 \mathcal{E}^2}{B} \quad \text{must be negative: ground state can only be pushed down.}\end{aligned}$$

In this approximation, the new eigenstate is:

$$\begin{aligned}|\Psi\rangle &= |0, 0\rangle - d\mathcal{E} \frac{\langle 1, 0 | \cos \vartheta | 0, 0 \rangle}{E_{0,0} - E_{1,0}} |1, 0\rangle \\ &= |0, 0\rangle + \frac{d\mathcal{E}}{\sqrt{3} 2 B} |1, 0\rangle\end{aligned}$$

The molecule starts to gain an orientation along the electric field, as illustrated in Fig. 26.

$$\langle \Psi | \hat{d} | \Psi \rangle = d\hat{z} \langle \Psi | \cos \vartheta | \Psi \rangle \quad \text{only the z-component is non-zero}$$

$$\begin{aligned}&= d\hat{z} \left(-d\mathcal{E} \frac{\langle 1, 0 | \cos \vartheta | 0, 0 \rangle}{E_{0,0} - E_{1,0}} \langle 0, 0 | \cos \vartheta | 1, 0 \rangle - d\mathcal{E} \frac{\langle 0, 0 | \cos \vartheta | 1, 0 \rangle}{E_{0,0} - E_{1,0}} \langle 1, 0 | \cos \vartheta | 0, 0 \rangle \right) \\ &= 2d^2 \frac{|\langle 1, 0 | \cos \vartheta | 0, 0 \rangle|^2}{-E_{0,0} + E_{1,0}} \mathcal{E} \\ &= \frac{2}{3} \frac{d^2}{2B} \mathcal{E} = \frac{1}{3} \frac{d^2}{B} \mathcal{E}.\end{aligned}$$

This is a non-zero dipole moment, but only a fraction ($\frac{d\mathcal{E}}{3B}$) of the maximum

dipole moment. We have $\langle \mathbf{d} \rangle = \alpha \mathbf{\mathcal{E}}$, and α , the polarizability, is given by $\alpha = \frac{d^2}{3B}$.

We see that $\Delta E_{00} = -\frac{1}{6} \frac{d^2 \mathcal{E}^2}{B} = -\frac{1}{2} \langle \hat{d} \cdot \mathbf{\mathcal{E}} \rangle$, while the expectation value of the perturbation is $\langle V \rangle = -\langle \hat{d} \rangle \cdot \mathbf{\mathcal{E}}$.

Why is there a factor of 1/2?

An analogy with a spring stretched under gravity: In the steady state, we have $mg = k\Delta x$; The gravitational energy is $-mg\Delta x$. Energy stored in stretched spring is $\frac{1}{2}k\Delta_x^2 = \frac{1}{2}mg\Delta x$. The total energy is given by $E = -\frac{1}{2}mg\Delta x$.

So that's a property of "stretching" a harmonic oscillator. In linear response, the molecule is like a harmonic oscillator (indeed, I may only need the first excited state, so I can describe it by a two-level system or a weakly excited harmonic oscillator).

Let's check:

$$\begin{aligned} E_{\text{total}} &= \langle \Psi | H | \Psi \rangle \\ &= \langle \Psi | H_0 | \Psi \rangle + \langle \Psi | V | \Psi \rangle \\ &= \langle 10 | H_0 | 10 \rangle \cdot \frac{|\langle 10 | V | 00 \rangle|^2}{|E_{00} - E_{10}|^2} - \langle \mathbf{d} \rangle \cdot \mathbf{\mathcal{E}} \\ &= \frac{|\langle 10 | V | 00 \rangle|^2}{E_{10}} - \langle \mathbf{d} \rangle \cdot \mathbf{\mathcal{E}} \\ &= \frac{1}{2} \langle \mathbf{d} \rangle \cdot \mathbf{\mathcal{E}} - \langle \mathbf{d} \rangle \cdot \mathbf{\mathcal{E}} \\ &= -\frac{1}{2} \langle \mathbf{d} \rangle \cdot \mathbf{\mathcal{E}} \end{aligned}$$

We see that the extra energy (of the "spring") is stored in the excited state admixture of the molecule.

Chapter 7

Interaction of an Atom with an Electro-Magnetic Field

7.1 Introduction: Spontaneous and Stimulated Emission

Einstein's 1917 paper on the theory of radiation ⁹ provided seminal concepts for the quantum theory of radiation. It also anticipated devices such as the laser, and pointed the way to the field of laser-cooling of atoms. In it, he set out to answer two questions:

- 1) How do the internal states of an atom that radiates and absorbs energy come into equilibrium with a thermal radiation field? (In answering this question Einstein invented the concept of spontaneous emission)

⁹A. Einstein, Z. Phys. **18**, 121 (1917), translated in *Sources of Quantum Mechanics*, B. L. Van der Waerden, Cover Publication, Inc., New York, 1967. This book is a gold mine for anyone interested in the development of quantum mechanics.

2) How do the translational states of an atom in thermal equilibrium (i.e. states obeying the Maxwell-Boltzmann Law for the distribution of velocities) come into thermal equilibrium with a radiation field? (In answering this question, Einstein introduced the concept of photon recoil. He also demonstrated that the field itself must obey the Planck radiation law.)

The first part of Einstein's paper, which addresses question 1), is well known, but the second part, which addresses question 2), is every bit as germane for contemporary atom/optical physics. Because the paper preceded the creation of quantum mechanics there was no way for him to calculate transition rates. However, his arguments are based on general statistical principles and provide the foundation for interpreting the quantum mechanical results.

Einstein considered a system of N atoms in thermal equilibrium with a radiation field. The system has two levels¹⁰ with energies E_b and E_a , with $E_b > E_a$, and $E_b - E_a = \hbar\omega$. The numbers of atoms in the two levels are related by $N_b + N_a = N$. Einstein assumed the Planck radiation law for the spectral energy density temperature. For radiation in thermal equilibrium at temperature T , the energy per unit volume in wavelength range $d\omega$ is:

$$\rho_E(\omega)d\omega = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{\exp(\hbar\omega/kT) - 1} d\omega. \quad (7.1)$$

The mean occupation number of a harmonic oscillator at temperature T , which can be interpreted as the mean number of photons in one mode of the radiation field, is

$$\bar{n} = \frac{1}{\exp(\hbar\omega/kT) - 1}. \quad (7.2)$$

According to the Boltzmann Law of statistical mechanics, in thermal equilibrium the populations of the two levels are related by

$$\frac{N_b}{N_a} = \frac{g_b}{g_a} e^{-(E_b - E_a)/kT} = \frac{g_b}{g_a} e^{-\hbar\omega/kT}. \quad (7.3)$$

Here g_b and g_a are the multiplicities of the two levels. The last step assumes the Bohr frequency condition, $\omega = (E_b - E_a) \hbar$. However, Einstein's paper actually derives this relation independently.

According to classical theory, an oscillator can exchange energy with the radiation field at a rate that is proportional to the spectral density of radiation. The rates for absorption and emission are equal. The population transfer rate equation is thus predicted to be

$$\dot{N}_b = -\rho_E(\omega)B_{ba}N_b + \rho_E(\omega)B_{ab}N_a = -\dot{N}_a. \quad (7.4)$$

This equation is incompatible with Eq. 7.3. To overcome this problem, Einstein postulated that atoms in state b must spontaneously radiate to state a, with a constant radiation rate A_{ba} . Today such a process seems quite natural:

¹⁰An energy level consists of all of the states that have a given energy. The number of quantum states in a given level is its multiplicity.

the language of quantum mechanics is the language of probabilities and there is nothing jarring about asserting that the probability of radiating in a short time interval is proportional to the length of the interval. At that time such a random fundamental process could not be justified on physical principles. Einstein, in his characteristic Olympian style, brushed aside such concerns and merely asserted that the process is analogous to radioactive decay. With this addition, Eq. 7.4 becomes

$$\dot{N}_b = -[\rho_E(\omega)B_{ba} + A_{ba}]N_b + \rho_E(\omega)B_{ab}N_a = -\dot{N}_a. \quad (7.5)$$

It follows that

$$\begin{aligned} g_b B_{ba} &= g_a B_{ab} \\ \frac{\hbar\omega^3}{\pi^2 c^3} B_{ba} &= A_{ba} \\ \rho_E(\omega)B_{ba} &= \bar{n}A_{ba} \end{aligned} \quad (7.6)$$

Consequently, the rate of transition $b \rightarrow a$ is

$$B_{ba}\rho_E(\omega) + A_{ba} = (\bar{n} + 1)A_{ba}, \quad (7.7)$$

while the rate of absorption is

$$B_{ab}\rho_E(\omega) = \frac{g_b}{g_a}\bar{n}A_{ba} \quad (7.8)$$

If we consider emission and absorption between single states by taking $g_b = g_a = 1$, then the ratio of rate of emission to rate of absorption is $(\bar{n} + 1)/\bar{n}$.

This argument reveals the fundamental role of spontaneous emission. Without it, atomic systems could not achieve thermal equilibrium with a radiation field. Thermal equilibrium requires some form of dissipation, and dissipation is equivalent to having an irreversible process. Spontaneous emission is the fundamental irreversible process in nature. The reason that it is irreversible is that once a photon is radiated into the vacuum, the probability that it will ever be reabsorbed is zero: there are an infinity of vacuum modes available for emission but only one mode for absorption. If the vacuum modes are limited, for instance by cavity effects, the number of modes becomes finite and equilibrium is never truly achieved. In the limit of only a single mode, the motion becomes reversible.

The identification of the Einstein A coefficient with the rate of spontaneous emission is so well established that we shall henceforth use the symbol A_{ba} to denote the spontaneous decay rate from state b to a . The radiative lifetime for such a transition is $\tau_{ba} = A_{ba}^{-1}$.

Here, Einstein came to a halt. Lacking quantum theory, there was no way to calculate A_{ba} .

7.2 Quantum Theory of Absorption and Emission

We shall start by describing the behavior of an atom in a classical electromagnetic field. Although treating the field classically while treating the atom quantum mechanically is fundamentally inconsistent, it provides a natural and intuitive approach to the problem. Furthermore, it is completely justified in cases where the radiation fields are large, in the sense that there are many photons in each mode, as for instance, in the case of microwave or laser spectroscopy. There is, however, one important process that this approach cannot deal with satisfactorily. This is spontaneous emission, which we shall treat later using a quantized field. Nevertheless, phenomenological properties such as selection rules, radiation rates and cross sections, can be developed naturally with this approach.

7.3 The classical E-M field

Our starting point is Maxwell's equations (S.I. units):

$$\begin{aligned}\nabla \cdot \mathbf{E} &= \rho/\epsilon_0 \\ \nabla \cdot \mathbf{B} &= 0 \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \\ \nabla \times \mathbf{B} &= \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{J}\end{aligned}\tag{7.9}$$

The charge density ρ and current density \mathbf{J} obey the continuity equation

$$\nabla \cdot \mathbf{J} + \frac{\partial \rho}{\partial t} = 0\tag{7.10}$$

Introducing the vector potential \mathbf{A} and the scalar potential ψ , we have

$$\begin{aligned}\mathbf{E} &= -\nabla\psi - \frac{\partial \mathbf{A}}{\partial t} \\ \mathbf{B} &= \nabla \times \mathbf{A}\end{aligned}\tag{7.11}$$

We are free to change the potentials by a gauge transformation:

$$\mathbf{A}' = \mathbf{A} + \nabla\Lambda, \quad \psi' = \psi - \frac{\partial\Lambda}{\partial t}\tag{7.12}$$

where Λ is a scalar function. This transformation leaves the fields invariant, but changes the form of the dynamical equation. We shall work in the *Coulomb gauge* (often called the radiation gauge), defined by

$$\nabla \cdot \mathbf{A} = 0\tag{7.13}$$

In free space, \mathbf{A} obeys the wave equation

$$\nabla^2 \mathbf{A} = \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2}\tag{7.14}$$

Because $\nabla \cdot \mathbf{A} = 0$, \mathbf{A} is transverse. We take a propagating plane wave solution of the form

$$\mathbf{A}(r, t) = A\hat{\mathbf{e}} \cos(\mathbf{k} \cdot \mathbf{r} - \omega t) = A\hat{\mathbf{e}} \frac{1}{2} [e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} + e^{-i(\mathbf{k} \cdot \mathbf{r} - \omega t)}], \quad (7.15)$$

where $k^2 = \omega^2/c^2$ and $\hat{\mathbf{e}} \cdot \mathbf{k} = 0$. For a linearly polarized field, the polarization vector $\hat{\mathbf{e}}$ is real. For an elliptically polarized field it is complex, and for a circularly polarized field it is given by $\hat{\mathbf{e}} = (\hat{\mathbf{x}} \pm i\hat{\mathbf{y}})/\sqrt{2}$, where the + and – signs correspond to positive and negative helicity, respectively. (Alternatively, they correspond to left and right hand circular polarization, respectively, the sign convention being a tradition from optics.) The electric and magnetic fields are then given by

$$\mathbf{E}(r, t) = \omega A \hat{\mathbf{e}} \sin(\mathbf{k} \cdot \mathbf{r} - \omega t) = -i\omega A \hat{\mathbf{e}} \frac{1}{2} [e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} - e^{-i(\mathbf{k} \cdot \mathbf{r} - \omega t)}]. \quad (7.16)$$

$$\mathbf{B}(r, t) = k(\hat{\mathbf{k}} \times \hat{\mathbf{e}}) \sin(\mathbf{k} \cdot \mathbf{r} - \omega t) = -ik A (\hat{\mathbf{k}} \times \hat{\mathbf{e}}) \frac{1}{2} [e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} - e^{-i(\mathbf{k} \cdot \mathbf{r} - \omega t)}]. \quad (7.17)$$

The time average Poynting vector is

$$\mathbf{S} = \frac{\epsilon_0 c^2}{2} (\mathbf{E} \times \mathbf{B}^*) = \frac{\epsilon_0 c}{2} \omega^2 A^2 \hat{\mathbf{k}}. \quad (7.18)$$

The average energy density in the wave is given by

$$u = \omega^2 \frac{\epsilon_0}{2} A^2 \hat{\mathbf{k}}. \quad (7.19)$$

7.4 Interaction of an electromagnetic wave and an atom

The behavior of charged particles in an electromagnetic field is correctly described by Hamilton's equations provided that the canonical momentum is re-defined:

$$\mathbf{p}_{\text{can}} = \mathbf{p}_{\text{kin}} + q\mathbf{A} \quad (7.20)$$

The kinetic energy is $\mathbf{p}_{\text{kin}}^2/2m$. Taking $q = -e$, the Hamiltonian for an atom in an electromagnetic field in free space is

$$H = \frac{1}{2m} \sum_{j=1}^N (\mathbf{p}_j + e\mathbf{A}(r_j))^2 + \sum_{j=1}^N V(\mathbf{r}_j), \quad (7.21)$$

where $V(\mathbf{r}_j)$ describes the potential energy due to internal interactions. We are neglecting spin interactions.

Expanding and rearranging, we have

$$\begin{aligned} H &= \sum_{j=1}^N \frac{\mathbf{p}_j^2}{2m} + V(\mathbf{r}_j) + \frac{e}{2m} \sum_{j=1}^N (\mathbf{p}_j \cdot \mathbf{A}(\mathbf{r}_j) + \mathbf{A}(\mathbf{r}_j) \cdot \mathbf{p}_j) + \frac{e^2}{2m} \sum_{j=1}^N A_j^2(\mathbf{r}) \\ &= H_0 + H_{\text{int}} + H^{(2)}. \end{aligned} \quad (7.22)$$

Here, $\mathbf{p}_j = -i\hbar\nabla_j$. Consequently, H_0 describes the unperturbed atom. H_{int} describes the atom's interaction with the field. $H^{(2)}$, which is second order in \mathbf{A} , plays a role only at very high intensities. (In a static magnetic field, however, $H^{(2)}$ gives rise to diamagnetism.)

Because we are working in the Coulomb gauge, $\nabla \cdot \mathbf{A} = 0$ so that \mathbf{A} and \mathbf{p} commute. We have

$$H_{\text{int}} = \frac{eA}{mc} \hat{\mathbf{e}} \cdot \mathbf{p} \cos(\mathbf{k} \cdot \mathbf{r} - \omega t). \quad (7.23)$$

It is convenient to write the matrix element between states $|a\rangle$ and $|b\rangle$ in the form

$$\langle b|H_{\text{int}}|a\rangle = \frac{1}{2}H_{ba}e^{-i\omega t} + \frac{1}{2}H_{ba}e^{+i\omega t}, \quad (7.24)$$

where

$$H_{ba} = \frac{eA}{m} \hat{\mathbf{e}} \langle b|\mathbf{p}|e^{i\mathbf{k}\cdot\mathbf{r}}|a\rangle. \quad (7.25)$$

Atomic dimensions are small compared to the wavelength of radiation involved in optical transitions. The scale of the ratio is set by $\alpha \approx 1/137$. Consequently, when the matrix element in Eq. 7.25 is evaluated, the wave function vanishes except in the region where $\mathbf{k} \cdot \mathbf{r} = 2\pi r/\lambda \ll 1$. It is therefore appropriate to expand the exponential:

$$H_{ba} = \frac{eA}{mc} \hat{\mathbf{e}} \cdot \langle b|\mathbf{p}(1 + i\mathbf{k} \cdot \mathbf{r} - 1/2(\mathbf{k} \cdot \mathbf{r})^2 + \dots)|a\rangle \quad (7.26)$$

Unless $\langle b|\mathbf{p}|a\rangle$ vanishes, for instance due to parity considerations, the leading term dominates and we can neglect the others. For reasons that will become clear, this is called the dipole approximation. This is by far the most important situation, and we shall defer consideration of the higher order terms. In the dipole approximation we have

$$H_{ba} = \frac{eA}{m} \hat{\mathbf{e}} \cdot \langle b|\mathbf{p}|a\rangle = \frac{-ieE}{m\omega} \hat{\mathbf{e}} \cdot \langle b|\mathbf{p}|a\rangle \quad (7.27)$$

where we have used, from Eq. 7.16, $A = -iE/\omega$. It can be shown (i.e. left as exercise) that the matrix element of \mathbf{p} can be transformed into a matrix element for \mathbf{r} :

$$\langle b|\mathbf{p}|a\rangle = -im\omega_{ab} \langle b|\mathbf{r}|a\rangle = +im\omega_{ba} \langle b|\mathbf{r}|a\rangle \quad (7.28)$$

This results in

$$H_{ba} = \frac{eE\omega_{ba}}{\omega} \hat{\mathbf{e}} \cdot \langle b|\mathbf{r}|a \rangle \quad (7.29)$$

We will be interested in resonance phenomena in which $\omega \approx \omega_{ba}$. Consequently,

$$H_{ba} = +e\mathbf{E}_0 \cdot \langle b|\mathbf{r}|a \rangle = -\mathbf{d}_{ba} \cdot \mathbf{E} \quad (7.30)$$

where \mathbf{d} is the dipole operator, $\mathbf{d} = -e\mathbf{r}$. Displaying the time dependence explicitly, we have

$$H'_{ba} = -\mathbf{d}_{ba} \cdot \mathbf{E}_0 e^{-i\omega t}. \quad (7.31)$$

However, it is important to bear in mind that this is only the first term in a series, and that if it vanishes the higher order terms will contribute a perturbation at the driving frequency.

H_{ba} appears as a matrix element of the momentum operator \mathbf{p} in Eq. 7.27, and of the dipole operator \mathbf{r} in Eq. 7.30. These matrix elements look different and depend on different parts of the wave function. The momentum operator emphasizes the curvature of the wave function, which is largest at small distances, whereas the dipole operator evaluates the moment of the charge distribution, i.e. the long range behavior. In practice, the accuracy of a calculation can depend significantly on which operator is used.

7.5 Quantization of the radiation field

We shall consider a single mode of the radiation field. This means a single value of the wave vector \mathbf{k} , and one of the two orthogonal transverse polarization vectors $\hat{\mathbf{e}}$. The radiation field is described by a plane wave vector potential of the form Eq. 7.15. We assume that \mathbf{k} obeys a periodic boundary or condition, $k_x L_x = 2\pi n_x$, etc. (For any \mathbf{k} , we can choose boundaries L_x, L_y, L_z to satisfy this.) The time averaged energy density is given by Eq. 7.19, and the total energy in the volume V defined by these boundaries is

$$U = \frac{\epsilon_0}{2} \omega^2 A^2 V, \quad (7.32)$$

where A^2 is the mean squared value of A averaged over the spatial mode. We now make a formal connection between the radiation field and a harmonic oscillator. We define variables Q and P by

$$A = \frac{1}{\omega} \sqrt{\frac{1}{\epsilon_0 V}} (\omega Q + iP), \quad A^* = \frac{1}{\omega} \sqrt{\frac{1}{\epsilon_0 V}} (\omega Q - iP). \quad (7.33)$$

Then, from Eq. 7.32, we find

$$U = \frac{1}{2} (\omega^2 Q^2 + P^2). \quad (7.34)$$

This describes the energy of a harmonic oscillator having unit mass. We quantize the oscillator in the usual fashion by treating Q and P as operators,

with

$$P = -i\hbar \frac{\partial}{\partial Q}, \quad [Q, P] = i\hbar. \quad (7.35)$$

We introduce the operators a and a^\dagger defined by

$$a = \frac{1}{\sqrt{2\hbar\omega}}(\omega Q + iP) \quad (7.36)$$

$$a^\dagger = \frac{1}{\sqrt{2\hbar\omega}}(\omega Q - iP) \quad (7.37)$$

The fundamental commutation rule is

$$[a, a^\dagger] = 1 \quad (7.38)$$

from which the following can be deduced:

$$H = \frac{1}{2}\hbar\omega[a^\dagger a + aa^\dagger] = \hbar\omega\left[a^\dagger a + \frac{1}{2}\right] = \hbar\omega\left[N + \frac{1}{2}\right] \quad (7.39)$$

where the number operator $N = a^\dagger a$ obeys

$$N|n\rangle = n|n\rangle \quad (7.40)$$

We also have

$$\begin{aligned} \langle n-1|a|n\rangle &= \sqrt{n} \\ \langle n+1|a^\dagger|n\rangle &= \sqrt{n+1} \\ \langle n|a^\dagger a|n\rangle &= n \\ \langle n|aa^\dagger|n\rangle &= n+1 \\ \langle n|H|n\rangle &= \hbar\omega\left(n + \frac{1}{2}\right) \\ \langle n|a|n\rangle &= \langle n|a^\dagger|n\rangle = 0 \end{aligned} \quad (7.41)$$

The operators a and a^\dagger are called the annihilation and creation operators, respectively. We can express the vector potential and electric field in terms of a and a^\dagger as follows

$$A = \frac{1}{\omega\sqrt{\epsilon_o V}}(\omega Q + iP) = \sqrt{\frac{2\hbar}{\omega\epsilon_o V}}a \quad (7.42)$$

$$A^\dagger = \frac{1}{\omega\sqrt{\epsilon_o V}}(\omega Q - iP) = \sqrt{\frac{2\hbar}{\omega\epsilon_o V}}a^\dagger$$

$$\mathbf{E} = -i\sqrt{\frac{\hbar\omega}{2\epsilon_o V}}[a\hat{\mathbf{e}}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} - a^\dagger\hat{\mathbf{e}}^*e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)}] \quad (7.43)$$

In the dipole limit we can take $e^{i\mathbf{k}\cdot\mathbf{r}} = 1$. Then

$$\mathbf{E} = -i\sqrt{\frac{\hbar\omega}{2\epsilon_0 V}} [a\hat{\mathbf{e}}e^{-i\omega t} - a^\dagger\hat{\mathbf{e}}^*e^{+i\omega t}] \quad (7.44)$$

The interaction Hamiltonian is,

$$H_{\text{int}} = -ie\sqrt{\frac{\hbar\omega}{2\epsilon_0 V}} \mathbf{r} \cdot [a\hat{\mathbf{e}}e^{-i\omega t} - a^\dagger\hat{\mathbf{e}}^*e^{+i\omega t}], \quad (7.45)$$

where we have written the dipole operator as $\mathbf{d} = -e\mathbf{r}$.

7.6 Interaction of a two-level system and a single mode of the radiation field

We consider a two-state atomic system $|a\rangle, |b\rangle$ and a radiation field described by $|n\rangle, n = 0, 1, 2 \dots$. The states of the total system can be taken to be

$$|I\rangle = |a, n\rangle = |a\rangle |n\rangle, \quad |F\rangle = |b, n'\rangle = |b\rangle |n'\rangle. \quad (7.46)$$

We shall take $\hat{\mathbf{e}} = \hat{\mathbf{z}}$. Then

$$\langle F|H_{\text{int}}|I\rangle = ie z_{ab} \sqrt{\frac{\hbar\omega}{2\epsilon_0 V}} \langle n'|ae^{-i\omega t} - a^\dagger e^{i\omega t}|n\rangle e^{-i\omega_{ab}t} \quad (7.47)$$

The first term in the bracket obeys the selection rule $n' = n-1$. This corresponds to loss of one photon from the field and absorption of one photon by the atom. The second term obeys $n' = n+1$. This corresponds to emission of a photon by the atom. Using Eq. 7.41, we have

$$\langle F|H_{\text{int}}|I\rangle = -ie z_{ab} \sqrt{\frac{\hbar\omega}{2\epsilon_0 V}} (\sqrt{n} \delta_{n, n-1} e^{-i\omega t} - \sqrt{n+1} \delta_{n, n+1} e^{+i\omega t}) e^{-i\omega_{ab}t} \quad (7.48)$$

Transitions occur when the total time dependence is zero, or near zero. Thus absorption occurs when $\omega = -\omega_{ab}$, or $E_a + \hbar\omega = E_b$. As we expect, energy is conserved. Similarly, emission occurs when $\omega = +\omega_{ab}$, or $E_a - \hbar\omega = E_b$.

A particularly interesting case occurs when $n = 0$, i.e. the field is initially in the vacuum state, and $\omega = \omega_{ab}$. Then

$$\langle F|H_{\text{int}}|I\rangle = ie z_{ab} \sqrt{\frac{\hbar\omega}{2\epsilon_0 V}} \equiv H_{FI}^0 \quad (7.49)$$

The situation describes a constant perturbation H_{FI}^0 coupling the two states $I = |a, n=0\rangle$ and $F = |b, n'=1\rangle$. The states are degenerate because $E_a = E_b + \hbar\omega$. Consequently, E_a is the upper of the two atomic energy levels.

The system is composed of two degenerate eigenstates, but due to the coupling of the field, the degeneracy is split. The eigenstates are symmetric and antisymmetric combinations of the initial states, and we can label them as

$$|\pm\rangle = \frac{1}{\sqrt{2}}(|I\rangle \pm |F\rangle) = \frac{1}{\sqrt{2}}(|a, 0\rangle \pm |b, 1\rangle). \quad (7.50)$$

1227.7 Interaction of a two-level system and a single mode of the radiation field

The energies of these states are

$$E_{\pm} = \pm|H_{FI}^0| \quad (7.51)$$

If at $t = 0$, the atom is in state $|a\rangle$, which means that the radiation field is in state $|0\rangle$, then the system is in a superposition state:

$$\psi(0) = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle). \quad (7.52)$$

The time evolution of this superposition is given by

$$\psi(t) = \frac{1}{\sqrt{2}} \left(|+\rangle e^{i\Omega/2t} + |-\rangle e^{-i\Omega/2t} \right) \quad (7.53)$$

where $\Omega/2 = |H_{FI}^0|/\hbar = ez_{ab}\sqrt{\omega/(2\epsilon_o V \hbar)}$. The probability that the atom is in state $|b\rangle$ at a later time is

$$P_b = \frac{1}{2}(1 + \cos \Omega t). \quad (7.54)$$

The frequency Ω is called the vacuum Rabi frequency.

The dynamics of a 2-level atom interacting with a single mode of the vacuum were first analyzed in Ref. [1] and the oscillations are sometimes called *Jaynes-Cummings* oscillations.

The atom-vacuum interaction H_{FI}^0 , Eq. 7.49, has a simple physical interpretation. The electric field amplitude associated with the zero point energy in the cavity is given by

$$\epsilon_o E^2 V = \frac{1}{2} \hbar \omega \quad (7.55)$$

Consequently, $|H_{FI}^0| = Ed_{ab} = ez_{ab}E$. The interaction frequency $|H_{FI}^0|/\hbar$ is sometimes referred to as the vacuum Rabi frequency, although, as we have seen, the actual oscillation frequency is $2 \times H_{FI}^0/\hbar$.

Absorption and emission are closely related. Because the rates are proportional to $|\langle F | H_{\text{int}} | I \rangle|^2$, it is evident from Eq. 7.48 that

$$\frac{\text{Rate of emission}}{\text{Rate of absorption}} = \frac{n+1}{n} \quad (7.56)$$

This result, which applies to radiative transitions between any two states of a system, is general. In the absence of spontaneous emission, the absorption and emission rates are identical.

The oscillatory behavior described by Eq. 7.53 is exactly the opposite of free space behavior in which an excited atom irreversibly decays to the lowest available state by spontaneous emission. The distinction is that in free space there are an infinite number of final states available to the photon, since it can go off in any direction, but in the cavity there is only one state. The natural way to regard the atom-cavity system is not in terms of the atom and cavity separately, as in Eq. 7.46, but in terms of the coupled states $|+\rangle$ and $|-\rangle$ (Eq.

7.50). Such states, called *dressed atom* states, are the true eigenstates of the atom-cavity system.

7.7 Absorption and emission

In Chapter 6, first-order perturbation theory was applied to find the response of a system initially in state $|a\rangle$ to a perturbation of the form $H_{ba}e^{-i\omega t}$. The result is that the amplitude for state $|b\rangle$ is given by

$$a_b(t) = \frac{1}{i\hbar} \int_0^t H_{ba} e^{-i(\omega - \omega_{ba})t'} dt' = \frac{H_{ba}}{\hbar} \left[\frac{e^{-i(\omega - \omega_{ba})t} - 1}{\omega - \omega_{ba}} \right] \quad (7.57)$$

There will be a similar expression involving the time-dependence $e^{+i\omega t}$. The $-i\omega$ term gives rise to resonance at $\omega = \omega_{ba}$; the $+i\omega$ term gives rise to resonance at $\omega = \omega_{ab}$. One term is responsible for absorption, the other is responsible for emission.

The probability that the system has made a transition to state $|b\rangle$ at time t is

$$W_{a \rightarrow b} = |a_b(t)|^2 = \frac{|H_{ba}|^2}{\hbar^2} \frac{\sin^2[(\omega - \omega_{ba})t/2]}{((\omega - \omega_{ba})/2)^2} \quad (7.58)$$

In the limit $\omega \rightarrow \omega_{ba}$, we have

$$W_{a \rightarrow b} \approx \frac{|H_{ba}|^2}{\hbar^2} t^2. \quad (7.59)$$

So, for short time, $W_{a \rightarrow b}$ increases quadratically. This is reminiscent of a Rabi resonance in a 2-level system in the limit of short time.

However, Eq. 7.58 is only valid provided $W_{a \rightarrow b} \ll 1$, or for time $T \ll \hbar/H_{ba}$. For such a short time, the incident radiation will have a spectral width $\Delta\omega \sim 1/T$. In this case, we must integrate Eq. 7.58 over the spectrum. In doing this, we shall make use of the relation

$$\int_{-\infty}^{+\infty} \frac{\sin^2(\omega - \omega_{ba})t/2}{[(\omega - \omega_{ba})/2]^2} d\omega = 2t \int_{-\infty}^{+\infty} \frac{\sin^2(u - u_o)}{(u - u_o)^2} du \rightarrow 2\pi t \int_{-\infty}^{+\infty} \delta(\omega - \omega_{ba}) d\omega. \quad (7.60)$$

Eq. 7.58 becomes

$$W_{a \rightarrow b} = \frac{|H_{ba}|^2}{\hbar^2} 2\pi t \delta(\omega - \omega_{ba}) \quad (7.61)$$

The δ -function requires that eventually $W_{a \rightarrow b}$ be integrated over a spectral distribution function. $W_{a \rightarrow b}$ can also be written

$$W_{a \rightarrow b} = \frac{|H_{ba}|^2}{\hbar} 2\pi t \delta(E_b - E_a - \hbar\omega). \quad (7.62)$$

Because the transition probability is proportional to the time, we can define the transition rate

$$\begin{aligned}\Gamma_{ab} &= \frac{d}{dt}W_{a \rightarrow b} = 2\pi \frac{|H_{ba}|^2}{\hbar^2} \delta(\omega - \omega_{ba}) \\ &= 2\pi \frac{|H_{ba}|^2}{\hbar} \delta(E_b - E_a - \hbar\omega)\end{aligned}\quad (7.63)$$

The δ -function arises because of the assumption in first order perturbation theory that the amplitude of the initial state is not affected significantly. This will not be the case, for instance, if a monochromatic radiation field couples the two states, in which case the amplitudes oscillate between 0 and 1. However, the assumption of perfectly monochromatic radiation is in itself unrealistic.

Radiation always has some spectral width. $|H_{ba}|^2$ is proportional to the intensity of the radiation field at resonance. The intensity can be written in terms of a spectral density function

$$S(\omega') = S_0 f(\omega')$$

where S_0 is the incident Poynting vector, and $f(\omega')$ is a normalized line shape function centered at the frequency ω' which obeys $\int f(\omega') d\omega' = 1$. We can define a characteristic spectral width of $f(\omega')$ by

$$\Delta\omega = \frac{1}{f(\omega_{ab})} \quad (7.64)$$

Integrating Eq. 7.63 over the spectrum of the radiation gives

$$\Gamma_{ab} = \frac{2\pi|H_{ba}|^2}{\hbar^2} f(\omega_{ab}) \quad (7.65)$$

If we define the effective Rabi frequency by

$$\Omega_R = \frac{2|H_{ba}|}{\hbar} \quad (7.66)$$

then

$$\Gamma_{ab} = \frac{\pi}{2} \frac{\Omega_R^2}{\Delta\omega} \quad (7.67)$$

Another situation that often occurs is when the radiation is monochromatic, but the final state is actually composed of many states spaced close to each other in energy so as to form a continuum. If such is the case, the density of final states can be described by

$$dN = \rho(E) dE \quad (7.68)$$

where dN is the number of states in range dE . Taking $\hbar\omega = E_b - E_a$ in Eq. 7.63, and integrating gives

$$\Gamma_{ab} = 2\pi \frac{|H_{ba}|^2}{\hbar} \rho(E_b) \quad (7.69)$$

This result remains valid in the limit $E_b \rightarrow E_a$, where $\omega \rightarrow 0$. In this static situation, the result is known as *Fermi's Golden Rule*.

Note that Eq. 7.65 and Eq. 7.69 both describe a uniform rate process in which the population of the initial state decreases exponentially in time. If the population of the initial state is $P(0)$, then

$$P(t) = P(0)e^{-\Gamma_{ba}t} \quad (7.70)$$

Applying this to the dipole transition described in Eq. 7.30, we have

$$\Gamma_{ab} = 2\pi \frac{E^2 d_{ba}^2}{\hbar^2} f(\omega) \quad (7.71)$$

The arguments here do not distinguish whether $E_a < E_b$ or $E_a > E_b$ (though the sign of $\omega = (E_b - E_a)/\hbar$ obviously does). In the former case the process is absorption, in the latter case it is emission.

7.8 Spontaneous emission rate

The rate of absorption for the transition $a \rightarrow b$, where $E_b > E_a$, is, from Eq. 7.45 and Eq. 7.63,

$$\Gamma_{ab} = \frac{4\pi^2}{\hbar V} |\hat{\mathbf{e}} \cdot \mathbf{d}_{ba}|^2 n \omega \delta(\omega_0 - \omega). \quad (7.72)$$

where $\omega_0 = (E_b - E_a)/\hbar$. To evaluate this we need to let $n \rightarrow n(\omega)$, where $n(\omega)d\omega$ is the number of photons in the frequency interval $d\omega$, and integrate over the spectrum. The result is

$$\Gamma_{ab} = \frac{4\pi^2}{\hbar V} |\hat{\mathbf{e}} \cdot \mathbf{d}_{ba}|^2 \omega_0 n(\omega_0) \quad (7.73)$$

To calculate $n(\omega)$, we first calculate the mode density in space by applying the usual periodic boundary condition

$$k_j L = 2\pi n_j, \quad j = x, y, z. \quad (7.74)$$

The number of modes in the range $d^3 k = dk_x dk_y dk_z$ is

$$dN = dn_x dn_y dn_z = \frac{V}{(2\pi)^3} d^3 k = \frac{V}{(2\pi)^3} k^2 dk d\Omega = \frac{V}{(2\pi)^3} \frac{\omega^2}{c^3} \frac{d\omega}{\hbar} d\Omega \quad (7.75)$$

Letting $\bar{n} = \bar{n}(\omega)$ be the average number of photons per mode, then

$$n(\omega) = \bar{n} \frac{dN}{d\omega} = \frac{\bar{n} V \omega^2 d\Omega}{(2\pi)^3 c^3} \quad (7.76)$$

Introducing this into Eq. 7.73 gives

$$\Gamma_{ab} = \frac{\bar{n} \omega^3}{2\pi \hbar c^3} |\hat{\mathbf{e}} \cdot \mathbf{d}_{ba}|^2 d\Omega \quad (7.77)$$

We wish to apply this to the case of isotropic radiation in free space, as, for instance, in a thermal radiation field. We can take \mathbf{d}_{ba} to lie along the z axis and

describe \mathbf{k} in spherical coordinates about this axis. Since the wave is transverse, $\hat{\mathbf{e}} \cdot \hat{\mathbf{D}} = \sin \theta$. However, there are 2 orthogonal polarizations. Consequently,

$$\int |\hat{\mathbf{e}} \cdot \mathbf{d}_{ba}|^2 d\Omega = 2|\mathbf{d}_{ba}|^2 \int \sin^2 \theta d\Omega = \frac{8\pi}{3} |\mathbf{d}_{ba}|^2 \quad (7.78)$$

Introducing this into Eq. 7.77 yields the absorption rates

$$\Gamma_{ab} = \frac{4}{3} \frac{\omega^3}{\hbar c^3} |\mathbf{d}_{ba}|^2 \bar{n} \quad (7.79)$$

It follows that the emission rate for the transition $b \rightarrow a$ is

$$\Gamma_{ba} = \frac{4}{3} \frac{\omega^3}{\hbar c^3} |\mathbf{d}_{ba}|^2 (\bar{n} + 1) \quad (7.80)$$

If there are no photons present, the emission rate—called the rate of spontaneous emission—is

$$\Gamma_{ba}^0 = \frac{4}{3} \frac{\omega^3}{\hbar c^3} |\mathbf{d}_{ba}|^2 = \frac{4}{3} \frac{e^2 \omega^3}{\hbar c^3} |\langle b | \mathbf{r} | a \rangle|^2 \quad (7.81)$$

In atomic units, in which $c = 1/\alpha$, we have

$$\Gamma_{ba}^0 = \frac{4}{3} \alpha^3 \omega^3 |\mathbf{r}_{ba}|^2. \quad (7.82)$$

Taking, typically, $\omega = 1$, and $r_{ba} = 1$, we have $\Gamma^0 \approx \alpha^3$. The “ Q ” of a radiative transition is $Q = \omega/\Gamma \approx \alpha^{-3} \approx 3 \times 10^6$. The α^3 dependence of Γ indicates that radiation is fundamentally a weak process: hence the high Q and the relatively long radiative lifetime of a state, $\tau = 1/\Gamma$. For example, for the $2P \rightarrow 1S$ transition in hydrogen (the L_α transition), we have $\omega = 3/8$, and taking $r_{2p,1s} \approx 1$, we find $\tau = 3.6 \times 10^7$ atomic units, or 0.8 ns. The actual lifetime is 1.6 ns.

The lifetime for a strong transition in the optical region is typically 10–100 ns. Because of the ω^3 dependence of Γ^0 , the radiative lifetime for a transition in the microwave region—for instance an electric dipole rotational transition in a molecule—is longer by the factor $(\lambda_{\text{microwave}}/\lambda_{\text{optical}})^3 \approx 10^{15}$, yielding lifetimes on the order of months. Furthermore, if the transition moment is magnetic dipole rather than electric dipole, the lifetime is further increased by a factor of α^{-2} , giving a time of thousands of years.

7.9 Line Strength

Because the absorption and stimulated emission rates are proportional to the spontaneous emission rate, we shall focus our attention on the Einstein A coefficient:

$$A_{ba} = \frac{4}{3} \frac{e^2 \omega^3}{\hbar c^3} |\langle b | \mathbf{r} | a \rangle|^2 \quad (7.83)$$

where

$$|\langle b | \mathbf{r} | a \rangle|^2 = |\langle b | x | a \rangle|^2 + |\langle b | y | a \rangle|^2 + |\langle b | z | a \rangle|^2 \quad (7.84)$$

For an isolated atom, the initial and final states will be eigenstates of total angular momentum. (If there is an accidental degeneracy, as in hydrogen, it

is still possible to select angular momentum eigenstates.) If the final angular momentum is J_a , then the atom can decay into each of the $2J_a + 1$ final states, characterized by the azimuthal quantum number $m_a = -J_a, -J_a + 1, \dots, +J_a$. Consequently,

$$A_{ba} = \frac{4 e^2 \omega^3}{3 \hbar c^3} \sum_{m_a} |\langle b, J_b | \mathbf{r} | a, J_a, m_a \rangle|^2 \quad (7.85)$$

The upper level, however, is also degenerate, with a $(2J_b + 1)$ -fold degeneracy. The lifetime cannot depend on which state the atom happens to be in. This follows from the isotropy of space: m_b depends on the orientation of \mathbf{J}_b with respect to some direction in space, but the decay rate for an isolated atom can't depend on how the atom happens to be oriented. Consequently, it is convenient to define the *line strength* S_{ba} , given by

$$S_{ba} = S_{ab} = \sum_{m_b} \sum_{m_a} |\langle b, J_b, m_b | \mathbf{r} | a, J_a, m_a \rangle|^2 \quad (7.86)$$

Then,

$$A_{ba} = \frac{4 e^2 \omega^3}{3 \hbar c^3} \frac{S_{ba}}{g_b} = \frac{4 e^2 \omega^3}{3 \hbar c^3} \frac{S_{ba}}{2J_b + 1} \quad (7.87)$$

The line strength is closely related to the average oscillator strength \bar{f}_{ab} . \bar{f}_{ab} is obtained by averaging f_{ab} over the initial state $|b\rangle$, and summing over the values of m in the final state, $|a\rangle$. For absorption, $\omega_{ab} > 0$, and

$$\bar{f}_{ab} = \frac{2m}{3\hbar} \omega_{ab} \frac{1}{2J_b + 1} \sum_{m_b} \sum_{m_a} |\langle b, J_b, m_b | \mathbf{r} | a, J_a, m_a \rangle|^2 \quad (7.88)$$

It follows that

$$\bar{f}_{ba} = -\frac{2J_b + 1}{2J_a + 1} \bar{f}_{ab}. \quad (7.89)$$

In terms of the oscillator strength, we have

$$\bar{f}_{ab} = \frac{2m}{3\hbar} \omega_{ab} \frac{1}{2J_b + 1} S_{ab}. \quad (7.90)$$

$$\bar{f}_{ba} = -\frac{2m}{3\hbar} |\omega_{ab}| \frac{1}{2J_a + 1} S_{ab}. \quad (7.91)$$

7.10 Excitation by narrow and broad band light sources

We have calculated the rate of absorption and emission of an atom in a thermal field, but a more common situation involves interaction with a light beam, either monochromatic or broad band. Here “broad band” means having a spectral width that is broad compared to the natural line width of the system—the spontaneous decay rate.

For an electric dipole transition, the radiation interaction is

$$|H_{ba}| = e |\mathbf{r}_{ba}| \cdot \hat{\mathbf{e}} E / 2, \quad (7.92)$$

where E is the amplitude of the field. The transition rate, from Eq. 7.78, is

$$W_{ab} = \frac{\pi}{2} \frac{e^2 |\hat{\mathbf{e}} \cdot \mathbf{r}_{ba}|^2 E^2}{\hbar^2} f(\omega_0) = \frac{\pi}{2} \frac{e^2 |\hat{\mathbf{e}} \cdot \mathbf{r}_{ba}|^2 E^2}{\hbar} f(E_b - E_a) \quad (7.93)$$

where $\omega_0 = (E_b - E_a)/\hbar$ and $f(\omega)$ is the normalized line shape function, or alternatively, the normalized density of states, expressed in frequency units. The transition rate is proportional to the intensity I_0 of a monochromatic radiation source. I_0 is given by the Poynting vector, and can be expressed by the electric field as $E^2 = 8\pi I_0/c$. Consequently,

$$W_{ab} = \frac{4\pi^2}{c} \frac{e^2 |\hat{\mathbf{e}} \cdot \mathbf{r}_{ba}|^2}{\hbar^2} I_0 f(\omega_0) \quad (7.94)$$

In the case of a Lorentzian line having a FWHM of Γ_0 centered on frequency ω_0 ,

$$f(\omega) = \frac{1}{\pi} \frac{(\Gamma_0/2)}{(\omega - \omega_0)^2 + (\Gamma_0/2)^2} \quad (7.95)$$

In this case,

$$W_{ab} = \frac{8\pi e^2}{c\hbar^2 \Gamma_0} |\langle b | \hat{\mathbf{e}} \cdot \mathbf{r} | a \rangle|^2 I_0 \quad (7.96)$$

Note that W_{ab} is the rate of transition between two particular quantum states, not the total rate between energy levels. Naturally, we also have $W_{ab} = W_{ba}$.

An alternative way to express Eq. 7.93 is to introduce the Rabi frequency,

$$\Omega_R = \frac{2H_{ba}}{\hbar} = \frac{e|\hat{\mathbf{e}} \cdot \mathbf{r}_{ba}|E}{\hbar} \quad (7.97)$$

In which case

$$W_{ab} = \frac{\pi}{2} \Omega_R^2 f(\omega_0) = \Omega_R^2 \frac{1}{\Gamma_0} \quad (7.98)$$

If the width of the final state is due solely to spontaneous emission, $\Gamma_0 = A = (4e^2\omega^3/3\hbar c^3)|r_{ba}|^2$. Since W_{ab} is proportional to $|r_{ba}|^2/A_0$, it is independent of $|r_{ba}|^2$. It is left as a problem to find the exact relationship, but it can readily be seen that it is of the form

$$W_{ab} = X\lambda^2 I_0/\hbar\omega \quad (7.99)$$

where X is a numerical factor. $I/\hbar\omega$ is the photon flux—i.e. the number of photons per second per unit area in the beam. Since W_{ab} is an excitation rate, we interpret $X\lambda^2$ as the resonance absorption cross section for the atom, σ_0 .

At first glance it is puzzling that σ_0 does not depend on the structure of the atom; one might expect that a transition with a large oscillator strength—i.e. a large value of $|r_{ab}|^2$ —should have a large absorption cross section. However, the absorption rate is inversely proportional to the linewidth, and since that also increases with $|r_{ab}|^2$, the two factors cancel out. This behavior is not limited to electric dipole transitions, but is quite general.

There is, however, an important feature of absorption that does depend on the oscillator strength. σ_0 is the cross subsection assuming that the radiation is monochromatic compared to the natural line width. As the spontaneous decay rate becomes smaller and smaller, eventually the natural linewidth becomes narrower than the spectral width of the laser, or whatever source is used. In that case, the excitation becomes broad band.

We now discuss broad band excitation. Using the result of the last subsection, finding the excitation rate or the absorption cross subsection for broad band excitation is trivial. From Eq. 7.93, the absorption rate is proportional to $f(\omega_0)$. For monochromatic excitation, $f(\omega_0) = (2/\pi)A^{-1}$ and $W_{\text{mono}} = X\lambda^2 I_0/\hbar\omega$. For a spectral source having linewidth $\Delta\omega_s$, defined so that the normalized line shape function is $f(\omega_0) = (2/\pi)\Delta\omega_s^{-1}$, then the broad band excitation rate is obtained by replacing Γ_0 with $\Delta\omega_s$ in Eq. 7.99. Thus

$$W_B = \left(X\lambda^2 \frac{\Gamma_0}{\Delta\omega_s} \right) \frac{I_0}{\hbar\omega} \quad (7.100)$$

Similarly, the effective absorption cross section is

$$\sigma_{\text{eff}} = \sigma_0 \frac{\Gamma_0}{\Delta\omega_s} \quad (7.101)$$

This relation is valid provided $\Delta\omega_s \gg \Gamma_0$. If the two widths are comparable, the problem needs to be worked out in detail, though the general behavior would be for $\Delta\omega_s \rightarrow (\Delta\omega_s^2 + \Gamma_0^2)^{1/2}$. Note that $\Delta\omega_s$ represents the actual resonance width. Thus, if Doppler broadening is the major broadening mechanism then

$$\sigma_{\text{eff}} = \sigma_0 \Gamma_0 / \Delta\omega_{\text{Doppler}}. \quad (7.102)$$

Except in the case of high resolution laser spectroscopy, it is generally true that $\Delta\omega_s \gg \Gamma_0$, so that $\sigma_{\text{eff}} \ll \sigma_0$.

7.11 Higher-order radiation processes

The atom-field interaction is given by Eq. 7.25

$$H_{ba} = \frac{e}{mc} \langle b | \mathbf{p} \cdot \mathbf{A}(\mathbf{r}) | a \rangle \quad (7.103)$$

For concreteness, we shall take $\mathbf{A}(\mathbf{r})$ to be a plane wave of the form

$$\mathbf{A}(\mathbf{r}) = A \hat{\mathbf{z}} e^{ikx} \quad (7.104)$$

Expanding the exponential, we have

$$H_{ba} = \frac{eA}{mc} \langle b | p_z (1 + ikz + (ikz)^2/2 + \dots) | a \rangle \quad (7.105)$$

If dipole radiation is forbidden, for instance if $|a\rangle$ and $|b\rangle$ have the same parity, then the second term in the parentheses must be considered. We can rewrite it

as follows:

$$p_z x = (p_z x - z p_x)/2 + (p_z x + z p_x)/2. \quad (7.106)$$

The first term is $-\hbar L_y/2$, and the matrix element becomes

$$-\frac{ieAk}{2m} \langle b | \hbar L_y | a \rangle = -iAk \langle b | \mu_B L_y | a \rangle \quad (7.107)$$

where $\mu_B = e\hbar/2m$ is the Bohr magneton.

The magnetic field, is $\mathbf{B} = -ikA\hat{\mathbf{y}}$. Consequently, Eq. 7.107 can be written in the more familiar form $-\boldsymbol{\mu} \cdot \mathbf{B}$ (The orbital magnetic moment is $\boldsymbol{\mu} = -\mu_B \mathbf{L}$: the minus sign arises from our convention that e is positive.)

We can readily generalize the matrix element to

$$H_{\text{int}}(M1) = \mathbf{B} \cdot \langle b | \mu_B \mathbf{L} | a \rangle \quad (7.108)$$

where $M1$ indicates that the matrix element is for a magnetic dipole transition.

The second term in Eq. 7.106 involves $(p_z x + z p_x)/2$. Making use of the commutator relation $[\mathbf{r}, H_0] = i\hbar\mathbf{p}/m$, we have

$$\frac{1}{2}(p_z x + z p_x) = \frac{m}{2i\hbar}([z, H_0]x + z[x, H_0]) = \frac{m}{2i\hbar}(-H_0 zx + zx H_0) \quad (7.109)$$

So, the contribution of this term to the matrix element in Eq. 7.105 is

$$\frac{ieA}{m} \frac{km}{2\hbar i} \langle b | -H_0 zx + zx H_0 | a \rangle = -\frac{eAk}{2c} \frac{E_b - E_a}{\hbar} \langle b | zx | a \rangle = \frac{ieE\omega}{2c} \langle b | zx | a \rangle \quad (7.110)$$

where we have taken $E = ikA$. This is an electric quadrupole interaction, and we shall denote the matrix element by

$$H_{\text{int}}(E2)' = \frac{ie\omega}{2c} \langle b | xz | a \rangle E \quad (7.111)$$

The prime indicates that we are considering only one component of a more general expression.

The total matrix element from Eq. 7.105 can be written

$$H_{\text{int}}^{(2)} = H_{\text{int}}(M1) + H_{\text{int}}(E2). \quad (7.112)$$

where the superscript (2) indicates that we are looking at the second term in the expansion of Eq. 7.105. Note that $H_{\text{int}}(M1)$ is real, whereas $H_{\text{int}}(E2)$ is imaginary. Consequently,

$$|H_{\text{int}}^{(2)}|^2 = |H_{\text{int}}(M1)|^2 + |H_{\text{int}}(E2)|^2 \quad (7.113)$$

The magnetic dipole and electric quadrupole terms do not interfere.

The magnetic dipole interaction,

$$H(M1) \sim \mathbf{B} \cdot \langle b | \boldsymbol{\mu} | a \rangle \quad (7.114)$$

is of order α compared to an electric dipole interaction because $\mu = \alpha/2$ atomic

units.

The electric quadrupole interaction

$$H(E2) \sim e \frac{\omega}{c} \langle b | xz | a \rangle \quad (7.115)$$

is also of order α . Because transitions rates depend on $|H_{ba}|^2$, the magnetic dipole and electric quadrupole rates are both smaller than the dipole rate by $\alpha^2 \sim 5 \times 10^{-5}$. For this reason they are generally referred to as *forbidden* processes. However, the term is used somewhat loosely, for there are transitions which are much more strongly suppressed due to other selection rules, as for instance triplet to singlet transitions in helium.

7.12 Selection rules

The dipole matrix element for a particular polarization of the field, $\hat{\mathbf{e}}$, is

$$\hat{\mathbf{e}} \cdot \mathbf{r}_{ba} = \hat{\mathbf{e}} \cdot \langle b, J_b, m_b | \mathbf{r} | a, J_a, m_a \rangle. \quad (7.116)$$

It is straightforward to calculate x_{ba}, y_{ba}, z_{ba} , but a more general approach is to write \mathbf{r} in terms of a spherical tensor. This yields the selection rules directly, and allows the matrix element to be calculated for various geometries using the Wigner-Eckart theorem, as discussed in various quantum mechanics text books.

The orbital angular momentum operator of a system with total angular momentum L can be written in terms of a spherical harmonic $Y_{L,M}$. Consequently, the spherical harmonics constitute spherical tensor operators. A vector can be written in terms of spherical harmonics of rank 1. This permits the vector operator \mathbf{r} to be expressed in terms of the spherical tensor $T_{1,M}(\mathbf{r})$.

The spherical harmonics of rank 1 are

$$Y_{1,0} = \sqrt{\frac{3}{4\pi}} \cos \theta; \quad Y_{1,+1} = -\sqrt{\frac{3}{8\pi}} \sin \theta e^{+i\phi}; \quad Y_{1,-1} = \sqrt{\frac{3}{8\pi}} \sin \theta e^{-i\phi} \quad (7.117)$$

These are normalized so that

$$\int Y_{1,m'}^* Y_{1,m} \sin \theta d\theta d\phi = \delta_{m',m} \quad (7.118)$$

We can write the vector \mathbf{r} in terms of components r_m , $m = +1, 0, -1$,

$$r_0 = r \sqrt{\frac{4\pi}{3}} Y_{1,0}, \quad r_{\pm} = r \sqrt{\frac{4\pi}{3}} Y_{1,\pm 1}, \quad (7.119)$$

or, more generally

$$r_M = r T_{1M}(\theta, \phi) \quad (7.120)$$

Consequently,

$$\langle b, J_b, m_b | r_M | a, J_a, m_a \rangle = \langle b, J_b, m_b | r T_{1M} | a, J_a, m_a \rangle \quad (7.121)$$

$$= \langle b, J_b | r | a, J_a \rangle \langle J_b, m_b | T_{1M} | J_a, m_a \rangle \quad (7.122)$$

The first factor is independent of m . It is

$$r_{ba} = \int_0^\infty R_{b,J_b}^*(r) r R_{a,J_a}(r) r^2 dr \quad (7.123)$$

where r_{ba} contains the radial part of the matrix element. It vanishes unless $|b\rangle$ and $|a\rangle$ have opposite parity. The second factor in Eq. 7.122 yields the selection rule

$$|J_b - J_a| = 0, 1; \quad m_b = m_a \pm M = m_a, m_a \pm 1 \quad (7.124)$$

Similarly, for magnetic dipole transition, Eq. 7.108, we have

$$H_{ba}(M1) = \mu_B B \langle b, J_b, m_b | T_{LM}(L) | a, J_a, m_a \rangle \quad (7.125)$$

It immediately follows that parity is unchanged, and that

$$|\Delta J| = 0, 1 \quad (J = 0 \rightarrow J = 0 \text{ forbidden}); \quad |\Delta m| = 0, 1 \quad (7.126)$$

The electric quadrupole interaction Eq. 7.111, is not written in full generality. Nevertheless, from Slichter, Table 9.1, it is evident that xz is a superposition of $T_{2,1}(\mathbf{r})$ and $T_{2,-1}(\mathbf{r})$. (Specifically, $xz = (T_{2,-1}(\mathbf{r}) - T_{2,1}(\mathbf{r}))/4$.)

In general, then, we expect that the quadrupole moment can be expressed in terms of $T_{2,M}(\mathbf{r})$. There can also be a scalar component which is proportional to $T_{0,0}(r)$.

Consequently, for quadrupole transition we have: parity unchanged

$$|\Delta J| = 0, 1, 2, \quad (J = 0 \rightarrow J = 0 \text{ forbidden}) \quad |\Delta m| = 0, 1, 2. \quad (7.127)$$

This discussion of matrix elements, selection rules, and radiative processes barely skims the subject. For an authoritative treatment, the books by Shore and Manzel, and Sobelman are recommended.

References

- [1] E.T. Jaynes and F.W. Cummings, Proc. IEEE, **51**, 89 (1963).
- [2] A. Einstein, Z. Phys. **18**, 121 (1917), reprinted in English by D. ter Haar, *The Old Quantum Theory*, Pergamon, Oxford.