

# Quantum Mechanics in Astrophysics

Neal J. Evans II

August 28, 2014

# Contents

<b>0 Preliminaries: Introduction to Astronomical Spectroscopy</b>	<b>3</b>
0.1 Radiative Transfer . . . . .	7
0.2 LTE . . . . .	10
0.3 Non - LTE . . . . .	12
0.4 Basic Collision Theory . . . . .	15
<b>1 Basic Concepts of Quantum Mechanics</b>	<b>21</b>
1.1 Hamiltonian Formulation . . . . .	21
1.2 Schrodinger Equation . . . . .	22
1.3 Operating with Quantum Mechanics . . . . .	24
1.4 Dirac Notation . . . . .	26
1.5 Illustrative Solutions . . . . .	27
1.6 Time Independent Perturbation Theory . . . . .	34
1.7 Variational Method . . . . .	35
<b>2 Radiative Transitions</b>	<b>39</b>
2.1 Time Dependent Perturbation Theory . . . . .	39
2.2 The Electromagnetic Hamiltonian . . . . .	42
2.3 The Einstein Coefficients . . . . .	46
2.4 The Dipole Approximation . . . . .	48
2.5 Magnetic Dipole and Electric Quadrupole . . . . .	48
2.6 Symmetry Operations and Some Selection Rules . . . . .	49
2.7 Natural Line Broadening . . . . .	51
2.8 Summary . . . . .	52
<b>3 Atomic Structure – Hydrogen</b>	<b>55</b>
3.1 Review of the Hydrogen Atom without Spin . . . . .	55
3.2 Spin . . . . .	63
3.3 The Hydrogen Atom with Spin . . . . .	65
3.3.1 The Zeeman Effect . . . . .	69
3.4 The Hydrogen Atom with Nuclear Spin . . . . .	71
3.5 The Spectrum of Hydrogen . . . . .	73
3.5.1 Bound-Bound ( $\Delta n \neq 0$ ) Transitions . . . . .	73
3.5.2 Bound-Free Transitions . . . . .	76
3.5.3 Free-Free Transitions . . . . .	78

3.5.4	Fine Structure Transitions ( $\Delta n = 0$ ) . . . . .	79
3.5.5	Hyperfine Transition . . . . .	81
<b>4</b>	<b>Atomic Structure with Two Electrons— Helium</b>	<b>82</b>
4.1	The Many Particle Problem . . . . .	82
4.2	Identical Particles and Exchange Symmetry . . . . .	83
4.3	Two Identical Particles . . . . .	84
4.4	The Helium Atom . . . . .	86
<b>5</b>	<b>Atomic Structure with Many Electrons</b>	<b>90</b>
5.1	N-Electron Systems . . . . .	90
5.2	The Shell Model of the Atom . . . . .	92
5.3	Angular Momentum Coupling in Atoms . . . . .	95
5.4	Structure of Energy Levels in Atoms . . . . .	101
5.5	Transitions and Spectra . . . . .	102
5.5.1	General Considerations . . . . .	102
5.5.2	Magnetic Dipole Transitions . . . . .	103
5.5.3	Electric Pole Transitions . . . . .	103
5.6	Elements of Astrophysical Interest . . . . .	105
5.6.1	Carbon . . . . .	105
<b>6</b>	<b>Introduction to Molecules in Astronomy</b>	<b>109</b>
6.1	Why, Where, and What? . . . . .	109
6.2	The Born-Oppenheimer Equation . . . . .	109
6.3	Symmetries and Nomenclature . . . . .	113
<b>7</b>	<b>Diatomic Molecules</b>	<b>116</b>
7.1	Introduction . . . . .	116
7.2	The Non-Rigid Rotator and the Anharmonic Oscillator . . . . .	116
7.3	Vibrational and Rotational Transitions . . . . .	119
7.4	Spin and Statistics . . . . .	123
<b>8</b>	<b>Polyatomic Molecules: Rotation</b>	<b>125</b>
8.1	Introduction . . . . .	125
8.2	Linear Molecules . . . . .	125
8.3	The Symmetric Top . . . . .	125
8.4	The Asymmetric Top . . . . .	132
8.5	Spin and Statistics . . . . .	135
<b>9</b>	<b>Polyatomic Molecules: Vibration</b>	<b>140</b>
9.1	Degrees of Freedom, Normal Modes . . . . .	140
9.2	Rovibrational Transitions of Polyatomic Molecules . . . . .	143
<b>10</b>	<b>Including Other Angular Momenta</b>	<b>148</b>
10.1	Molecules with Residual Electronic Angular Momentum . . . . .	148

<b>11 Electronic Transitions</b>	<b>155</b>
11.1 Generalities and Nomenclature . . . . .	155
11.2 Selection Rules and Strengths . . . . .	157
11.2.1 Selection Rules Applying to Both Hund's Case a and Hund's Case b . . . . .	159
11.2.2 Selection Rules Applying only to Case a . . . . .	159
11.2.3 Selection Rules Applying only to Case b . . . . .	160
11.2.4 Summary . . . . .	160
<b>12 Hyperfine Structure in Molecules</b>	<b>162</b>
12.1 Introduction . . . . .	162
12.2 Molecules with Residual Electronic Angular Momentum . . . . .	162
12.3 Hyperfine Structure - Linear ${}^1\Sigma$ State . . . . .	165
12.3.1 Single Coupling Nucleus . . . . .	166
12.3.2 Frequency of Hyperfine Transitions and Relative Strength . . . . .	169
12.3.3 Multiple Coupling Nuclei . . . . .	172
12.3.4 Multiple Coupling Nuclei - Energy Levels & Strength . . . . .	173
12.4 Hyperfine Structure in Non-linear Molecules . . . . .	176
<b>A Angular Momentum</b>	<b>180</b>
A.1 Clebsch-Gordon Coefficients . . . . .	180
A.2 3-J Symbols . . . . .	181
A.3 Behavior of Operators under Rotation: Spherical Tensors . . . . .	182
A.4 The Wigner-Eckart Theorem . . . . .	183
A.5 The Dipole Moment Operator in Polarization Form . . . . .	184

# List of Figures

1.1	The rigid rotor as a dumbbell. . . . .	30
1.2	Spherical polar coordinates . . . . .	31
1.3	Here is example in which the perturbation is a magnetic field, which breaks the degeneracy in the azimuthal quantum number. . . . .	38
3.1	The H atom energy level diagram with a few transitions shown. . . . .	58
3.2	Bohr model with dubious distances, but showing more transitions. . . . .	59
3.3	The vector model for angular momentum . . . . .	60
3.4	The vector model for $l = 2$ with 5 possible projections . . . . .	60
3.5	Shows $J$ precessing around the z axis, $L$ and $S$ precessing around $J$ . . . . .	66
3.6	Splitting of $p_{\frac{1}{2}}$ in a magnetic field. . . . .	70
3.7	Figure from Dopita and Sutherland, <b>Astrophysics of the Diffuse Universe</b> , showing allowed transitions and the $\Delta\ell = \pm 1$ selection rule. . . . .	73
4.1	The He atom energy level diagram with a few transitions shown. . . . .	89
5.1	Madelung's Rule, also called Klechkowski's Rule. This gives the order of filling of atomic orbitals. . . . .	93
5.2	The Periodic Table of the Elements . . . . .	94
5.3	Radial probability distributions for He, Ne, and Ar. Taken from <i>Principles of General Chemistry</i> , available through the Creative Commons. . . . .	95
5.4	Energy levels for carbon from Herzberg's book, <i>Atomic Spectra and Atomic Structure</i> , pg. 143. . . . .	106
5.5	Splitting of ground configuration of carbon into terms and sublevels. Figure from Bransden and Joachain, <i>Physics of Atoms and Molecules</i> , pg. 348. . . . .	107
6.1	The lowest electronic energy state is shown with multiple vibrational energy levels. . . . .	113
6.2	The electronic and some vibrational levels are shown for $H_2$ . . . . .	114
6.3	The hierarchy of energy levels is indicated schematically. . . . .	115
7.1	The energy level diagram of CO, $^{13}\text{CO}$ , $\text{C}^{18}\text{O}$ , CS, $\text{HCO}^+$ , and $\text{DCO}^+$ for all energy levels less than 40 K above the ground state. The vertical axis is the energy above ground, $E/k$ , in K. All energy level diagrams are plotted on the same vertical scale. Notice that heavier molecules, such as CS, have more closely spaced energy levels. Figure supplied by Yancy Shirley. . . . .	120

8.1	A prolate symmetric top. . . . .	126
8.2	An oblate symmetric top. . . . .	127
8.3	The energy levels of prolate (left) and oblate (right) symmetric tops. Taken from Barrow, <i>Introduction to Molecular Spectroscopy</i> , 1962 . . . . .	128
8.4	The three Euler angles. Only one axis (labeled $N$ ) of the intermediate coordinate system is shown. . . . .	129
8.5	A spectrum showing different $K$ levels for the $J = 7 - 6$ transition of $\text{CH}_3\text{CN}$ . . . . .	131
8.6	A correlation diagram for asymmetric top energy levels. Prolate limit is on the left, oblate limit on the right. Taken from Townes and Schawlow. . . . .	133
8.7	Symmetry operations for $C_{2v}$ and $C_{3v}$ species. Taken from Barrow. . . . .	136
8.8	The potential for a harmonic oscillator with a quartic perturbation. Taken from Townes and Schawlow. . . . .	137
8.9	The wavefunctions for a harmonic potential with a hump in the middle. Taken from Townes and Schawlow. . . . .	138
8.10	The energy levels of $\text{NH}_3$ with the parity of the levels indicated by $\pm$ . Taken from Townes and Schawlow. . . . .	139
9.1	Two stretching modes (symmetric and antisymmetric) and two degenerate bending modes. . . . .	140
9.2	Three modes of water molecule with fundamental frequencies. . . . .	143
9.3	The parallel band of the linear molecule $\text{CO}_2$ , showing P and R branches. The other, weaker bands are caused by other species. . . . .	144
9.4	A perpendicular band of acetylene, showing a Q branch as well as P and R branches. Note the alternation of strong and weak lines, with a ratio of about 3, caused by the ortho-para effects of the identical hydrogens. . . . .	145
9.5	A parallel band of a symmetric top, showing the sub-bands caused by different $K$ ladders and the net result at the bottom. Figure and caption taken from Barrow, <i>Introduction to Molecular Spectroscopy</i> . . . . .	145
9.6	A perpendicular mode of a symmetric top, showing the sub-band structure, with two displaced sub-bands for each $K \neq 0$ . Figure and caption taken from Barrow, <i>Introduction to Molecular Spectroscopy</i> . . . . .	146
9.7	The resulting spectrum of a perpendicular band, in this case of $\text{CH}_3\text{Cl}$ . Figure and caption taken from Barrow, <i>Introduction to Molecular Spectroscopy</i> . . . . .	147
10.1	Hund's case a. $\vec{L}$ and $\vec{S}$ precess rapidly around the molecular axis, which precesses more slowly around $\vec{J}$ , the total angular momentum. $\vec{O}$ is the usual molecular rotation, which is perpendicular to the molecular axis. Figure from Townes and Schawlow. . . . .	149
10.2	Hund's case b. Precession of the molecular axis is slower than the precession of $\vec{L}$ about the axis, but faster than the precession of $\vec{N}$ and $\vec{S}$ about $\vec{J}$ , the total angular momentum. $\vec{O}$ is the usual molecular rotation, which is perpendicular to the molecular axis, and $\vec{N} = \vec{\Lambda} + \vec{O}$ . Figure from Townes and Schawlow. . . . .	150
10.3	The energy levels for CN. For now, we have explained only the first splitting. The further (hyperfine) splitting to the right is due to interactions with nuclear spins. . . . .	151
10.4	The energy level diagram of $\text{O}_2$ , showing transitions, frequencies, and Einstein A values. The transitions used to detect the molecule in space are indicated in blue. This figure is taken from Goldsmith et al. 2011, Ap. J., 737, 96. . . . .	152

10.5 OH energy levels with $\Lambda$ doubling. . . . .	153
11.1 A Fortrat parabola. This illustration assumes $B' = 0.8$ and $B'' = 1.0 \text{ cm}^{-1}$ . In this case, the bandhead appears at relatively low $J$ in the R branch. Figure from Wikipedia page on vibronic spectroscopy. . . . .	156
11.2 Illustration of the Franck-Condon principle. A transition between two electronic states tends to occur at constant internuclear separation. Absorption from the ground state to an excited electronic state with larger internuclear separation will likely occur to higher vibrational levels. Emission from the ground vibrational state of the excited electronic state is likely to result in a transition to the ground vibrational state of the ground electronic state. Figure from Wikipedia page by Mark Somoza. . . . .	158
12.1 The spin of nuclei in common astrophysically observed molecules. . . . .	163
12.2 The possible Hunds coupling cases including nuclear spin, $\vec{I}$ and total spin, $\vec{F}$ . Figure from Townes and Schawlow, figure 8.1. . . . .	164
12.3 OH energy levels with hyperfine splitting. . . . .	166
12.4 The vector model for the hyperfine structure of a linear molecule in a ${}^1\Sigma$ state. . . . .	167
12.5 The hyperfine energy level structure of HCN and $\text{C}^{17}\text{O}$ for $J = 0, 1$ , and $2$ . The vertical axis is the shift in energy ( $\Delta E = E_Q + E_M$ ) in kHz. Note the vertical scale is different for HCN and $\text{C}^{17}\text{O}$ . . . . .	168
12.6 The hyperfine spectrum of $\text{C}^{17}\text{O}$ $J = 2 \rightarrow 1$ and $J = 1 \rightarrow 0$ . The vertical axis is relative intensity and the horizontal axis is velocity (km/s) in the rest frame of the molecule. . . . .	169
12.7 Diagrammatic representation of the hyperfine energy levels for $\text{N}_2\text{H}^+$ $J= 1 \rightarrow 0$ . The $J= 0$ level is not actually split into separate energy levels. The seven observed transitions from the seven upper hyperfine levels are shown. . . . .	174
12.8 Observed spectra of the $J_{K-1K+1} = 1_{10} \rightarrow 1_{11}$ transition of $\text{H}_2\text{CO}$ on the left, with hyperfine structure indicated, toward three starless, quiescent molecular cores, using the Arecibo telescope. The spectra on the right are the $J_{K-1K+1} = 3_{12} \rightarrow 2_{11}$ transition. No hyperfine structure is visible because of the much higher frequency (225 GHz). Figure from Young et al. (2004), Ap. J. 614, 252. . . . .	177
12.9 Energy levels of the $(J, K) = (1, 1) \rightarrow (0, 0)$ inversion levels of $\text{NH}_3$ , with hyperfine structure due to the nitrogen and hydrogen nuclei. From Rydbeck et al. (1977) Ap. J. (Letters), 215, L35. . . . .	178
12.10 Spectrum of the $(J, K) = (1, 1) \rightarrow (0, 0)$ inversion transition of $\text{NH}_3$ , with hyperfine structure due to the nitrogen and hydrogen nuclei. From Rydbeck et al. (1977) Ap. J. (Letters), 215, L35. . . . .	179

# Preliminaries: Introduction to Astronomical Spectroscopy

## 0.1 Radiative Transfer

Here we want to develop some simple results from radiative transfer and develop the nomenclature and equations that will allow us to deal with astronomical problems. We will follow the approach in Chapter 3 of Spitzer’s book, *Physical Processes in the Interstellar Medium*.

We define a specific intensity  $I_\nu(\vec{k}, \vec{r}, t)$  so that  $I_\nu d\nu d\omega dA dt$  is the energy of the photons which during a time interval  $dt$ , pass through the area  $d\vec{A}$ , whose frequency lies within  $d\nu$  about  $\nu$  and whose direction lies within  $d\omega$  about  $\vec{k}$ ;  $d\vec{A}$  is located at  $\vec{r}$  and oriented perpendicular to  $\vec{k}$ .

How does  $I_\nu$  change as it interacts with matter? Consider a cylinder of length  $ds$  and volume  $dV = dsdA$ . Then  $\kappa_\nu dV I_\nu d\nu d\omega dt$  is the energy absorbed from a beam of specific intensity  $I_\nu$  in the cylinder;  $\kappa_\nu$  is the absorption coefficient. The energy emitted by the cylinder is  $j_\nu dV d\nu d\omega dt$  (in the intervals  $d\nu d\omega dt$ ). Then the change in  $I_\nu$  in distance  $ds$  is

$$\frac{dI_\nu}{ds} = -\kappa_\nu I_\nu + j_\nu. \quad (1)$$

Equation (1) is the fundamental equation of radiative transport. We define the “optical depth”  $d\tau'_\nu$  backward along the ray path:

$$d\tau'_\nu = -\kappa_\nu ds; \quad (2)$$

$\tau'_\nu = 0$  at the observer and increases toward the source; it is a measure of how far we see into the source. If we consider a region of total “optical depth”  $\tau_\nu$ , we integrate (1) to get

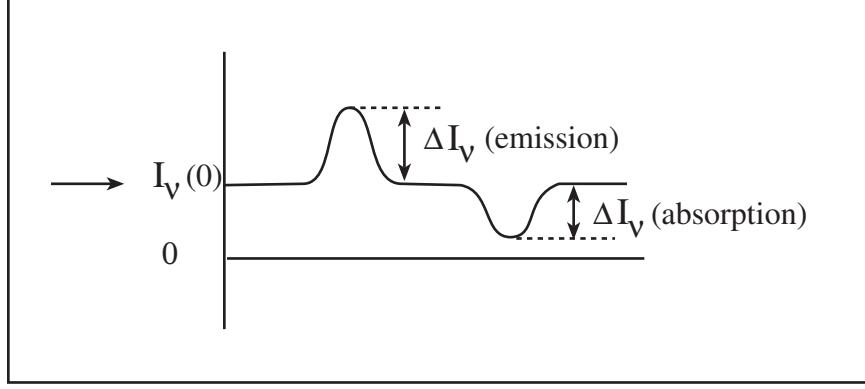
$$I_\nu = I_\nu(0)e^{-\tau_\nu} + \int_0^{\tau_\nu} e^{-\tau'_\nu} \frac{j_\nu}{\kappa_\nu} d\tau'_\nu, \quad (3)$$

where  $I_\nu(0)$  is the intensity at the back side of the region. We define the source function  $S_\nu \equiv j_\nu/\kappa_\nu$ . If  $S_\nu$  is constant through the region, we can take it out of the integral:

$$I_\nu = I_\nu(0)e^{-\tau_\nu} + S_\nu(1 - e^{-\tau_\nu}). \quad (4)$$

For spectral lines, we are usually interested in the *difference* between the line peak and the continuum level,  $I_\nu(0)$ . So we measure  $\Delta I_\nu = I_\nu - I_\nu(0)$ . Then

$$\Delta I_\nu = (S_\nu - I_\nu(0))(1 - e^{-\tau_\nu}). \quad (5)$$



If the radiation is a blackbody, we can replace  $I_\nu(0)$  by a Planck function at temperature,  $T$ :

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \left( e^{h\nu/kT} - 1 \right)^{-1}. \quad (6)$$

Likewise, if the material in the region is in thermal equilibrium,  $j_\nu$  and  $\kappa_\nu$  have a fixed ratio:

$$\frac{j_\nu}{\kappa_\nu} \equiv S_\nu = B_\nu(T). \quad (7)$$

In thermal equilibrium, the  $T$  in the Planck function is the kinetic temperature of the gas, and this equality is known as Kirchoff's law. Even if the system is not in thermal equilibrium, the source function at a specific frequency can be written as a Planck function at some temperature that is characteristic of the region. So we can rewrite (5) as

$$\Delta I_\nu = [B_\nu(T) - B_\nu(T_{bg})](1 - e^{-\tau_\nu}), \quad (8)$$

where we characterize the background intensity  $I_\nu(0)$  as a blackbody at temperature  $T_{bg}$ .

It turns out that equation 8 can also be used if the region is not in thermal equilibrium as long as we generalize  $T$  to be an excitation temperature, which need not be the physical temperature of the region. Likewise, as long as we deal with only a small range of  $\nu$ ,  $I_\nu(0)$  can always be approximated by a blackbody at some *effective* temperature,  $T_{bg}$ . If we are dealing with a spectral line from a transition between levels  $i$  and  $j$ , then the excitation temperature  $T_{ij}$  is defined by

$$\frac{n_j}{n_i} = \frac{g_j}{g_i} e^{-E_{ij}/kT_{ij}}, \quad (9)$$

$$\begin{array}{ccc} j & \xrightarrow{\hspace{1cm}} & g_j \\ i & \xrightarrow{\hspace{1cm}} & g_i \end{array} \left. \right\} E_{ij}$$

where  $n_i$  is the population in level  $i$ , which has degeneracy  $g_i$ , and  $E_{ij}$  is the energy difference. Then equation (8) is generalized to

$$\Delta I_\nu = [B_\nu(T_{ij}) - B_\nu(T_{bg})](1 - e^{-\tau_\nu}). \quad (10)$$

At radio wavelengths ( $h\nu \ll kT$ ),

$$B_\nu(T) \rightarrow \frac{2k\nu^2}{c^2} T. \quad (11)$$

Radio astronomers use a thermal calibration scale in which  $B_\nu(T_B) = I_\nu$  defines  $T_B$ , so that equation (10) becomes

$$\Delta T_B = (T_{ij} - T_{bg})(1 - e^{-\tau_\nu}). \quad (12)$$

This form has some intuitive advantages: if  $T_{ij} > T_{bg}$ , you get  $\Delta T_B > 0$  (emission); if  $T_{ij} < T_{bg}$ ,  $\Delta T_B < 0$  (absorption). At short radio wavelengths (millimeter, submillimeter)  $h\nu$  may not be much less than  $kT$ , but the thermal calibration scale is still used. One defines a radiation temperature ( $T_R$ ) by

$$T_R \equiv \frac{\lambda^2}{2k} I_\nu, \quad (13)$$

and equation (10) is written as

$$\Delta T_R = [J_\nu(T_{ij}) - J_\nu(T_{bg})](1 - e^{-\tau_\nu}), \quad (14)$$

where  $J_\nu$  is a Planck function in temperature units [ $J_\nu(T) = h\nu/k(e^{h\nu/kT} - 1)^{-1}$ ]. Note that this  $J_\nu$  is not the same as the mean intensity in some books, such as **Rybicki and Lightman**. Note that the  $\Delta$  is usually dropped and some people use  $T_B$  for  $T_R$ , but the “official” convention is that  $T_R = J_\nu(T_B)$ ; i.e.,  $T_B$  is the real “brightness temperature”.  $T_R \rightarrow T_B$  as  $h\nu/kT \rightarrow 0$ .

Now, let us consider some special cases:

- a) strong background, negligible emission (e.g., stellar atmosphere, interstellar lines seen in absorption at visible or ultraviolet wavelengths.) Then

$$\Delta I_\nu = -I_\nu(0)(1 - e^{-\tau_\nu}) \quad (15)$$

and a useful quantity is the equivalent width:

$$W_\nu = \int (1 - e^{-\tau_\nu}) d\nu = - \int \frac{\Delta I_\nu d\nu}{I_\nu(0)}; \quad (16)$$

or, in wavelength units,

$$W_\lambda = \int (1 - e^{-\tau_\lambda}) d\lambda = \frac{\lambda^2}{c} W_\nu. \quad (17)$$

- b) A very optically thin region producing line emission. Then since  $\kappa_\nu \rightarrow 0$ ,

$$\frac{dI_\nu}{ds} = j_\nu \Rightarrow I_\nu = \int j_\nu ds. \quad (18)$$

By now it should be clear that we need to be able to compute  $j_\nu$  and  $\kappa_\nu$  – that is one of the main points of this course. First, we relate  $j_\nu$  and  $\kappa_\nu$  to the Einstein coefficients  $A_{ji}$  and  $B_{ji}$ .  $A_{ji}$  is the probability per second that the system makes a spontaneous downward transition from  $j$  to  $i$ . We have

$$\int j_\nu d\nu = \frac{h\nu_{ij}n_j A_{ji}}{4\pi}. \quad (19)$$

The relation of  $\kappa_\nu$  to the Einstein B's depends on one's convention. If the rate of stimulated absorption is  $B_{ij}I_\nu$ , then,

$$\int \kappa_\nu d\nu = \frac{h\nu_{ij}}{4\pi}(n_i B_{ij} - n_j B_{ji}). \quad (20)$$

If the rate of stimulated absorption is  $B_{ij}U_\nu$ , where  $U_\nu$  is the energy density,  $U_\nu = (4\pi/c)I_\nu$ , then

$$\int \kappa_\nu d\nu = \frac{h\nu_{ij}}{c}(n_i B_{ij} - n_j B_{ji}). \quad (21)$$

For this course, I adopt the energy density convention and use equation (21), following Spitzer. Most courses in stellar atmospheres make the opposite choice. Note that stimulated emission enters as a negative absorption.

Relations between B's and A's can be derived by noting that, in thermal equilibrium, the rate of energy absorption must equal the rate of energy emission. This fact leads to

$$g_i B_{ij} = g_j B_{ji} = \frac{c^3}{8\pi h\nu_{ij}^3} g_j A_{ji}, \quad (22)$$

where the last equality applies only to the energy density convention for  $B$ .

Now, we can see that interpretation of astronomical data, or theoretical modeling of astronomical radiative processes, will require knowledge of the Einstein coefficients, the frequencies of transitions ( $\nu_{ij}$ ) and the populations of the relevant levels ( $n_i$  and  $n_j$ ). Much of the course will be devoted to the first two items; we now turn our attention briefly to the last item.

## 0.2 LTE

Local Thermodynamic Equilibrium (LTE) describes a state of affairs in which some (usually those relevant to the problem at hand) of the physical properties assume their thermodynamic equilibrium values. For example, if the populations in different levels of an atom or molecule are in LTE, their ratios are given by the Boltzmann equation:

$$\frac{n_j}{n_i} = \frac{g_j}{g_i} e^{-E_{ij}/kT}, \quad (23)$$

where  $T$  is the kinetic temperature (the temperature characterizing the random motions of particles). Comparison to equation 0.9 indicates that LTE implies  $T_{ij} = T$ , which is often described by saying that the transition is thermalized. Note that some transitions may be thermalized while others are not.

When several ionization stages may exist, one would like to know the relative numbers of species in different ionization stages. If we denote by  $X^{(r)}$  the abundance of element  $X$  ionized  $r$  times then we have the Saha equation when thermal equilibrium applies:

$$\frac{n(X^{(r+1)})n_e}{n(X^{(r)})} = \frac{Z_{r+1}Z_e}{Z_r}, \quad (24)$$

where  $Z_r$  is the partition function for  $X^{(r)}$  and  $Z_e$  is the partition function for electrons.

Partition functions have to do with the degrees of freedom of a system, so they include translational (external) and internal contributions:  $Z_r = Z_{tr}Z_{int}$ .

$$Z_{tr} = g_s \left( \frac{2\pi m_{X^{(r)}} kT}{h^2} \right)^{\frac{3}{2}}, \quad (25)$$

where  $m_{X^{(r)}}$  is the mass of  $X^{(r)}$  and  $g_s$  is its spin degeneracy. Thus for the electron, which has no internal excitation,

$$Z_e = 2 \left( \frac{2\pi m_e kT}{h^2} \right)^{\frac{3}{2}} = 4.829 \times 10^{15} T^{\frac{3}{2}}. \quad (26)$$

For particles with internal excitations,  $Z_{int}$  is the sum of the populations in all the internal states

$$Z_{int} = \sum_i n_i = \sum_i g_i e^{-E_{i1}/kT_{i1}}, \quad (27)$$

where  $E_{i1}$ , is the energy above the ground state and  $T_{i1}$  is the excitation temperature. In LTE,

$$Z_{int} = \sum_i g_i e^{-E_{i1}/kT}. \quad (28)$$

Since species  $X^{(r)}$  and  $X^{(r+1)}$  have essentially the same mass, their  $Z_{tr}$  will cancel out in equation (24), leaving  $Z_e$  and the internal partition functions:

$$\frac{n(X^{(r+1)})n_e}{n(X^{(r)})} = \frac{Z_e \sum_i g_i^{(r+1)} e^{-E_{i1}^{(r+1)}/kT}}{\sum_i g_i^{(r)} e^{-E_{i1}^{(r)}/kT}}, \quad (29)$$

where  $E_{i1}^{(r+1)}$  is the energy of state  $i$  of  $X^{(r+1)}$  above the ground state of  $X^{(r)}$ ; that is, it includes the energy needed to ionize  $X^{(r)}$ . If the  $Z_{int}$  are approximated by their first terms,

$$Z_e \frac{g_1^{(r+1)} e^{-E^{(r+1)}/kT}}{g_1^{(r)} e^{-E^{(r)}/kT}} = \frac{g_1^{(r+1)}}{g_1^{(r)}} Z_e e^{-\Phi_r/kT}, \quad (30)$$

where  $\Phi_r$  is the energy required to ionize  $X^{(r)}$  to  $X^{(r+1)}$ .

So, in this approximation, valid for many atoms,

$$\frac{n(X^{(r+1)})n_e}{n(X^{(r)})} = 2 \frac{g_1^{(r+1)}}{g_1^{(r)}} \left( \frac{2\pi m_e kT}{h^2} \right)^{\frac{3}{2}} e^{-\Phi_r/kT}. \quad (31)$$

The Saha and Boltzmann Equations are often combined into the Saha-Boltzmann equation. This is most often used when most atoms of element  $X$  are ionized one stage higher than the atoms which

are emitting the line which is observed (e.g., in ionized nebulae, where we observe lines from the few atoms of neutral H). Then

$$n_j(X^{(r)}) = \frac{g_j^{(r)} e^{-E_j^{(r)}/kT} n_e n(X^{(r+1)})}{Z_{r+1} Z_e}. \quad (32)$$

Note that, in this equation,  $Z_{r+1}$  includes only the internal partition function.

### 0.3 Non - LTE

Situations in which LTE actually applies are rare in astronomy, so one has to consider how to generalize the LTE results. We have already referred to one of these methods, the use of an excitation temperature to replace the kinetic temperature; this method is the usual choice in molecular line studies. In atomic problems, it is more common to use a departure coefficient, defined by

$$b_j \equiv \frac{n_j}{n_j(\text{LTE})}. \quad (33)$$

where it is assumed that one can calculate  $n_j(\text{LTE})$  from the Boltzmann equation or the Saha-Boltzmann equation. Then,

$$\frac{n_j}{n_i} = \frac{b_j}{b_i} \frac{g_j}{g_i} e^{-E_{ij}/kT}, \quad (34)$$

and one would multiply the Saha-Boltzmann equation (32) by  $b_j$ .

So much for definitions. How do we compute  $T_{ij}$  or  $b_j$ ? This is the problem of excitation. One must consider, for each state of relevance, all the processes which may change the population of that state. The relevant equations are

$$\frac{dn_j}{dt} = \underbrace{-n_j \sum_y \sum_k (R_{jk})_y}_{\text{depopulation processes}} + \underbrace{\sum_y \sum_k n_k (R_{kj})_y}_{\text{population processes}}, \quad (35)$$

where  $(R_{jk})_y$  is the rate of transitions from  $j$  to  $k$  due to process  $y$ .

In a dynamical situation, such as a flow, a jet, or a wind,  $dn_j/dt$  may have contributions due to “convection” [ $\vec{\nabla} \cdot (n\vec{v})$ ] which we ignore here. In time dependent situations,  $dn_j/dt \neq 0$ , but we usually have a steady state condition,

$$\frac{dn_j}{dt} = 0. \quad (36)$$

Then one solves the “equation of statistical equilibrium”. If  $M$  levels are relevant, there are  $M - 1$  linearly independent equations which determine the  $M - 1$  population ratios. Determination of the populations then requires knowledge of the total number density of the species in question. The processes of relevance may include ionization and recombination, in which case one solves a combined ionization-excitation problem. For simplicity, let's ignore ionization and recombination

for now. Then we have only excitation processes in a single species. We will consider both collisional and radiative processes. We know that radiative de-excitation from level  $j$  is given by

$$n_j \left( \sum_{i < j} A_{ji} + \sum_i B_{ji} U_\nu \right), \quad (37)$$

since spontaneous emission must be to lower levels, but either stimulated emission or absorption may cause a transition out of level  $j$ . Similarly collisions may remove population from level  $j$  by excitation or de-excitation:

$$n_j \sum_i C_{ji} \quad (38)$$

where  $C_{ji}$  is the rate at which collisions induce transitions from level  $j$  to  $i$ . The collisions are usually with a different species than the one whose excitation we are calculating; in ionized regions, electron collisions are usually most important, while in neutral regions, H atoms or H<sub>2</sub> molecules dominate. For generality, we will denote the density of colliding particles as  $n$  and define a collisional rate coefficient ( $\gamma_{ji}$ ) by

$$C_{ji} = n \gamma_{ji}. \quad (39)$$

Then our steady state rate equations become (for each level  $j$ )

$$n_j \left\{ \sum_{i < j} A_{ji} + \sum_i B_{ji} U_\nu + \sum_i n \gamma_{ji} \right\} = \sum_i n_i n \gamma_{ij} + \sum_i n_i B_{ij} U_\nu + \sum_{i > j} n_i A_{ij}. \quad (40)$$

The special case of the two-level atom or molecule provides a clear illustration of the principles involved. In this case, equations (40) now become for level 1:

$$\begin{aligned} n_1 (\underbrace{n \gamma_{12} + B_{12} U_\nu}_{\text{depopulation of level 1}}) &= n_2 (\underbrace{n \gamma_{21} + B_{21} U_\nu + A_{21}}_{\text{population of level 1}}); \end{aligned} \quad (41)$$

and for level 2:

$$n_2 (n \gamma_{21} + A_{21} + B_{21} U_\nu) = n_1 (n \gamma_{12} + B_{12} U_\nu). \quad (42)$$

Now it is obvious that these are really the same equation so we will only get the ratio of  $n_2/n_1$ :

$$\frac{n_2}{n_1} = \frac{n \gamma_{12} + B_{12} U_\nu}{n \gamma_{21} + B_{21} U_\nu + A_{21}}. \quad (43)$$

In the “Rayleigh-Jeans limit” ( $E_{12} \ll kT$  for all relevant  $T$ ), using relations between upward and downward collision rates derived later (see equation 67), equation 43 can be solved for the excitation temperature characterizing  $n_2/n_1$  to provide an intuitively useful formula:

$$T_{12} = T_K \frac{T_{bg} + D_{12}}{T_K + D_{12}}, \quad (44)$$

where

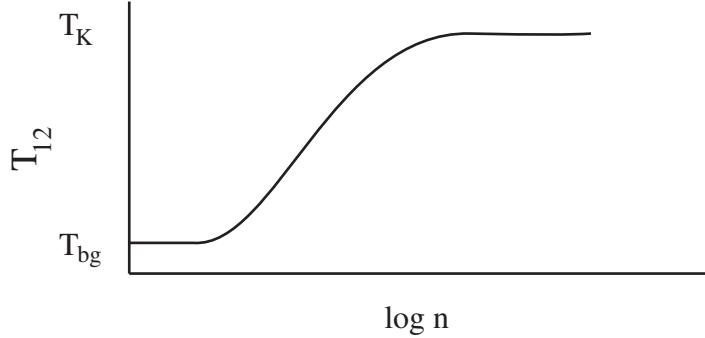
$$D_{12} = \frac{E_{12}}{k} \frac{n \gamma_{21}}{A_{21}}, \quad (45)$$

and  $T_{bg}$  is the brightness temperature of the radiation field,  $T_K$  is the kinetic temperature, and  $E_{12}$  is the energy difference between levels 1 and 2.  $A_{21}/\gamma_{21}$  has dimensions of  $\text{cm}^{-3}$  so we define the critical density  $n_c$ ,

$$n_c = \frac{A_{21}}{\gamma_{21}}. \quad (46)$$

Consider the two limits:  $n \ll n_c$  and  $n \gg n_c$ .

- a)  $n \ll n_c \Rightarrow D_{12} \rightarrow 0 \Rightarrow T_{12} \rightarrow T_{bg}$  and the populations approach equilibrium with the radiation field.
- b)  $n \gg n_c \Rightarrow D_{12} \gg T_{bg}$  and  $D_{12} \gg T_K \Rightarrow T_{12} \rightarrow T_K$  and the populations approach equilibrium with the collisions (LTE).



Another useful limit is that of high  $\nu$  (e.g., optical, ultraviolet lines) where  $A_{ji} \gg B_{ji}U_\nu$  unless  $U_\nu$  is very strong. (see equation 22) Then we can neglect the  $BU$  terms in equation 37, and, in particular, for the two-level system, we have

$$\frac{n_2}{n_1} = \frac{n\gamma_{12}}{n\gamma_{21} + A_{21}}, \quad (47)$$

which leads to

$$\frac{b_2}{b_1} = \frac{1}{1 + \frac{A_{21}}{n\gamma_{21}}}, \quad (48)$$

which is applicable, for example, to optical, permitted lines. In terms of excitation temperature,

$$T_{12} = T_K \left[ 1 + \frac{T_K}{E_{12}/k} \ln \left( 1 + \frac{A_{21}}{n\gamma_{21}} \right) \right]^{-1}. \quad (49)$$

Then the behavior of these equations in the limits  $n \gg n_c$  and  $n \ll n_c$  are illuminating:

$$\begin{aligned} n \gg n_c : \quad & \frac{b_2}{b_1} \rightarrow 1; T_{12} \rightarrow T_K \quad \text{LTE} \\ n \ll n_c : \quad & \frac{b_2}{b_1} \rightarrow \frac{n}{n_c} \quad \text{NLTE} \end{aligned} \quad (50)$$

implying a linear dependence of  $b_2$  on  $n$  and  $b_2/b_1 \ll 1$  for  $n \ll n_c$ .

Another situation of relevance occurs when stimulated absorption is important and optical depths are high. Then photons emitted spontaneously may excite other systems. This can be viewed as decreasing the effectiveness of spontaneous decay by a factor  $\beta$ , the escape probability. One can account for this by replacing  $A_{ji}$  by  $\beta_{ji}A_{ji}$ .

One may also need to deal with dilute radiation fields. For example, at a distance  $r$  from a star, the radiation field is still a blackbody at  $T_*$ , but the energy density is less by  $1/4(R_*/r)^2$  (assuming  $r \gg R_*$ ). More generally, this dilution factor is

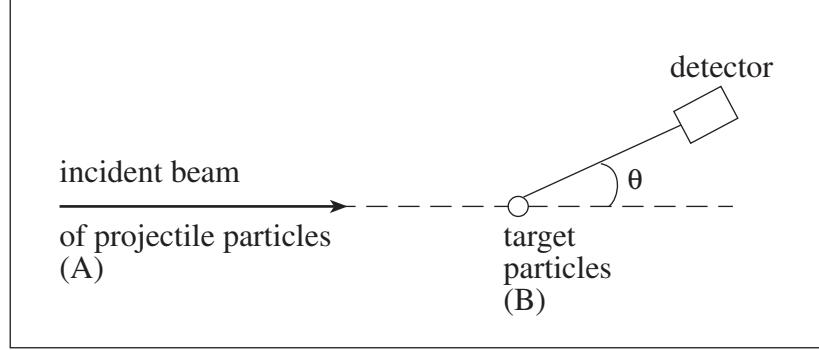
$$W = \frac{1}{2} \left[ 1 - \sqrt{1 - \frac{R_*^2}{r^2}} \right], \quad (51)$$

where  $U_\nu(r) = (4\pi/c)B_\nu(T_*)W(r)$ .

**Summary:** We have discussed a few simple ideas about non-LTE conditions. From these, it should be clear that in order to interpret observations in non-LTE situations, we need to be able to compute populations, and this may require a knowledge of collision rates,  $\gamma_{ij}$ , as well as radiative rates.

## 0.4 Basic Collision Theory

The terminology of collision theory is based largely on laboratory scattering experiments of the following general type:



The various kinds of processes that are possible include the following:

1. Elastic Scattering:  $A + B \rightarrow A + B$  (no change in internal energy state)
2. Nonelastic Scattering:  $A + B \rightarrow A + B^*$  ( $B$  is left in a different state)
3. Reactive Scattering:  $A + B \rightarrow C + D$ , or  $\rightarrow C_1 + C_2 + \dots + C_n$  ( $A + B$  turn into different particles).

A channel is a term applied to initial and final states of the combination of particles. In elastic scattering, the final channel is the same as the initial channel. In inelastic or reactive scattering the final channel is different. A channel is said to be open if it is allowed by relevant conservation

laws, such as energy conservation. Otherwise, it is closed. In astronomical problems, we are usually interested in the total cross section (integrated over all scattering angles).

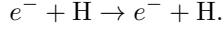
It is convenient to work in the center of mass system (theoretically, not experimentally!), so both particles in this system have initial kinetic energies. The total kinetic energy is  $E_i$  (initial) and  $E_f$  (final).

$$E_f = E_i + Q_{if}, \quad (52)$$

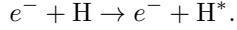
where  $Q_{if}$  is the change in internal energy. A reaction cannot occur unless  $E_f \geq 0$ . Thus if  $Q_{if} < 0$ , there is a threshold ( $E_i > |Q_{if}|$ ); such reactions are endothermic. If  $Q_{if} \geq 0$ , the reaction is exothermic and there is no threshold. In the laboratory frame of reference, where  $B$  is at rest, the threshold is

$$E_L^t = \frac{m_A + m_B}{m_B} E_{cm}^t, \quad (53)$$

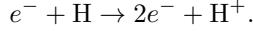
where  $E_{cm}^t$  is the threshold in the center of mass frame of reference. To illustrate, consider electrons scattering off H atoms, initially in the ground state. If  $E_i < 10.2$  eV, only elastic scattering is possible



Only the kinetic energy of the  $e^- + \text{H}$  change in the collision. If  $10.2 < E_i < 13.6$  eV, excitation of bound states is possible,



If  $E_i > 13.6$  eV, ionization (reactive scattering) is possible



Consider a target object which is spherical with radius  $r$  and thus cross sectional area  $\sigma = \pi r^2$ . If we have a flux of infinitesimal projectiles given by  $n(\text{cm}^{-3})$  and traveling at speed  $v$  relative to our target object, then the target will suffer collisions at a rate  $C(s^{-1})$ .

$$C = n\sigma v (= n\gamma) \quad (54)$$

This basic relation between collision rates and cross sections can be generalized, and the idea of a cross section can be generalized to include cross sections for particular processes (e.g., absorption of a photon, elastic collisions, inelastic collisions, a chemical reaction, etc.) The cross section may then be a function of the relative velocity, or energy [ $\sigma(v)$ ]. In laboratory experiments,  $v$  is usually controlled, but in astronomy,  $v$  usually varies over a wide range. One would want to average  $\sigma(v)$  over this range.

In most, but not all, astronomical situations the velocity distribution is a Maxwellian, characteristic of thermalization of the translational degrees of freedom. In fact, it is the elastic collisions which establish the distribution (cf. Spitzer, Chapter 2).<sup>1</sup> Since the random motions of emitting and absorbing particles are a source of line broadening, we will consider first the astronomical consequences of these elastic collisions: Doppler broadening by a Maxwellian.

---

<sup>1</sup>See also Ap.J. **206**, 435.

The Maxwellian velocity distribution is

$$f(v) = \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} e^{-(mv^2/2kT)}, \quad (55)$$

where  $f(v)d\vec{v}$  is the fractional number of particles whose velocity lies in volume element  $d\vec{v} = dv_x dv_y dv_z$ . The distribution function,  $f(v)$ , is isotropic. For Doppler broadening, however, only the velocity component along our line of sight contributes, so we need the distribution function along one axis, say the  $z$  axis:

$$g(v_z) = \left(\frac{m}{2\pi kT}\right)^{\frac{1}{2}} e^{-\frac{1}{2}mv_z^2/kT}, \quad (56)$$

so the  $1/e$  width is

$$\Delta v_{1/e} = \sqrt{\frac{2kT}{m}}, \quad (57)$$

and the full width at half maximum is

$$\Delta v = 2\sqrt{\frac{2(\ln 2)kT}{m}} = 0.215\sqrt{\frac{T}{m(\text{a.m.u.})}} \text{ km s}^{-1}, \quad (58)$$

It will also be useful to know the *mean speed*, independent of direction, in a Maxwellian. It is

$$\langle v \rangle = \left(\frac{8kT}{\pi m}\right)^{\frac{1}{2}} \simeq 0.15\sqrt{\frac{T}{m(\text{a.m.u.})}} \text{ km s}^{-1} \quad (59)$$

Note that the one-dimensional *rms velocity*,  $\sigma_{v_z} = (\langle v_z^2 \rangle^{1/2})$  is the  $1/\sqrt{e}$  width, or  $\sigma_{v_z} = \sqrt{kT/m}$ , so  $g(v_z) = 1/(\sqrt{2\pi}\sigma_{v_z})e^{-(v_z^2/2\sigma_{v_z}^2)}$  and the rms 3-d velocity is  $\sqrt{3}\sigma_{v_z} = \sqrt{3kT/m}$ . This is the quantity used in the virial theorem, for example. Another speed of interest is the sound speed,  $a$ . This is given by the equation,

$$a^2 = \frac{dP}{d\rho}, \quad (60)$$

where  $P$  is the pressure and  $\rho$  is the density. For an isothermal equation of state,  $P = nkT = \rho kT/\mu$ ,

$$a = \sqrt{kT/\mu}, \quad (61)$$

where  $\mu$  is the mean molecular weight.

Now let us return to our problem of averaging  $\sigma(v)$  over the velocity distribution. We will deal with  $\gamma(v) = \langle \sigma(v)v \rangle$  to remove the dependence on  $n$ , the density of projectiles. Since the *direction* of motion is irrelevant, we want to average over the distribution of *speeds*, given by

$$f(v)4\pi v^2 dv. \quad (62)$$

Then the average rate of collisions is

$$\begin{aligned} \gamma(T) &= \int_0^\infty \sigma(v)v4\pi v^2 f(v)dv \\ &= 4\pi \left(\frac{m}{2\pi kT}\right)^{\frac{3}{2}} \int_0^\infty \sigma(v)v^3 e^{-mv^2/2kT} dv \end{aligned} \quad (63)$$

Since the kinetic energy of the projectile is  $E = 1/2 mv^2$ ,  $dE = mvdv$ ,  $v^3dv = 2E/m^2dE$ , and we can convert this into an average over  $E$ :

$$\gamma(T) = \left(\frac{8kT}{\pi m}\right)^{\frac{1}{2}} \left[ \frac{1}{(kT)^2} \int_0^\infty \sigma(E) E e^{-E/kT} dE \right], \quad (64)$$

where we can recognize the first factor as the mean speed and regard the second factor as an average of  $\sigma$  over energy – convenient because  $\sigma$  is often computed as a function of energy.

Now consider inelastic collisions specifically. Imagine that the target is initially in state  $j$  and we want the rate at which collisions cause it to make a transition to state  $k$ . (When we use  $j$  and  $k$  as the indices, we generally allow either state to be the upper energy state.) The rate coefficient, averaged over the velocity distribution is  $\gamma_{jk}$ . If the state  $k$  lies above  $j$  in energy, a threshold kinetic energy is required. It must at least equal  $E_{jk} \equiv E_k - E_j$

$$\begin{array}{c} E_k \text{---} k \\ E_j \text{---} j \end{array} \quad \left. \right\} E_{jk}$$

Let  $E_j^{\text{kin}}$  equal the kinetic energy of the colliders when the system is in state  $j$  and  $E_k^{\text{kin}}$  the kinetic energy when the system is in state  $k$ . Conservation of *total* energy requires

$$E_j^{\text{kin}} + E_j = E_k^{\text{kin}} + E_k. \quad (65)$$

Time reversal invariance can be shown to imply a relation between forward and backward processes:

$$E_j^{\text{kin}} g_j \sigma_{jk}(E_j^{\text{kin}}) = E_k^{\text{kin}} g_k \sigma_{kj}(E_k^{\text{kin}}). \quad (66)$$

If our velocity distribution is a Maxwellian, we can use (65) and (66) to derive a relation between forward and backward rates:

$$\frac{\gamma_{jk}}{\gamma_{kj}} = \frac{g_k}{g_j} e^{-E_{jk}/kT} \quad (67)$$

This result was used to derive (44), (47), (49). It is usually easier to work with downward collision rates to avoid complications with thresholds and use (67) to get upward rates.

Now let's consider some specific cases: inelastic collisions with electrons and inelastic collisions with neutrals.

### Inelastic collisions with electrons

In ionized regions, we are usually concerned with collisions of electrons with ions. Electrons are usually more important as projectiles than protons because their mean speed is about 43 times higher. Unless  $\sigma$  for electrons is low for some reason, electron collisions will dominate. In this field, it is conventional to express the results in terms of a canonical cross section ( $\sigma_o$ ) times a fudge factor, called a “collision strength”,  $\Omega(j, k)$  [Note:  $\Omega(j, k) = \Omega(k, j)$ .]:

$$\sigma_{jk}(v_j) = \sigma_0 \Omega(j, k), \quad (68)$$

where  $v_j = \sqrt{\frac{2E_j}{m}}$ .

Table 1  
Collision Strengths for Excitation by Electrons

Number of p electrons	Ion	Levels		$E_{jk}$ (eV)	$\Omega(j, k)$	$\Sigma_j A_{kj}$ (s <sup>-1</sup> )
		Lower	Upper			
1,5	C II	$^2P_{1/2}$	$^2P_{3/2}$	0.0079	1.33	$2.4 \times 10^{-6}$
	Ne II	$^2P_{3/2}$	$^2P_{1/2}$	0.097	0.37	$8.6 \times 10^{-3}$
	Si II	$^2P_{1/2}$	$^2P_{3/2}$	0.036	7.70	$2.1 \times 10^{-4}$
2	N II	$^3P_0$	$^3P_1$	0.0061	0.41	$2.1 \times 10^{-6}$
		$^3P_0$	$^3P_2$	0.0163	0.28	$7.5 \times 10^{-6}$
		$^3P_1$	$^3P_2$	0.0102	1.38	$7.5 \times 10^{-6}$
		$^3P$	$^1D_2$	1.90	2.99	$4.0 \times 10^{-3}$
		$^3P$	$^1S_0$	4.05	0.36	1.1
	O III	$^3P_0$	$^3P_1$	0.014	0.39	$2.6 \times 10^{-5}$
		$^3P_0$	$^3P_2$	0.038	0.21	$9.8 \times 10^{-5}$
		$^3P_1$	$^3P_2$	0.024	0.95	$9.8 \times 10^{-5}$
		$^3P$	$^1D_2$	2.51	2.50	$2.8 \times 10^{-2}$
3	O II	$^4S_{3/2}$	$^2D_{5/2}$	3.32	0.88	$4.2 \times 10^{-5}$
		$^4S_{3/2}$	$^2D_{3/2}$	3.32	0.59	$1.8 \times 10^{-4}$
		$^2D_{3/2}$	$^2D_{5/2}$	0.0025	1.16	$4.2 \times 10^{-5}$

The canonical cross section may be motivated by considering the electron to be in a quasi-bound state of momentum  $p = mv_j$ . Then the uncertainty principle provides a size

$$\Delta r = \frac{\hbar}{p},$$

and a cross section

$$\pi\Delta r^2 = \pi \left( \frac{\hbar}{mv_j} \right)^2, \quad (69)$$

or, dividing by the degeneracy ( $g_j$ ) of state  $j$ , we arrive at the standard definition of  $\sigma_0$ :

$$\sigma_0 = \frac{\pi}{g_j} \left( \frac{\hbar}{2\pi mv_j} \right)^2 \quad (70)$$

In the absence of any better information, one can set  $\Omega = 1$ .

Integration over a Maxwellian leads to

$$\gamma_{kj} = \frac{h^2 \Omega(k, j)}{g_k (2\pi m)^{3/2} (kT)^{1/2}} = \frac{8.63 \times 10^{-6} \Omega(k, j)}{g_k T^{1/2}} \text{ cm}^3 \text{ s}^{-1}, \quad (71)$$

if  $\Omega(j, k)$  is not a function of velocity or is an average over velocity. Detailed calculations of  $\Omega$  have resulted in values ranging from 0.2 to 8. (See Table 1, from Spitzer, pg. 74). Newer calculations include more subtle temperature dependences and tables of these can be found in Appendix F of Draine's book, *Physics of the Interstellar and Intergalactic Medium*.

Table 2  
Rate Coefficients ( $\text{cm}^3\text{s}^{-1}$ ) for Deexcitation by  $\text{H}_2$

Temperature $T(\text{°K})$	10°	30°	100°	300°	1000°
$10^{11} \times \gamma_{10}$ for p- $\text{H}_2$ -CO <sup>a</sup>	3.3	3.3	3.5	3.5	3.4
$10^{11} \times \gamma_{10}$ for H <sub>2</sub> -CS <sup>b</sup>	3.5	3.0	2.8	3.0	...
$10^{11} \times \gamma_{10}$ for H <sub>2</sub> -HCN <sup>c</sup>	2.4	1.6	0.9	0.6	...

<sup>a</sup> Yang et al. 2010, ApJ, 718, 1062; p- $\text{H}_2$  refers to para species.

<sup>b</sup> Lique et al. 2006, Astr. Ap., 451, 1125

<sup>c</sup> Dumouchel et al. 2010, MNRAS, 406, 2488, scaled up by 1.37 from calculations for helium.

### Inelastic Collisions with Neutrals

Since neutral particles do not have very long range interactions, collision cross sections are more similar to the geometric cross sections. Since typical sizes are a few Å, a standard cross section is  $\sim 10^{-15} \text{ cm}^2$ . Cross sections for some processes, such as  $\text{H}_2 - \text{H}_2$  inelastic collisions, are notably less than this. Using equation (0.59) and assuming that  $\text{H}_2$  is the projectile, we would estimate a “typical” rate coefficient to be

$$\gamma = 1.0 \times 10^{-11} \sqrt{T} \text{ cm}^3 \text{ s}^{-1}. \quad (72)$$

Some actual rates are given in Table 2. Note that helium was assumed to be the projectile in the collisions with HCN. In most cases,  $\text{H}_2$  will be much more important, but collisions between  $\text{H}_2$  and other molecules are harder to calculate. In most cases, we have to assume that collision rates for  $\text{H}_2$  are the same as those with helium, with perhaps a scaling by 1.37 to account for the higher mean speed of the  $\text{H}_2$  compared to helium. Some calculations with  $\text{H}_2$  are now available, and the results show differences between collisions with  $\text{H}_2$  in different rotational states, so rates are given for para- $\text{H}_2$  and ortho- $\text{H}_2$  separately. A database that keeps up to date with new rate calculations exists at <http://home.strw.leidenuniv.nl/~moldata/> in Leiden. Some characteristic rates from that data base are given in Table 2.

# Chapter 1

## Basic Concepts of Quantum Mechanics

### 1.1 Hamiltonian Formulation

In classical mechanics, the trajectory of a particle could be described by solving Newton's equation:

$$\frac{md^2\vec{r}}{dt^2} = \vec{F} \quad (1.1)$$

In principle, the equation is completely deterministic: "sufficiently" accurate knowledge of the initial conditions allow one to integrate the equations as far into the future as one likes. In fact, the question of what is sufficient accuracy may be fundamental. Further, many systems which obey apparently simple equations of motion quickly assume chaotic behavior. Thus the notion of classical mechanics as being completely deterministic, as opposed to the random, probabilistic nature of quantum mechanics, is misleading. The vast majority of systems turn out to be chaotic, where an initial observational error grows exponentially, rendering determinism meaningless on short time scales. For more on this subject, see Ford, *Physics Today*, **36**, 40 (April 1983) and Prigogine, *From Being to Becoming* 1980, Freeman. For our purposes we merely note that quantum mechanics also assumes a certain determinism, as we shall see.

To make the transition to quantum mechanics, it is useful to recast Newton's law into the Hamiltonian formulation. We imagine a system of particles with coordinates,  $q_1, q_2 \dots q_s$ , and momenta  $p_1, p_2 \dots p_s$ . The Hamiltonian is an ordinary function which describes the total energy of the system. We can break it into two parts, consisting of kinetic energy and potential energy:

$$H = T(p_1, p_2, \dots p_s) + V(q_1, q_2, \dots q_s). \quad (1.2)$$

Note that  $T$  is a function only of momenta, and  $V$  is a function only of coordinates.  $V$  may also be an explicit function of time, but we shall not consider this class of functions at present. The usual laws of motion can be derived from the Hamiltonian:

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}; \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i} \quad (i = 1, 2, \dots s). \quad (1.3)$$

To make this more explicit, substitute  $p = mv$ ,  $T = p^2/2m$ ,  $q = x$ . Then

$$\begin{aligned}\frac{dx}{dt} &= \frac{\partial(p^2/2m)}{\partial p} = \frac{p}{m} = v; \\ \frac{dmv}{dt} &= -\frac{\partial V(x)}{\partial x} \text{ or, } m \frac{d^2x}{dt^2} = F.\end{aligned}$$

Now consider any arbitrary function of the coordinates,  $f(q_1 \dots q_s, p_1 \dots p_s)$ . The time evolution of  $f$  is given by

$$\begin{aligned}\frac{df}{dt} &= \sum_{i=1}^s \left( \frac{\partial f}{\partial q_i} \frac{dq_i}{dt} + \frac{\partial f}{\partial p_i} \frac{dp_i}{dt} \right) \\ &= \sum_{i=1}^s \left( \frac{\partial f}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i} \frac{\partial H}{\partial q_i} \right) \text{ by use of (1.3)} \\ &\equiv [f, H],\end{aligned}\tag{1.4}$$

which is called the Poisson bracket of  $f$  with  $H$ . The function  $f$  will be a constant of the motion if  $[f, H] = 0$ . Obviously  $[H, H] = 0$ , representing conservation of energy.

## 1.2 Schrodinger Equation

The observation that atoms emit light only at certain frequencies led to the Bohr model in which atoms have discrete energy levels. Since the classical Hamiltonian is a continuous function of coordinates and momenta, it cannot adequately represent such systems. We can, however, make use of the Hamiltonian if we view it as an operator and the energy states as eigenvalues. Then we must solve the eigenvalue problem,

$$H\psi = E\psi\tag{1.5}$$

where  $E$  is simply a number and  $\psi$  is an eigenvector, or eigenfunction.

How do we transform a classical function to an operator? We must have rules to change from classical variables to operators. The rule we use is

$$q_c \rightarrow q \quad p_c \rightarrow \frac{\hbar}{i} \frac{\partial}{\partial q}\tag{1.6}$$

where we use the subscript  $c$  to indicate classical variables. This rule can be *motivated* in various ways, but it cannot be *derived*. The only justification is that it leads to equations which are useful in predicting the results of experiments on atoms and similar systems. Notice that the coordinate and momentum operators just defined do not commute:

$$[q, p] \equiv q \frac{\hbar}{i} \frac{\partial}{\partial q} - \frac{\hbar}{i} \frac{\partial}{\partial q} q \neq 0,\tag{1.7}$$

where the brackets are analogous to the Poisson brackets, but are now called the commutator. Variables whose operators do not commute cannot be measured simultaneously to arbitrarily high precision. Hence, the Heisenberg uncertainty relations:

$$\Delta q \Delta p \geq \hbar\tag{1.8}$$

If we also posit that the operator for the Hamiltonian is

$$H = i\hbar \frac{\partial}{\partial t}, \quad (1.9)$$

we get the time dependent Schrodinger equation,

$$i\hbar \frac{\partial}{\partial t} \Psi = H\Psi \quad (1.10)$$

Since we can write  $H = \sum_i (p_i^2/2m) + V(q_1, \dots, q_s, t)$ , we can express (1.10) more explicitly:

$$i\hbar \frac{\partial}{\partial t} \Psi = -\frac{\hbar^2}{2m} \sum_i \frac{\partial^2}{\partial q_i^2} \Psi + V(q_1, \dots, q_s, t)\Psi. \quad (1.11)$$

Thus we have an equation for the time evolution of a function  $\Psi$ , which is commonly called the wave function. Note that this equation is deterministic; probabilities enter when we try to interpret  $\Psi$  in terms of observable quantities. Solution of equation (1.10) in quantum mechanics replaces the solution of equations (1.3) in classical mechanics. The wave function is all you know and all you need to know, to paraphrase Keats. A formal solution to 1.10 is given by

$$\Psi(t) = e^{-iHt/\hbar} \Psi(t=0). \quad (1.12)$$

The unitary operator  $e^{-iHt/\hbar}$  gives the time evolution of the wave function. A unitary operator ( $U$ ) is characterized by its inverse ( $U^{-1}$ ) equaling its adjoint ( $U^\dagger$ ); thus  $UU^{-1} = UU^\dagger = 1$ . An analogous equation exists in classical mechanics, where  $H/\hbar$  is replaced by something called the Liouville operator (Prigogine, pp. 58).

Now suppose  $H$  does not depend explicitly on time. We can try to separate  $\Psi$  into functions of coordinates and time. For simplicity consider only a single coordinate,  $x$ . So  $\Psi(t) \rightarrow \Psi(x, t) = \psi(x)\varphi(t)$ . If  $\psi(x)$  are eigenfunctions of  $H$ , equation (1.5) allows us to change  $H$  into  $E$ , leading to

$$\Psi(x, t) = e^{-iEt/\hbar} \psi(x). \quad (1.13)$$

Thus if  $H$  does not depend explicitly on time, the solution to the time *dependent* Schrodinger equation is trivially reduced to solving the time *independent* equation. The only time dependence is an oscillatory phase factor which makes no difference when we compute observables. It is often said that the wave function simply “rotates” in Hilbert space, analogous to a vector rotating but maintaining the same length. It is sometimes useful to transform to this rotating frame of reference – this corresponds to the Heisenberg representation, as opposed to the Schrodinger representation, which we are using.

Clearly, time independent potentials are going to be desirable. We can generally take  $V$  to be time independent if we are considering the energy level structure of an isolated system. Thus we can generally study the energy levels and thus spectra (in the sense of what frequencies are possible) of atoms, molecules, etc., in the context of the time independent Schrodinger equation (1.5). To understand the transitions from one state to another (and hence the *strength* of transitions), we must deal with the time dependent form, equation (1.10). We should recognize however that using the time independent form is a bit of a cop-out: excited states are usually not infinitely long-lived and thus not strictly time independent. The key is that for systems we will consider, the life is long enough that the wave function “rotates” (in the sense described above) many times before other, non-periodic time dependences become important. Thus we can treat those other time dependences

in approximations that consider explicitly time variable potentials to be perturbations, which are small compared to the other relevant potentials. This gets us close to the nasty question of how a system that can only exist in discrete energy states ever makes a transition between them.

This issue is related to the fact that time is *not* an operator – it is a classical quantity, just as in classical mechanics. Thus it flows continuously and is not quantized. Quantum mechanics is also time reversible. These issues lie at the heart of attempts to modify quantum mechanics. A relativistic quantum mechanics can be developed, but quantum mechanics and relativity are still not fully integrated. Prigogine argues that time must be made an operator in order to relate the physics of “being” (mechanics) which the physics of “becoming” (irreversible processes, the second law of thermodynamics, etc.). Ford (cited above) argues that the notion of a continuous variable implies infinite measurement precision, which is impossible. While quantum mechanics does not allow arbitrary precision in measuring positions and momenta simultaneously, it does not rule out and, Ford argues, depends upon the assumption of infinite precision in measuring one quantity. Ford argues that ultimately, all physical variables will have to be quantized. We know that time is likely to be quantized on some level – the Planck time, if no other. The Planck time is  $\sqrt{\hbar G/c^5}$  where  $G$  is the gravitational constant; on time scales less than the Planck time ( $10^{-43}s$ ), time has no meaning. There may be more restrictive constraints as well.

While all these considerations indicate that quantum mechanics is not a complete theory, it has been remarkably successful in the problems to which we shall apply it – those of atoms, molecules, and solids. In fact, experiments have validated the predictions of quantum mechanics and ruled out those of its competitor, hidden variables theory. Not only hidden variables theories, but all theories based on naive realism are apparently ruled out. Thus one of the following must be untrue:

1. atoms, photons, etc. have objective reality (they exist whether or not we observe them).
2. Events separated by space-like separations cannot be causally related.
3. Inductive logic works.

Take your pick. For a review of this fascinating subject, I recommend the following works:

Bernard d’Espagnat, “The Quantum Theory and Reality,” *Scientific American*, Nov. 1979;  
 N.D. Mermin, *American Journal of Physics*, **49**, 940, Oct. 1981;  
 and  
 J.F. Clauser and A. Shimony, *Reports on Progress in Physics*, **41**, 1881 (1978),

in order of increasing detail.

### 1.3 Operating with Quantum Mechanics

We have argued that classical variables must be replaced by quantum mechanical operators and that the solution of the Schrodinger equation for the wave function gives us all the available information. How does this actually work? The wave function is not observable; indeed being a complex function, it cannot be said to have a real existence. If we take its “length” in the vector picture, we obtain a real quantity, the probability density:

$$\rho(q, t) \equiv \Psi^*(q, t)\Psi(q, t), \quad (1.14)$$

where we use  $q$  as a shorthand for all coordinates and  $\Psi^*$  is the complex conjugate of  $\Psi$ . Note that for Hamiltonians that do not depend explicitly on time, equation (1.13) shows that the probability density is independent of time:

$$\rho(q, t) = e^{+iEt/\hbar} \psi^*(q) e^{-iEt/\hbar} \psi(q) = \psi^*(q) \psi(q). \quad (1.15)$$

Thus the probability of finding the system in a given configuration in a steady state is given by the modulus of the eigenfunction for that state.

For simplicity consider a single particle with position  $\vec{r}$ . Then  $\rho(\vec{r}, t) d\vec{r}$  gives the probability that the particle will be found in a volume element  $d\vec{r}$  centered on  $\vec{r}$  at the time  $t$ . Since the particle must be somewhere we can require that the total probability is one:

$$\int \rho(\vec{r}, t) d\vec{r} = \int \Psi^*(\vec{r}, t) \Psi(\vec{r}, t) d\vec{r} = 1. \quad (1.16)$$

Wave functions satisfying (1.16) are normalized.

Since we cannot describe a classical trajectory in quantum mechanics, we can only determine the probabilities that the particle lies at various places and the average result of many measurements of its location. The latter is called the expectation value and is calculated from,

$$\langle A_c \rangle \equiv \int \Psi^*(\vec{r}, t) A \Psi(\vec{r}, t) d\vec{r}, \quad (1.17)$$

where  $A$  is the operator corresponding to the classical variable  $A_c$ . For example, the position of the particle has an operator which is simply the classical variable:

$$\langle \vec{r}_c \rangle = \int \Psi^*(\vec{r}, t) \vec{r} \Psi(\vec{r}, t) d\vec{r}, \quad (1.18)$$

which is readily interpreted as an average of  $\vec{r}$ , weighted by the probability that the particle is at  $\vec{r}$ . Things get a bit more interesting if we ask about the expectation value of the momentum,

$$\begin{aligned} \langle \vec{p}_c \rangle &= \int \Psi^*(\vec{r}, t) \vec{p} \Psi(\vec{r}, t) d\vec{r} \\ &= \int \Psi^*(\vec{r}, t) \frac{\hbar}{i} \vec{\nabla} \Psi(\vec{r}, t) d\vec{r}, \\ \text{where } \vec{\nabla} &= \hat{x} \frac{\partial}{\partial x} + \hat{y} \frac{\partial}{\partial y} + \hat{z} \frac{\partial}{\partial z}. \end{aligned} \quad (1.19)$$

Thus we do not simply average over  $p_c$ , but instead we use the corresponding operator. In general, the expectation value of any “observable” (which corresponds to a classical variable) can be obtained from the prescription (1.17). Suppose we have a system in a given state, so that equation (1.15) applies; then the expectation values will be time independent. A further simplification occurs if the operator commutes with the Hamiltonian, since commuting operators can have simultaneous eigenfunctions. The commutator of two operators is denoted

$$[A, B] \equiv AB - BA. \quad (1.20)$$

If two operators commute, their commutator is zero. For example the position, being represented by a classical variable, does not commute with the Hamiltonian, which contains partial derivatives. Under some conditions, notably those with no potential energy, the momentum may commute with

the Hamiltonian. Then the eigenfunctions of the Hamiltonian will also be eigenfunctions of the momentum. In this case, the momentum will be a constant of the motion; every measurement of momentum will produce the same answer. These remarks have applied to the Schrodinger representation; if we transformed to the Heisenberg representation, the position rather than the momentum would be the operator. The eigenvalues of operators that commute with the Hamiltonian are called “good quantum numbers” and are often used to label states in atoms and molecules.

## 1.4 Dirac Notation

It can become very tedious to write out the integrations and wave functions explicitly. Furthermore, the wave functions are merely conventional representations for more abstract concepts and can in fact be done away with entirely, as in Heisenberg’s matrix mechanics. We can use the more abstract concept of a state, which can be represented by the wave function. The state can be thought of as being a vector in Hilbert space; then the integrals we have been writing can be seen as inner products in this space. To use those notions, we will work with a notation introduced by Dirac.

$$\begin{aligned} \langle \psi | &= \text{bra;} \\ |\psi\rangle &= \text{ket;} \\ \langle \psi_n | \psi_m \rangle &= \text{braket} = \int \psi_n^*(q) \psi_m(q) dq; \end{aligned} \tag{1.21}$$

and  $\langle \psi_n | A\psi_n \rangle$  is the expectation value of  $A$ .

We will use this notation to establish a few important facts. Just as we can represent an arbitrary vector by a suitable linear combination of basic vectors, we can represent an arbitrary state  $|\psi\rangle$  by a suitable combination of basis states (eigenfunctions):

$$|\psi\rangle = \sum_n c_n |\psi_n\rangle, \tag{1.22}$$

where  $c_n = \langle \psi_n | \psi \rangle$ .

It is useful to use an orthonormal basis set:

$$\langle \psi_n | \psi_m \rangle = \delta_{nm}, \quad \delta_{nm} = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{if } n \neq m \end{cases} \tag{1.23}$$

The operators corresponding to observables turn out to be Hermitian (or self-adjoint)

$$A = A^\dagger, \tag{1.24}$$

where  $A^\dagger$  is defined by

$$\langle \psi_n | A\psi_m \rangle = \langle A^\dagger \psi_n | \psi_m \rangle. \tag{1.25}$$

Hermitian operators have some nice properties: their eigenvalues are real; the eigenfunctions corresponding to distinct eigenvalues are orthogonal; and the eigenfunctions form a complete basis set. Thus we are assured that we can use (1.22) to represent *any* state. This *completeness relation* can be expressed compactly by

$$1 = \sum_n |\psi_n\rangle \langle \psi_n| \tag{1.26}$$

since  $|\psi\rangle = \sum_n |\psi_n\rangle\langle\psi_n| \psi\rangle$  then holds for *any*  $|\psi\rangle$ . Note that if we want to represent a time variable  $\Psi(t)$  we must use

$$|\Psi(t)\rangle = \sum_n C_n |\psi_n\rangle e^{-iE_n t/\hbar}, \quad (1.27)$$

where  $E_n$  is the energy of state  $|\psi_n\rangle$ .

While we have used the notation  $|\psi_n\rangle$  to maintain contact with earlier notation, it is clear that all we need is a label for the state; thus  $|n\rangle$  would do just as well. In more complex situations, a generalized quantum number, or a list of quantum numbers specifying the state would be used. Also note that we have assumed all states are nondegenerate – i.e., that only one state corresponds to a given energy,  $E_n$ . The more general, degenerate, situation can also be dealt with at the cost of decreased clarity.

Finally, we introduce another concept which will be useful in discussing transitions, the matrix element. The matrix element is similar to the expectation value,  $\langle\psi_n | A\psi_n\rangle$ , except that the states differ:

$$A_{mn} \equiv \langle\psi_n | A\psi_m\rangle \equiv \langle\psi_n | A | \psi_m\rangle. \quad (1.28)$$

Note that  $A_{mn}$  is the opposite of the usual convention ( $A_{nm}$ ) but will be more consistent with astronomical applications.  $|A_{mn}|^2$  will turn out to enter the expression for the probability of a transition from  $m$  to  $n$ .

## 1.5 Illustrative Solutions

After all this abstraction, some concrete examples will be helpful. These are important because (1) there are almost no problems with exact solutions in quantum mechanics and (2) these idealized problems are good approximations to real problems we will encounter.

Let us consider first a single, structureless particle moving in a force-free environment – the “free particle.” In this case the potential energy is a constant, which we can set to zero. The Schrodinger equation then becomes

$$i\hbar \frac{\partial}{\partial t} \Psi = -\frac{\hbar^2}{2m} \nabla^2 \Psi, \quad (1.29)$$

where

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}, \quad (1.30)$$

This equation is separable so we can, without loss of generality, restrict ourselves to a single dimension:

$$i\hbar \frac{\partial}{\partial t} \Psi = \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \Psi. \quad (1.31)$$

Substitution will verify that  $\Psi(x, t) = e^{i(kx - \omega t)}$  is a solution, if  $\omega = E/\hbar$  and  $k = p/\hbar$ , since the equation becomes

$$\hbar\omega\Psi(x, t) = \frac{p^2}{2m}\Psi(x, t), \quad (1.32)$$

or

$$E = \frac{p^2}{2m}.$$

So the free particle wave functions are plane waves. Removing the time variable part,  $e^{-iEt/\hbar}$ , as discussed above, we are left with the eigenfunctions  $e^{ikx}$ . In what sense do these correspond to our classical notion of a particle moving freely with constant momentum? Let us compute the expectation value of  $p$ .

$$\begin{aligned}\langle p \rangle &= \langle \psi | p\psi \rangle \\ &= \int_{-\infty}^{\infty} e^{-ikx} \frac{\hbar}{i} \frac{\partial}{\partial x} e^{ikx} dx \\ &= \hbar k \int_{-\infty}^{\infty} e^{-ikx} e^{ikx} dx.\end{aligned}\tag{1.33}$$

If we neglect the ugly fact that the integral is infinite and notice that it is also  $\int_{-\infty}^{\infty} \psi^* \psi dx$ , and recall that we argued that the total probability is 1, we can pretend that  $\langle p \rangle = \hbar k$  as expected. In fact free particle wave functions can never be properly normalized since they extend, with finite amplitude, to infinity; nonetheless they are extremely useful. In particular, we can use a suitable linear combination to construct a wave packet. Why would we do this? Consider the notion of a particle. Our free particle wave functions extend with equal amplitude to  $\pm\infty$ . Our “particle” is not localized at all. If we use a linear combination of free particle wave functions of *similar*  $k$ , we know by Fourier analysis that we can make a peaked function. By increasing our uncertainty in  $p$ , we can decrease our uncertainty in  $x$ . The wave packet corresponds to our notion of a particle. We can also circumvent the normalization problems of free particles, by considering the particles to be confined within a certain region – a box. Thus we solve the problem of the “particle in a box”. This stratagem allows us to normalize the wave functions by dividing by the length of the box. While it might seem that we now have an arbitrary constant, in fact the length of the box always cancels out in any calculation of relevance. Thus we can always let the length of the box go to infinity. This called box normalization.

Since some interesting things can be illustrated by following this procedure through, let us consider the “particle in a box” problem. We consider a box with infinitely high walls – the infinite square well potential:

$$V(x) = \begin{cases} \infty & \text{for } x < -\frac{a}{2} \text{ or } x > \frac{a}{2} \\ 0 & \text{for } -\frac{a}{2} < x < \frac{a}{2} \end{cases}\tag{1.34}$$

Since we must have zero probability of finding the particle outside the box, the traveling wave solutions of the free particle are not satisfactory. To satisfy the boundary conditions,  $\psi(a/2) = \psi(-a/2) = 0$ , we can take linear combinations of the traveling waves:

$$\psi(x) = \sum_n c_n e^{ik_n x}.\tag{1.35}$$

In particular, if we use two terms ( $k_n$  and  $-k_n$ ), we can get two standing wave solutions

$$\sin k_n x = \frac{e^{ik_n x} - e^{-ik_n x}}{2i}; \cos k_n x = \frac{e^{ik_n x} + e^{-ik_n x}}{2}.\tag{1.36}$$

These can be made to satisfy the boundary conditions if

$$k_n = \frac{n\pi}{a} \quad n = 1, 3, 5, \dots; \quad k_n = \frac{n\pi}{a} \quad n = 2, 4, 6, \dots \\ \text{for } \cos k_n x \quad \text{for } \sin k_n x. \quad (1.37)$$

Notice that unlike the free particle wave functions, for which any  $k$  is allowed, the particle in a box wave functions can only have certain choices of  $k_n$ , or momenta, or energy:

$$E_n = \frac{p^2}{2m} = \frac{\hbar^2 k_n^2}{2m} = \frac{\pi^2 \hbar^2 n^2}{2ma^2} \quad n = 1, 2, 3 \dots \quad (1.38)$$

Simply putting the free particle in a box has changed its spectrum of eigenvalues from continuous to discrete! How do we ever recover the free particle wave functions from box normalization? The key is that  $a$ , the box length, appears in the denominator; if we increase  $a$ , the energy levels get closer together. As  $a \rightarrow \infty$  the discrete spectrum approaches a continuous one, and the sines and cosines can always be added appropriately to get back the traveling wave functions. This sort of slippery stuff is commonly encountered in this business.

Note that the lowest energy state is  $E_1 = \pi^2 \hbar^2 / 2ma^2$  for the infinite square well. This is the zero point energy – the particle cannot have zero energy ( $n = 0$  would correspond to  $\sin 0 = 0$  and cannot describe a particle in the box). The zero point energy can be related to the uncertainty principle; since we know the particle is in the box,  $\Delta x = a$ , so  $\Delta p \sim \hbar/a$ ,  $\Delta E \sim \hbar^2 / 2ma^2$ .

The free particle wave functions have obvious applications. The infinite square well can of course be applied to a real box, or to a conduction electron in a metal block of length  $a$  (in principle a finite square well, but the infinite square well is a good approximation if electrons rarely escape the metal). There are several other illustrative solutions, which will be explored in homework problems. These include the finite square well potential, the harmonic oscillator, and the periodic potential. These are all one dimensional problems. We will also need to move to three dimensional problems which do not readily separate into three Cartesian coordinates. An example of these is rotation; the simplest problem of this type is the rigid rotor.

The rigid rotor is conceived as a dumbbell: two mass points connected by a rigid, massless bond. The dumbbell is able to rotate freely in space. We essentially have a highly restricted two body problem: the distance between the two bodies must remain constant. At first glance, this seems a trivial problem, and it is not clear how quantization will enter since the dumbbell is “free”. However the constrained separation makes the dumbbell a complex system which, in addition to moving freely through space, can have internal excitations (namely rotation about the center of mass). In addition to its direct applicability to the diatomic molecule, the rigid rotor solutions turn out to be relevant to all central force problems – those where the potential energy is a function only of the distance from a suitably chosen origin:  $V(\vec{r}) = V(r)$ . Therefore, let us consider the rigid rotor in some detail.

First we establish the notation.

$$\begin{aligned} M &= m_1 + m_2 \\ \mu &= \frac{m_1 m_2}{m_1 + m_2} \end{aligned} \quad (1.39)$$

We can always reduce a two body problem to a one-body problem by separating the motion into the motion of the center of mass and the motion of a mythical body of mass  $\mu$  relative to the center

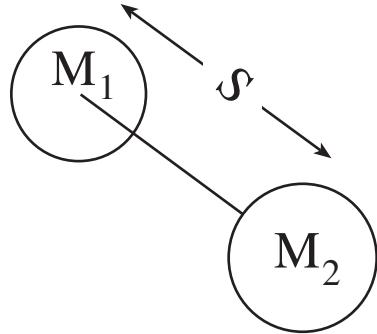


Figure 1.1: The rigid rotor as a dumbbell.

of mass. The center of mass motion is just that of a free particle, leaving us with the Schrodinger equation for a point mass  $\mu$  about the origin:

$$\frac{\hbar^2}{2\mu} \nabla^2 \Psi = -i\hbar \frac{\partial}{\partial t} \Psi, \quad (1.40)$$

and since the problem is time independent, we can make the usual  $\Psi = e^{-iEt/\hbar}\psi(\vec{r})$  (equation 1.12) separation to get

$$\frac{\hbar^2}{2\mu} \nabla^2 \psi(\vec{r}) + E\psi(\vec{r}) = 0, \quad (1.41)$$

where  $\vec{r}$  is the vector from the origin to the mass  $\mu$ . We want to use spherical polar coordinates as shown in the next figure.

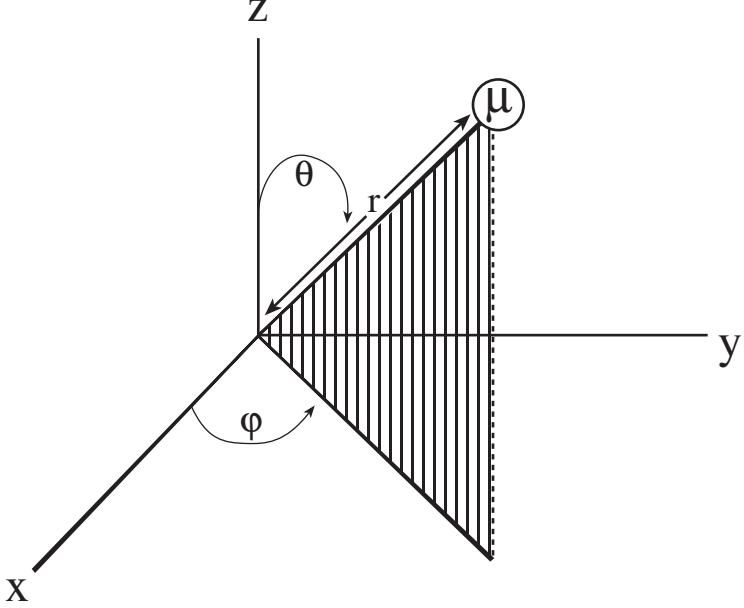


Figure 1.2: Spherical polar coordinates

The operator  $\nabla^2$  in these coordinates is

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \varphi^2}. \quad (1.42)$$

If we assume  $\psi(\vec{r}) = R(r)Y_{\ell m}(\theta, \varphi)$ , the problem separates (as it will for any central force problem) into separate equations for  $R(r)$  and  $Y_{\ell m}(\theta, \varphi)$ . The solution for  $R(r)$  is obvious for the rigid rotor – essentially a delta function for  $r$  equal to the distance ( $r_e$ ) of  $\mu$  from the center of the mass. So we only need to deal with the angular function. It is useful to define the operator  $L^2$ ,

$$L^2 = -\hbar^2 \left( \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \varphi^2} \right). \quad (1.43)$$

The angular part of the problem then reduces to

$$L^2 Y_{\ell m}(\theta, \varphi) = 2\mu r_e^2 E Y_{\ell m}(\theta, \varphi). \quad (1.44)$$

That is, we must solve the eigenvalue problem for the operator  $L^2$ . We have, of course, given away the game by our choice of notation.  $L^2$  is the operator for the square of total angular momenta, and  $Y_{\ell m}(\theta, \varphi)$  are spherical harmonics. We know that  $Y_{\ell m}(\theta, \varphi)$  are eigenfunctions for  $L^2$  with eigenvalues  $\hbar^2 \ell(\ell + 1)$  so  $2\mu r_e^2 E = \hbar^2 \ell(\ell + 1)$  or

$$E = \frac{\hbar^2 \ell(\ell + 1)}{2\mu r_e^2} = \frac{\hbar^2 \ell(\ell + 1)}{2I}, \quad (1.45)$$

where  $I$  = moment of inertia of mass  $\mu$  a distance  $r_e$  from the origin.

Note that  $L^2$  commutes with  $H$ , which is true for any central force problem, so the square of the angular momentum is a good quantum number. Indeed, the energy states correspond to different values of angular momentum, and we can label the eigenvalues of  $H$  by the eigenvalues of  $L^2$ , or more simply by  $\ell$ .

The other quantum number in the eigenfunctions is  $m$ . The energy does not depend on  $m$ , but the eigenfunctions do, so we have an example of degeneracy: more than one eigenfunction at a given energy. We can further separate  $Y_{\ell m}(\theta, \varphi)$  into functions only of  $\theta$  and  $\varphi$ :

$$Y_{\ell m}(\theta, \varphi) = \Theta(\theta)\Phi(\varphi). \quad (1.46)$$

The equation for  $\Phi(\varphi)$  is simple:

$$\frac{d^2}{d^2\varphi}\Phi(\varphi) = -m^2\Phi(\varphi), \quad (1.47)$$

which has the solution

$$\Phi_m(\varphi) = \frac{1}{\sqrt{2\pi}}e^{im\varphi} \quad m = 0, \pm 1, \pm 2, \dots \quad (1.48)$$

The solution for  $\Theta(\theta)$  is a bit messier and is not necessary: the solutions are suitably normalized associated Legendre functions, but it is always easier to deal with spherical harmonics directly. One lesson from the details is that  $\ell$  and  $m$  are related:  $m$  can take on the values  $-\ell, -\ell+1, \dots, 0, 1, \dots, \ell-1, \ell$ . It is clear that  $m\hbar$  is the projection of  $L$  on the  $z$  axis.

The spherical harmonics form an orthonormal and complete set and are the natural functions to use whenever central force problems are encountered. It is also useful to define operators for each component of angular momentum:

$$\begin{aligned} L_x &= -i\hbar \left( y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) \\ L_y &= -i\hbar \left( z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right) \\ L_z &= -i\hbar \left( x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right), \end{aligned} \quad (1.49)$$

which can be obtained by “quantizing”  $\vec{L} = \vec{r}\vec{x}\vec{p}$ . Transforming to spherical coordinates yields,

$$\begin{aligned} L_x &= i\hbar \left( \sin \varphi \frac{\partial}{\partial \theta} + \text{ctn} \theta \cos \varphi \frac{\partial}{\partial \varphi} \right) \\ L_y &= -i\hbar \left( \cos \varphi \frac{\partial}{\partial \theta} - \text{ctn} \theta \sin \varphi \frac{\partial}{\partial \varphi} \right) \\ L_z &= -i\hbar \frac{\partial}{\partial \varphi}. \end{aligned} \quad (1.50)$$

From these, one can verify the identification of  $L^2$  with the part of  $\nabla^2$  claimed in (1.42). Also useful are the raising and lowering operators,

$$L_{\pm} = L_x \pm iL_y. \quad (1.51)$$

The spherical harmonics are not eigenfunctions of  $L_{\pm}$ ; instead

$$L_{\pm} Y_{\ell m}(\theta, \varphi) = \hbar \sqrt{(\ell \mp m)(\ell \pm m + 1)} Y_{\ell m \pm 1}(\theta, \varphi). \quad (1.52)$$

So  $L_+$  raises the  $m$  quantum number by 1;  $L_-$  lowers it. The commutation relations are,

$$[L^2, L_i] = 0 \quad i = x, y, z, \quad (1.53)$$

but

$$[L_x, L_y] = i\hbar L_z, \quad [L_y, L_z] = i\hbar L_x, \quad [L_z, L_x] = i\hbar L_y. \quad (1.54)$$

So we can know the projection on only one axis at a time.

## 1.6 Time Independent Perturbation Theory

Since very few problems can be solved exactly in quantum mechanics, we need approximation methods. One very useful one is to consider a small change, or perturbation, to a problem whose solution is known. Since we are interested here in time independent perturbations, we need only deal with the time independent Schrodinger equation. Suppose the Hamiltonian can be divided into a part  $H^{(0)}$  for which we know the solution and a part  $H^{(1)}$  which is a small perturbation on  $H^{(0)}$ .

Then

$$H^{(0)}\psi_n^{(0)} = E_n^{(0)}\psi_n^{(0)} \quad (1.55)$$

where  $E_n^{(0)}$  and  $\psi_n^{(0)}$  are the eigenvalues and eigenfunctions of the unperturbed problem. These form a complete basis set ( $H^{(0)}$  is Hermitian) so we can expand  $\psi_n = \sum_k c_k \psi_k^{(0)}$  where  $\psi_n$  is the solution we seek:

$$H = H^{(0)} + H^{(1)}, \quad H\psi_n = E_n\psi_n \quad (1.56)$$

Since  $\psi_n \rightarrow \psi_n^{(0)}$  if  $H^{(1)} \rightarrow 0$ , there is a one-to-one correspondence between solutions to the unperturbed and perturbed problems. It is useful to introduce an expansion parameter  $\lambda$  which can go to zero, so  $H^{(1)} = \lambda H'$  where  $H'$  is then constant.

Assume that  $\psi_n \simeq \psi_n^{(0)} + \lambda\psi_n^{(1)}$ ;  $E_n = E_n^{(0)} + \lambda E_n^{(1)}$ . Then,

$$(H^{(0)} + \lambda H')(\psi_n^{(0)} + \lambda\psi_n^{(1)}) = (E_n^{(0)} + \lambda E_n^{(1)})(\psi_n^{(0)} + \lambda\psi_n^{(1)}). \quad (1.57)$$

Let us multiply this out, collect common powers of  $\lambda$ , and adopt Dirac notation:

$$\begin{aligned} & (H^{(0)} | \psi_n^{(0)} \rangle - E_n^{(0)} | \psi_n^{(0)} \rangle) \lambda^0 \\ & + (H' | \psi_n^{(0)} \rangle + H^{(0)} | \psi_n^{(1)} \rangle - E_n^{(1)} | \psi_n^{(0)} \rangle - E_n^{(0)} | \psi_n^{(1)} \rangle) \lambda^1 \\ & \quad + \text{terms of order } \lambda^2 = 0 \end{aligned} \quad (1.58)$$

The factor multiplying  $\lambda^0$  is zero (equation 1.55), and we drop  $\theta(\lambda^2)$  for first order perturbation theory. We are left with

$$H' | \psi_n^{(0)} \rangle + H^{(0)} | \psi_n^{(1)} \rangle - E_n^{(1)} | \psi_n^{(0)} \rangle - E_n^{(0)} | \psi_n^{(1)} \rangle = 0 \quad (1.59)$$

With a  $\langle \psi_n^{(0)} |$  on the left, we get

$$\begin{aligned} & \langle \psi_n^{(0)} | H' | \psi_n^{(0)} \rangle + \langle \psi_n^{(0)} | H^{(0)} | \psi_n^{(1)} \rangle - \langle \psi_n^{(0)} | E_n^{(1)} | \psi_n^{(0)} \rangle \\ & \quad - \langle \psi_n^{(0)} | E_n^{(0)} | \psi_n^{(1)} \rangle = 0, \end{aligned} \quad (1.60)$$

but  $H^{(0)}$  is Hermitian so (1.25) tells us that  $\langle \psi_n^{(0)} | H^{(0)} | \psi_n^{(1)} \rangle = \langle H^{(0)} \psi_n^{(0)} | \psi_n^{(1)} \rangle = E_n^{(0)} \langle \psi_n^{(0)} | \psi_n^{(1)} \rangle$  so we are left with

$$\langle \psi_n^{(0)} | H' | \psi_n^{(0)} \rangle - E_n^{(1)} \langle \psi_n^{(0)} | \psi_n^{(0)} \rangle,$$

or

$$E_n^{(1)} = \frac{\langle \psi_n^{(0)} | H' | \psi_n^{(0)} \rangle}{\langle \psi_n^{(0)} | \psi_n^{(0)} \rangle}. \quad (1.61)$$

If the  $|\psi_n^{(0)}\rangle$  are properly normalized,

$$\begin{aligned} E_n &= E_n^{(0)} + \lambda \langle \psi_n^{(0)} | H' | \psi_n^{(0)} \rangle \\ &= E_n^{(0)} + \langle \psi_n^{(0)} | H^{(1)} | \psi_n^{(0)} \rangle \end{aligned} \quad (1.62)$$

That is, the first order correction to the energy eigenvalue is the matrix element of the perturbing Hamiltonian. This is usually what we want, but if we want to know the perturbed wave function as well, we must write

$$|\psi_n^{(1)}\rangle = \sum_{m=1}^{\infty} a_{nm}^{(1)} |\psi_m^{(0)}\rangle.$$

A similar procedure, but using  $\langle \psi_k^{(0)} |$  with  $k \neq n$  gives us an expression for  $a_{nk}^{(1)}$ :

$$a_{nk}^{(1)} = \frac{\langle \psi_k^{(0)} | H^{(1)} | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_k^{(0)}} \quad n \neq k, \quad (1.63)$$

so

$$|\psi_n\rangle \approx |\psi_n^{(0)}\rangle + \sum_{k \neq n}^{\infty} \frac{\langle \psi_k^{(0)} | H^{(1)} | \psi_n^{(0)} \rangle}{E_n^{(0)} - E_k^{(0)}} |\psi_k^{(0)}\rangle. \quad (1.64)$$

Thus, a suitable combination of unperturbed wave functions gets mixed in to each initial wave function. The denominator indicates that states nearby in energy will have much bigger coefficients, so often only a few terms are necessary. Note however that if the states are degenerate, the coefficient blows up. For degenerate states, another method is necessary.

Before exploring this method, we can note that  $\langle \psi_n^{(0)} | H^{(1)} | \psi_n^{(0)} \rangle$  may be zero (typically for reasons of symmetry) and higher order terms must be retained. The resulting expression for the perturbed energy to second order is

$$\begin{aligned} E_n &= E_n^{(0)} + \langle \psi_n^{(0)} | H^{(1)} | \psi_n^{(0)} \rangle \\ &\quad + \sum_{k \neq n}^{\infty} \frac{|\langle \psi_k^{(0)} | H^{(1)} | \psi_n^{(0)} \rangle|^2}{E_n^{(0)} - E_k^{(0)}} \end{aligned} \quad (1.65)$$

## 1.7 Variational Method

We are often interested in the effect of perturbations on degenerate states, precisely because perturbations often result in splitting the degeneracy. However, the effect may be large, violating the conditions of perturbation theory. Fortunately, there is a solution to this problem, based on a variational method. The idea here is to use a trial wave function  $\varphi$ . Note that, if  $\psi_n$  is the actual, exact solution,

$$E_n = \frac{\langle \psi_n | H | \psi_n \rangle}{\langle \psi_n | \psi_n \rangle}, \quad (1.66)$$

so we define the functional,  $E'[\varphi]$  by

$$E'[\varphi] \equiv \frac{\langle \varphi | H | \varphi \rangle}{\langle \varphi | \varphi \rangle}. \quad (1.67)$$

It turns out that the energy is *stationary* with respect to small variations about the true wave function. That is, if  $\varphi$  differs from  $\psi_n$  by a term of order  $\alpha$ , then  $E'[\varphi]$  differs from  $E_n$  by a term of order  $\alpha^2$ . It is also true that for ground state energies,  $E'[\varphi]$  is always greater than or equal to the true ground state energy,  $E_1$ ; consequently, minimizing  $E'[\varphi]$  with respect to arbitrary variational parameters gives the best estimate for  $E_1$ . To prove this, we expand  $\varphi$  using the exact eigenfunctions:

$$\varphi = \sum_{k=1}^{\infty} a_k \psi_k. \quad (1.68)$$

Then

$$E'[\varphi] = \frac{\langle \varphi | H | \sum_{k=1}^{\infty} a_k \psi_k \rangle}{\langle \varphi | \varphi \rangle} = \frac{\sum_{k=1}^{\infty} |a_k|^2 E_k}{\sum_{k=1}^{\infty} |a_k|^2}. \quad (1.69)$$

But since  $E_k \geq E_1$  ( $E_1$  is the *ground state*),

$$E'[\varphi] \geq \frac{\sum_{k=1}^{\infty} |a_k|^2 E_1}{\sum_{k=1}^{\infty} |a_k|^2} \geq E_1 \quad (1.70)$$

This is all very nice, but how do we actually *use* this method, since we *don't* know  $\psi_n$ . The basic approach is to cook up a form for the trial functions which is a reasonable approximation and incorporate some variational parameters. A particularly useful form occurs when the variational parameters enter only linearly. The linear, or Rayleigh-Ritz, variational principle results. Suppose

$$\varphi = \sum_{i=1}^{\infty} c_i u_i, \quad (1.71)$$

where  $u_i$  satisfy the appropriate boundary conditions. The number of basis functions used ( $N$ ) depends on the problem. Then

$$E' \equiv E'[\varphi] = \frac{\langle \varphi | H \varphi \rangle}{\langle \varphi | \varphi \rangle} \Rightarrow E' \langle \varphi | \varphi \rangle = \langle \varphi | H | \varphi \rangle,$$

or

$$E' \sum_{j=1}^N \sum_{i=1}^N c_j^* c_i S_{ji} = \sum_{j=1}^N \sum_{i=1}^N c_j^* c_i H_{ji}, \quad (1.72)$$

where  $S_{ji} \equiv \langle u_j | u_i \rangle$  is the overlap matrix and  $H_{ji} = \langle u_j | H | u_i \rangle$  is the Hamiltonian matrix. Then

$$\sum_{j=1}^N \sum_{i=1}^N c_j^* c_i (H_{ji} - E' S_{ji}) = 0 \quad (1.73)$$

to achieve an estimate of the ground state energy, we minimize  $E'$  with respect to each  $c_j$ . Applying  $\frac{\partial}{\partial c_i}$  and requiring  $\partial E'/\partial c_i = 0$  yields  $N$  linear homogeneous equations:

$$\sum_{j=1}^N c_j^* (H_{ji} - E' S_{ji}) = 0 \quad i = 1, \dots, N \quad (1.74)$$

or

$$\sum_{j=1}^N c_j (H_{ij} - E' S_{ij}) = 0$$

using  $H_{ij} = H_{ji}^*$  and  $S_{ij} = S_{ji}^*$ . These are called the secular equations. They have non-trivial solutions for the  $c_j$  if and only if the determinant

$$| H_{ij} - E' S_{ij} | = 0 \quad (1.75)$$

By explicitly writing out the determinant, we get an  $N^{th}$  order algebraic equation, which has  $N$  roots,  $E'_n$ . We simply take the lowest of these as our upper bound on  $E_1$ . Substituting  $E'_n$  back into the secular equations

$$\sum_{j=1}^N c_j^{(n)} (H_{ij} - E'_n S_{ij}) = 0 \quad (1.76)$$

and solving for  $c_j^{(n)}$ . Together with the normalization condition, this yields all the coefficients. It also turns out that perturbation theory is a special case of the variational method, but the variational method does *not* require that the perturbation be small. In particular we can apply it to degenerate states. Consider a hypothetical system with two degenerate states. That is, the wave functions  $\psi_1^{(0)}$  and  $\psi_2^{(0)}$  differ, but  $E_1^{(0)} = E_2^{(0)}$ . Certain perturbations will split this degeneracy. Let's use  $\psi_1^{(0)}$  and  $\psi_2^{(0)}$  as our basis set.

$$\varphi = c_1 \psi_1^{(0)} + c_2 \psi_2^{(0)} \quad (1.77)$$

Then

$$\begin{vmatrix} H_{11} - E' & H_{12} \\ H_{21} & H_{22} - E' \end{vmatrix} = 0 \quad (1.78)$$

is the determinant equation.

$$\begin{aligned} (H_{11} - E')(H_{22} - E') - H_{21}H_{12} &= 0 \\ E'^2 - (H_{11} + H_{22})E' + H_{11}H_{22} - H_{21}H_{12} &= 0 \end{aligned}$$

or

$$\begin{aligned} ax^2 + bx + c = 0 \quad x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \\ E' = \frac{H_{11} + H_{22} \pm \sqrt{(H_{11} + H_{22})^2 - 4H_{11}H_{22} + 4H_{21}H_{12}}}{2} \\ E' = \frac{H_{11} + H_{22}}{2} \pm \frac{1}{2} \sqrt{(H_{11} - H_{22})^2 + 4H_{12}H_{21}} \end{aligned} \quad (1.79)$$

Call  $E'_1$  the solution with the minus sign and  $E'_2$  the solution with the plus sign. Substituting these into the secular equations yields

$$c_1^{(n)} (H_{11} - E'_n) + c_2^{(n)} H_{12} = 0 \quad n = 1, 2$$

$$c_1^{(n)} H_{21} + c_2^{(n)} (H_{22} - E'_n) = 0 \quad n = 1, 2 \quad (1.80)$$

because  $S_{ij} = \delta_{ij}$  for our orthonormal basis,  $\psi_1^{(0)}$  and  $\psi_2^{(0)}$ . Thus, we get values for  $c_1$  and  $c_2$  for each  $n = 1, 2$ . The result is two wave functions [eq. (1.80) gives us  $\frac{c_2^{(1)}}{c_1^{(1)}} = \frac{-H_{21}}{H_{22} - E'_1}$ ; that is, ratios only] Then one uses the normalization condition.

$$\begin{aligned} \psi'_1 &= N_1 [\psi_1^{(0)} + \frac{H_{21}}{E'_1 - H_{22}} \psi_2^{(0)}] \quad \text{for } E'_1 \\ \psi'_2 &= N_2 [\psi_2^{(0)} + \frac{H_{12}}{E'_2 - H_{11}} \psi_1^{(0)}] \quad \text{for } E'_2 \end{aligned} \quad (1.81)$$

Where  $N_1$  and  $N_2$  are normalization constants. Now since  $\psi_1^{(0)}$  and  $\psi_2^{(0)}$  are eigenfunctions of the unperturbed system with  $H^{(0)}$ , if  $H = H^{(0)} + H^{(1)}$ , we have  $H_{ij} = E_i^{(0)}\delta_{ij} + \langle\psi_i^{(0)}|H^{(1)}|\psi_j^{(0)}\rangle$ , and the problem reduces to calculating the matrix elements of the perturbation  $H^{(1)}$ .

Note that  $E'_1 \neq E'_2$ ; in fact  $E'_1 < E'_2$  -the degeneracy is split.

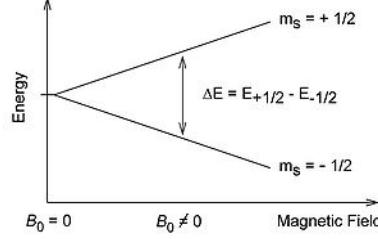


Figure 1.3: Here is example in which the perturbation is a magnetic field, which breaks the degeneracy in the azimuthal quantum number.

We can also include the effects of higher (non-degenerate) levels by using perturbation theory.

$$\begin{aligned} E_1 &\simeq E'_1 + \sum_{j \neq 1,2}^N \frac{|\langle\psi_j^{(0)}|H^{(1)}|\psi'_1\rangle|^2}{E_1^{(0)} - E_j^{(0)}} \\ E_2 &\simeq E'_2 + \sum_{j \neq 1,2}^N \frac{|\langle\psi_j^{(0)}|H^{(1)}|\psi'_2\rangle|^2}{E_2^{(0)} - E_j^{(0)}} \end{aligned} \quad (1.82)$$

# Chapter 2

## Radiative Transitions

### 2.1 Time Dependent Perturbation Theory

To deal with transitions from one state to another, we clearly have to involve the time dependent Schrodinger's equation. The easiest way to do this is to use the time - independent solutions as unperturbed eigenstates and consider the transition as being caused by a time dependent perturbation. Then the eigenfunctions of the unperturbed Hamiltonian form a complete set,  $|\psi\rangle$ , where

$$H^{(0)}|\psi\rangle = E|\psi\rangle \quad (2.1)$$

To make full use of the Dirac notation, let us assume that we can label these eigenfunctions by a single quantum number so we can write the  $j^{th}$  state as  $|j\rangle$ . Then, since  $e^{-i(E_j/\hbar)t}|j\rangle$  form a complete set (see pg. 25), we can expand the solution to the full Hamiltonian  $H = H^{(0)} + H^{(1)}(t)$  as

$$\Psi(\vec{r}, t) = \sum_{j=1}^{\infty} a_j(t) e^{-i(E_j/\hbar)t} |j\rangle \quad (2.2)$$

where  $|j\rangle$  carries the  $\vec{r}$  dependence, were we to write it out explicitly. substituting into the time dependent Schrodinger equation,

$$i\hbar \frac{\partial}{\partial t} \Psi(\vec{r}, t) = (H^{(0)} + H^{(1)}(t)) \Psi(\vec{r}, t) \quad (2.3)$$

we get

$$i\hbar \frac{\partial}{\partial t} \left( \sum_{j=1}^{\infty} a_j(t) e^{-i(E_j/\hbar)t} |j\rangle \right) = (H^{(0)} + H^{(1)}(t)) \sum_{j=1}^{\infty} a_j(t) e^{-i(E_j/\hbar)t} |j\rangle \quad (2.4)$$

which will have two terms on each side.

Now we want to "project out" the  $q^{th}$  component of the total wave function, so we "left multiply" by  $\langle q|e^{+i(E_q/\hbar)t}$  (recall this means multiply by complex conjugate and integrate). Then we use

equation 2.1 and the orthonormality of the  $|j\rangle$ 's to get rid of a term  $E_q a_q(t)$  that will appear on both sides of the equation. Then we obtain

$$\frac{d}{dt} a_q(t) = \frac{1}{i\hbar} \sum_{j=1}^{\infty} a_j(t) e^{-i(E_j - E_q)t/\hbar} H_{jq}^{(1)}(t) \quad (2.5)$$

where  $H_{jq}^{(1)}(t) \equiv \langle q | H^{(1)}(t) | j \rangle$  (eqn. 1.27). This is an infinite set of coupled equations which can formally be integrated to give

$$a_q(t) = \frac{-i}{\hbar} \sum_{j=1}^{\infty} \int_0^t a_j(t) e^{-i(E_j - E_q)t/\hbar} H_{jq}^{(1)}(t) dt + c_q \quad (2.6)$$

where  $c_q$  is a constant of integration corresponding to  $a_q(t)$  at  $t = 0$ .

Suppose we start with the system in state  $|k\rangle$  at  $t = 0$  and turn on the perturbation. Then  $a_q(0) = c_q = \delta_{qk}$ , and since the perturbation is small, the changes will be slow and we can drop all terms in the summation except the one with  $j = k$ . Then,

$$a_q(t) = \delta_{qk} - \frac{i}{\hbar} \int_0^t e^{-i(E_k - E_q)t/\hbar} H_{kq}^{(1)}(t) dt \quad (2.7)$$

We can solve for the expansion coefficients  $a_q(t)$  to first order and use the resulting wave function to improve the approximation if needed. Physically, the time evolution of  $a_q$  gives the rate at which the probability of finding the system in state  $q$  increases or decreases. To see this, project the  $q^{th}$  state onto the time dependent wave function.

$$\begin{aligned} P_q(t) &= |\langle q | e^{iE_q t/\hbar} \sum_{j=1}^{\infty} a_j(t) e^{-iE_j t/\hbar} | j \rangle|^2 \\ &= \left| \sum_{j=1}^{\infty} a_j(t) e^{i(E_q - E_j)t/\hbar} (\langle q | j \rangle = \delta_{qj}) \right|^2 = |a_q(t)|^2 \end{aligned} \quad (2.8)$$

Now consider a perturbation which is harmonic in time (e.g., a traveling electromagnetic wave) after switching on at time  $t = 0$ . We can represent it as,

$$H^{(1)}(\vec{r}, t) = H^+(\vec{r}) e^{i\omega t} + H^-(\vec{r}) e^{-i\omega t} \quad (t \geq 0) \quad (2.9)$$

Then

$$a_q(t) = \frac{-i}{\hbar} \int_0^t e^{-i(E_k - E_q)t/\hbar} H_{kq}^+ e^{i\omega t} dt - \frac{i}{\hbar} \int_0^t e^{-i(E_k - E_q)t/\hbar} H_{kq}^- e^{-i\omega t} dt \quad (2.10)$$

Where  $H_{kq}^{\pm} \equiv \langle q | H^{\pm}(\vec{r}) | k \rangle$ , and we integrate (2.10) to get

$$a_q(t) = \frac{H_{kq}^+}{(E_k - E_q - \hbar\omega)} [e^{-i(E_k - E_q - \hbar\omega)t/\hbar} - 1] + \frac{H_{kq}^-}{(E_k - E_q + \hbar\omega)} [e^{-i(E_k - E_q + \hbar\omega)t/\hbar} - 1] \quad (2.11)$$

The time evolution of  $a_q(t)$  depends on the frequency of the perturbation. As  $t$  increases, these functions become strongly peaked around  $\omega$  values which make the denominators zero. We can associate  $\hbar\omega$  with the photon energy  $E_{ph}$  and we expect significant contributions to  $a_q(t)$  only in the two cases,

$$\begin{aligned} 1. \quad \hbar\omega &= E_{ph} = E_k - E_q \quad (H^+) \quad \text{stimulated emission} \quad \downarrow_q^k \\ &\quad \text{or} \\ 2. \quad \hbar\omega &= E_{ph} = E_q - E_k \quad (H^-) \quad \text{stimulated absorption} \quad \overline{\downarrow}_k^q \end{aligned} \quad (2.12)$$

Using stimulated emission as an example, we find the probability for a transition from initial state  $k$  to state  $q$ :

$$P_{k \rightarrow q}(t) = |a_q(t)|^2 = \left| H_{kq}^+ \right|^2 \left| \frac{e^{-i(E_k - E_q - \hbar\omega)t/\hbar} - 1}{(E_k - E_q - \hbar\omega)} \right|^2 \quad (2.13)$$

Let  $\frac{E_k - E_q - \hbar\omega}{2\hbar} \equiv x$ ; then the factor on the right becomes

$$\begin{aligned} \frac{1}{\hbar^2} \left| \frac{e^{-2ixt} - 1}{2x} \right|^2 &= \frac{1}{\hbar^2} \left| \frac{e^{-ixt}(e^{-ixt} - e^{ixt})}{2x} \right|^2 = \left| -ie^{-ixt} \right|^2 \left| \frac{e^{ixt} - e^{-ixt}}{2ix} \right|^2 \frac{1}{\hbar^2} \\ &= \frac{1}{\hbar^2} \left( \frac{\sin xt}{x} \right)^2 \end{aligned} \quad (2.14)$$

This is a function that becomes increasingly peaked near  $x = 0$  as  $t \rightarrow \infty$ ; i.e., it looks like a Dirac delta function. We can change it to one as long as we watch the normalization. We imagine that we ultimately will want to integrate over a range of final state energies (transitions to a continuum) or photon frequencies ( $\omega$ ) or energies ( $\hbar\omega$ ). If we integrate over  $\omega$ , we will want  $\int_{-\infty}^{\infty} \delta(x)d\omega = 1$  and  $\int_{-\infty}^{\infty} \frac{\sin^2 xt}{x^2} d\omega = 2\pi t$  using a standard integral<sup>1</sup> since  $d\omega = 2dx$ . So the transition probability grows linearly with time and the rate (transition probability per unit time) is

$$R_{k \rightarrow q} = \frac{P_{k \rightarrow q}}{t} = \frac{2\pi}{\hbar^2} \left| H_{kq}^+ \right|^2 \delta \left( \frac{E_k - E_q}{\hbar} - \omega \right) \quad (2.15)$$

If we eventually want to integrate over final state (or photon) energies, we have

$$R_{k \rightarrow q} = \frac{2\pi}{\hbar} \left| H_{kq}^+ \right|^2 \delta(E_k - E_q - \hbar\omega) \quad (dE = \hbar d\omega) \quad (2.15b)$$

Of course, these blow up for a resonant condition, but we will always integrate over a distribution function. This result (Fermi's Golden Rule) is therefore sometimes written as

$$R_{k \rightarrow q} = \frac{2\pi}{\hbar} \left| H_{kq}^+ \right|^2 \rho_q(E) \quad (2.16)$$

where  $\rho_q(E)$  is the “density of states near  $q$ ”.

---

<sup>1</sup>See, e.g., Dwight, Tables of Integrals, pg 221

## 2.2 The Electromagnetic Hamiltonian

To consider transitions between states, we must consider the interaction with an outside influence. A very important case for astrophysics is that of radiative transitions, where the outside influence is an electromagnetic wave. Let us begin by specifying the interaction between the electromagnetic field and an electron (generalizable to any charged particle). The electric and magnetic fields may be expressed in terms of a scalar potential  $\phi(\vec{r}, t)$  and a vector potential  $\vec{A}(\vec{r}, t)$  such that

$$\vec{B}(\vec{r}, t) = \vec{\nabla} \times \vec{A}(\vec{r}, t); \quad \vec{E}(\vec{r}, t) = -\frac{1}{c} \frac{\partial \vec{A}(\vec{r}, t)}{\partial t} - \vec{\nabla} \phi(\vec{r}, t) \quad (2.17)$$

With the choice of the Coulomb gauge, (see Jackson, *Classical Electrodynamics*)

$$\vec{\nabla} \cdot \vec{A}(\vec{r}, t) = 0 \quad (2.18)$$

and assuming the scalar potential is time independent, the classical Lorentz force equation is

$$m \frac{d^2 \vec{r}}{dt^2} = -e \left[ \vec{E}(\vec{r}, t) + \frac{\vec{v}}{c} \vec{B}(\vec{r}, t) \right] \quad (2.19)$$

where  $m$  is the electron mass, and  $e$  is the electron charge.

Then  $\nabla^2 \phi = -4\pi\rho = 0$  if no sources are present and we can set  $\phi = 0$  and get

$$\nabla^2 \vec{A} - \frac{1}{c^2} \frac{\partial \vec{A}^2}{\partial t^2} = 0,$$

or more generally,

$$\vec{\nabla}^2 \vec{A} = -\frac{4\pi}{c} \vec{j} + \frac{1}{c} \vec{\nabla} \frac{\partial \phi}{\partial t}$$

and

$$\vec{E} = -\frac{1}{c} \frac{\partial \vec{A}}{\partial t}; \quad \vec{B} = \vec{\nabla} \times \vec{A}$$

To cast this into the Hamiltonian formulation, we make the substitution  $\vec{p} \rightarrow \vec{p} + \frac{e}{c} \vec{A}(\vec{r}, t)$ . Then adding the static scalar potential, we get

$$H = \frac{1}{2m} \left[ \vec{p} + \frac{e}{c} \vec{A}(\vec{r}, t) \right]^2 - e\phi(\vec{r}) \quad (2.20)$$

in analogy with the relation between the field-free Hamiltonian and Newton's Law outlined in Chapter 1. Now, we make the transition to quantum mechanics according to the prescription in Chapter 1 (eq. 1.6) and treat  $\vec{A}$  as a classical variable. This is known as the semi-classical treatment. The Schrodinger equation becomes

$$\frac{1}{2m} \left( \frac{\hbar}{i} \vec{\nabla} + \frac{e}{c} \vec{A} \right)^2 \Psi(\vec{r}, t) = [E + e\phi(\vec{r})] \Psi(\vec{r}, t) \quad (2.21)$$

Multiplying out the LHS, we get

$$-\frac{\hbar^2}{2m} \nabla^2 \Psi(\vec{r}, t) + \frac{\hbar e}{2imc} \vec{\nabla} \cdot (\vec{A} \Psi(\vec{r}, t)) + \frac{\hbar e}{2imc} \vec{A} \cdot (\vec{\nabla} \Psi(\vec{r}, t)) + \frac{e^2}{2mc^2} \vec{A}^2 \Psi(\vec{r}, t) \quad (2.22)$$

and by using the identity,  $(\vec{\nabla} \cdot (\vec{A}\Psi)) = \vec{A} \cdot (\nabla\Psi) + (\nabla \cdot \vec{A})\Psi$  the second term is

$$\frac{\hbar e}{2imc} [\Psi(\vec{r}, t) (\vec{\nabla} \cdot \vec{A}) + \vec{A} \cdot (\vec{\nabla}\Psi(\vec{r}, t))] \quad (2.23)$$

The first term in 2.23 is zero in the Coulomb gauge, leaving this as the Schrodinger equation:

$$\frac{-\hbar^2}{2m} \nabla^2 \Psi(\vec{r}, t) + \frac{\hbar e}{imc} \vec{A} \cdot \vec{\nabla}\Psi(\vec{r}, t) + \frac{e^2}{2mc^2} \vec{A}^2 \Psi(\vec{r}, t) = [E + e\phi(\vec{r})]\Psi(\vec{r}, t) \quad (2.24)$$

We assume the source of the fields to be very distant, so the region of interest is source free and  $\phi = 0$  for the electromagnetic fields. So, we can identify  $\phi$  as whatever potential is internal to the system (e.g., the Coulomb potential in an atom). Then,  $\frac{-\hbar^2}{2m} \nabla^2 - e\phi(r)$  can be identified as  $H^{(0)}$ , the unperturbed Hamiltonian, and  $H^{(1)}$ , the perturbation due to the electromagnetic field, is

$$H^{(1)} = \frac{\hbar e}{imc} \vec{A} \cdot \vec{\nabla} + \frac{e^2}{2mc^2} \vec{A}^2 \quad (2.25)$$

The ratio of the  $\vec{A}^2$  term to the  $\vec{A} \cdot \vec{\nabla}$  term is  $\frac{ie}{2\hbar c} \frac{\vec{A}^2}{\vec{A} \cdot \vec{\nabla}} = \frac{i}{2} \alpha \frac{\vec{A}^2/e}{\vec{A} \cdot \vec{\nabla}}$ , where  $\alpha = \frac{e^2}{\hbar c} \simeq \frac{1}{137}$ . Thus, unless the field is strong or  $\vec{A} \cdot \vec{\nabla}$  has a vanishing matrix element, the  $\vec{A}^2$  term will be negligible. In lasers, the field may be strong enough that the  $\vec{A}^2$  term cannot be ignored. The main place in astronomy where the  $\vec{A}^2$  term is relevant is when the  $\vec{A} \cdot \vec{\nabla}$  term has zero matrix element. The  $\vec{A}^2$  term corresponds to two photon transitions and is responsible for decay of the 2s to the 1s level in hydrogen. We will ignore it and consider only the first term in  $H^{(1)}$ .

Now we must specify the form of the vector potential. In a region devoid of sources,  $\vec{A}$  satisfies the wave equation:

$$\nabla^2 \vec{A} - \frac{1}{c^2} \frac{\partial^2 \vec{A}}{\partial t^2} = 0 \quad (2.26)$$

We posit a solution with the form of our harmonic perturbation (eq. 2.9):

$$\vec{A}(\vec{r}, t) = \vec{A}_0(\vec{r}) e^{i\omega t} + \vec{A}_0(\vec{r}) e^{-i\omega t} = 2\vec{A}_0(\vec{r}) \cos(\omega t) \quad (2.27)$$

and substitution into the wave equation shows that

$$\vec{A}_0(\vec{r}) = \vec{A}_0 e^{i\vec{k} \cdot \vec{r}} \quad (2.28)$$

With  $\vec{A}_0$  real and  $|\vec{k}|^2 = \frac{w^2}{c^2}$ , and  $\vec{k} \cdot \vec{A}_0 = 0$  (transverse wave). So we have

$$\vec{A}(\vec{r}, t) = \vec{A}_0 e^{i(\vec{k} \cdot \vec{r} - \omega t)} + c.c. \quad (2.29)$$

where c.c. stands for complex conjugate of the first term.

We could also add an arbitrary phase factor  $e^{i\delta}$  but it will make no difference in the end so we omit it. We can also specify the polarization of  $\vec{A}_0$  by using a unit vector  $\hat{\epsilon}$ , so that  $\vec{A}_0 = A_0 \hat{\epsilon}$ . Finally, we may want to represent a pulse of radiation, which we can do as a superposition of plane waves

$$\vec{A}(\vec{r}, t) = \int A_0(\omega) \hat{\epsilon} [e^{i(\vec{k} \cdot \vec{r} - \omega t)} + c.c.] d\omega \quad (2.30)$$

Now we pause to clarify some notation and develop some results that will be useful later. The spatial coordinate  $\vec{r}$  is, in principle, the distance vector from the source of the waves, but we now want to consider a coordinate fixed in the charge distribution that will interact with the wave. So, we relabel the coordinate we have called  $\vec{r}$  to  $\vec{R} = \vec{R}_0 + \vec{r}$ , where  $\vec{R}_0$  is the vector from the source of the waves to the center of the charge distribution, and  $\vec{r}$  is the vector to the point in question from the center of the charge distribution.

Then  $e^{i\vec{k}\cdot\vec{R}} = e^{i\vec{k}\cdot\vec{R}_0}e^{i\vec{k}\cdot\vec{r}}$  and the problem separates into center of mass motion, which does not concern us, and relative motion of charges in our charge distribution. So, we continue to use  $e^{i\vec{k}\cdot\vec{r}}$ , but we have changed the meaning of  $\vec{r}$ .

Next, we show how the electric field ( $\vec{E}$ ) and magnetic field ( $\vec{B}$ ) and the energy density  $U(\omega) \equiv \frac{1}{8\pi}(E^2 + B^2)$  are related to the vector potential. With equation 2.17, and noting that  $\nabla\phi = 0$  in our source-free region,

$$\vec{E} = -\frac{1}{c}\frac{\partial A}{\partial t} = -\frac{2\omega}{c}A_0\hat{\epsilon}\sin(\vec{k}\cdot\vec{r} - \omega t) = \frac{-i\omega}{c}\vec{A} \quad (2.31)$$

$$\vec{B} = \vec{\nabla}\times\vec{A} = -2A_0(\vec{k}\hat{\epsilon})\sin(\vec{k}\cdot\vec{r} - \omega t) = -iA_0\vec{k}\hat{\epsilon} \quad (2.32)$$

and  $U(\omega)$ , time averaged over a period of the oscillation is (recall  $|k|^2 = \frac{\omega^2}{c^2}$ )

$$U(\omega) = \frac{4\omega^2/c^2}{8\pi}|A_0|^2\left\langle 2\sin^2(\vec{k}\cdot\vec{r} - \omega t)\right\rangle = \frac{\omega^2}{2\pi c^2}|A_0|^2 \quad (2.33)$$

or per unit frequency interval ( $\nu = \frac{\omega}{2\pi}$ ;  $d\nu = \frac{d\omega}{2\pi}$ ), we have

$$U(\nu) = 2\pi U(\omega) = \frac{\omega^2}{c^2}|A_0|^2 = \frac{4\pi^2\nu^2}{c^2}|A_0|^2 \quad (2.33b)$$

Now we are ready to consider our perturbation (2.25), neglecting the  $A^2$  term; and using 2.29 and 2.30.

$$\begin{aligned} H^{(1)} &= \frac{\hbar e}{imc}\vec{A}\cdot\vec{\nabla} = \frac{\hbar e}{imc}\int A_0(\omega)\hat{\epsilon}e^{i(\vec{k}\cdot\vec{r}-\omega t)}\cdot\vec{\nabla} + c.c.\cdot\vec{\nabla} \\ &\equiv H^-(\vec{r})e^{-i\omega t} + H^+(\vec{r})e^{i\omega t} \end{aligned} \quad (2.34)$$

where

$$H^-(\vec{r}) = \frac{\hbar e}{imc}\int d\omega A_0(\omega)e^{i\vec{k}\cdot\vec{r}}\hat{\epsilon}\cdot\vec{\nabla} \quad (2.35)$$

We will consider only the  $H^-$  ( $\vec{r}$ ) term (stimulated absorption) and substitute into Fermi's Golden Rule in the form (2.15a), as modified to apply to  $H^-$  by using  $\delta\left(\frac{E_q - E_k}{\hbar} - \omega\right)$

$$R_{k\rightarrow q} = \frac{2\pi}{\hbar^2}|H_{kq}^-|^2\delta\left(\frac{E_q - E_k}{\hbar} - \omega\right) \quad \overline{\uparrow}_k^q$$

$$\begin{aligned}
&= \frac{2\pi}{\hbar^2} \int |\langle q | \frac{\hbar e}{imc} A_0(\omega) e^{i\vec{k}\cdot\vec{r}} \hat{\epsilon} \cdot \vec{\nabla} | k \rangle|^2 \delta\left(\frac{E_q - E_k}{\hbar} - \omega\right) d\omega \\
&= \frac{2\pi}{\hbar^2} \frac{\hbar^2 e^2}{m^2 c^2} \int A_0^2(\omega) \delta\left(\frac{E_q - E_k}{\hbar} - \omega\right) d\omega \left| \langle q | e^{i\vec{k}\cdot\vec{r}} \hat{\epsilon} \cdot \vec{\nabla} | k \rangle \right|^2 \\
&= \frac{2\pi e^2}{m^2 c^2} A_0^2(\omega_{kq}) \left| \langle q | e^{i\vec{k}\cdot\vec{r}} \hat{\epsilon} \cdot \vec{\nabla} | k \rangle \right|^2
\end{aligned} \tag{2.36}$$

with  $\omega_{kq} = (E_q - E_k)/\hbar$ .

If we took the other term ( $H^+(\vec{r})$ ) and arranged it so we dealt with the same initial state  $q$  above  $k$ , we would get the rate for stimulated emission  $R_{q \rightarrow k}$  and find  $R_{q \rightarrow k} = R_{k \rightarrow q}$  (we assume no degeneracy here). We don't get spontaneous emission out of this approach because we have treated the field classically. We can "motivate" it by recalling that  $U(\omega) = \frac{\omega^2}{2\pi c^2} |A_0|^2$  and noting that in a quantum description of the radiation field  $U(\omega) = n(\omega)\hbar\omega$  where  $n(\omega)$  is the photon occupation number, the number density of photons per "mode of vibration" of angular frequency  $\omega$ . So we can write  $|A_0|^2 = (n(\omega)\hbar\omega 2\pi c^2/\omega^2)$  and (2.36) could be written in terms of a photon occupation number. In QED, the vector potential gets used to construct creation and destruction operators, analogous to raising and lowering operators for a harmonic oscillator. The *creation* of a photon requires a modification of the above expression to  $|A_0|^2 = \frac{2\pi c^2}{\omega^2} \hbar\omega (n(\omega) + 1)$ . Note the analogy to the harmonic oscillator, where the photon occupation number is like the quantum number of the oscillator. So for creation of a photon (emission),  $|A_0|^2 \neq 0$  even if  $n(\omega) = 0$ . There is a zero point energy of the electromagnetic field that can "induce" a transition – this is the term that leads to spontaneous emission.

We have used  $H^{(1)} = \frac{\hbar e}{imc} \vec{A} \cdot \vec{\nabla} = \frac{e}{mc} \vec{A} \cdot \vec{p}$  partly because it led to expressions where we could see the connection to the more general QED formulations. But one often sees the perturbation expressed in terms of the electric field interacting with the moments of the charge distribution, and this will be easier for us to deal with. Since  $\vec{A}$  is proportional to  $\vec{E}$ , we can convert that part, but turning  $\vec{p}$  into a moment of the distribution is trickier. We use the Heisenberg equation of motion for the time evolution of an operator,  $C$ ,

$$\frac{dC}{dt} = \frac{\partial C}{\partial t} + \frac{i}{\hbar} [H, C] \tag{2.37}$$

where  $[H, C]$  is the commutator. Then we note that

$$\vec{p} = m\dot{\vec{r}} \quad \text{and} \quad \frac{\partial r}{\partial t} = 0 \quad \text{so} \quad \dot{\vec{r}} = \frac{i}{\hbar} [H, \vec{r}]$$

Recalling that  $\vec{E} = \frac{-i\omega}{c} \vec{A}$  (equation 2.31) we get

$$\begin{aligned}
H^{(1)} &= \frac{e}{mc} \vec{A} \cdot \vec{p} = \frac{e}{mc} m \vec{A} \cdot \dot{\vec{r}} = \frac{ie}{\hbar c} \vec{A} \cdot [H, \vec{r}] \\
&= \frac{ie}{\hbar c} \left( \frac{c}{-i\omega} \right) \vec{E} \cdot [H, \vec{r}] = \frac{-e}{\hbar\omega} \vec{E} \cdot (H\vec{r} - \vec{r}H)
\end{aligned} \tag{2.38}$$

Since we are dealing with first-order perturbation theory, we can replace  $H$  by  $H^{(0)}$  so

$$\begin{aligned}
\langle q|H^{(1)}|k\rangle &= \frac{-e}{\hbar\omega}\vec{E}\cdot\langle q|H^{(0)}\vec{r}-\vec{r}H^{(0)}|k\rangle \\
&= \frac{-e}{\hbar\omega}\vec{E}\cdot\langle q|\vec{r}|k\rangle(E_q^{(0)}-E_k^{(0)}) \\
&= -e\vec{E}\cdot\langle q|\vec{r}|k\rangle \\
&= \langle q|\vec{E}\cdot(-e\vec{r})|k\rangle
\end{aligned} \tag{2.39}$$

So we can use  $H^{(1)} = \vec{E}\cdot(-e\vec{r})$  as our perturbation. In this form, equation (2.36) would become

$$\begin{aligned}
R_{k\rightarrow q} &= \frac{2\pi}{\hbar^2}\int d\omega |\langle q|H^-(\omega)|k\rangle|^2 \delta\left(\frac{E_q-E_k}{\hbar}-\omega\right) \\
&= \frac{2\pi}{\hbar^2}|E_0(\omega_{kq})|^2|\langle q|e^{i\vec{k}\cdot\vec{r}}\hat{\epsilon}\cdot(-e\vec{r})|k\rangle|^2
\end{aligned} \tag{2.40}$$

and since  $|E_0(\omega_{kq})|^2 = \frac{\omega^2}{c^2}|A_0(\omega_{kq})|^2 = \frac{\omega^2}{c^2}\left(\frac{2\pi c^2}{\omega_{kq}^2}U(\omega_{kq})\right) = 2\pi U(\omega_{kq}) = U(\nu_{kq})$  we get

$$R_{k\rightarrow q} = \frac{2\pi}{\hbar^2}|\langle q|e^{i\vec{k}\cdot\vec{r}}\hat{\epsilon}\cdot(-e\vec{r})|k\rangle|^2U(\nu_{kq}) \tag{2.40b}$$

## 2.3 The Einstein Coefficients

We can rewrite (2.40b) as

$$R_{k\rightarrow q} = B_{kq}U(\nu) \tag{2.41}$$

where  $B_{kq}$  is the Einstein coefficient for absorption of a photon and a transition from state  $k$  to  $q$ . Likewise for emission, we have

$$R_{qk} = B_{qk}U(\nu) + A_{qk}, \tag{2.42}$$

where the ‘‘Einstein A’’ reflects the presence of the  $(N(\nu) + 1)$  factor in emission. For a system in thermodynamic equilibrium with a radiation field characterized by a temperature  $T$ , we must have a steady state: transitions to  $k$  must equal those to  $q$ . If we denote the number of atoms in  $q$  by  $n_q$  and the number in  $k$  by  $n_k$ , we get

$$n_q R_{qk} = n_k R_{kq} \Rightarrow \frac{n_q}{n_k} = \frac{B_{kq}U(\nu)}{B_{qk}U(\nu) + A_{qk}} \tag{2.43}$$

Now suppose state  $q$  has degeneracy  $g_q$  and state  $k$  has degeneracy  $g_k$ . The Boltzmann equation gives

$$\frac{n_q}{n_k} = \frac{g_q}{g_k} e^{-E_{qk}/kT} = \frac{g_q}{g_k} e^{-h\nu/kT}, \quad \text{where } \nu = \frac{E_{qk}}{h} \tag{2.44}$$

and

$$U(\nu) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{h\nu/kT} - 1} \quad (2.45)$$

To make this work, we must have

$$\frac{B_{kq}}{B_{qk}} = \frac{g_q}{g_k} \quad \text{and} \quad \frac{A_{qk}}{B_{qk}} = \frac{8\pi h\nu^3}{c^3} \quad (2.46)$$

These relations allow us to determine the rest of the Einstein coefficients if we know one. Note also that the ratio of stimulated to spontaneous emission is

$$\frac{B_{qk}U(\nu)}{A_{qk}} = (e^{h\nu/kT} - 1)^{-1} = n(\nu) = 2\pi n(w) \quad (2.47)$$

where  $n(\nu)$  number of photons per  $\nu$  mode, as expected. At low  $\nu(h\nu \ll kT)$ , stimulated emission will dominate, whereas at high  $\nu(h\nu \gg kT)$ , it will be negligible compared to spontaneous emission. Note that these relationships are appropriate only if we use the energy density. Many astronomical texts use the mean intensity, which is

$$I_\nu = \frac{c}{4\pi} U_\nu \quad (2.48)$$

in which case the expressions for  $B$ 's must be multiplied by  $4\pi/c$  to keep the rate the same. In this case

$$\frac{A_{qk}}{B_{qk}} = \frac{c}{4\pi} \left( \frac{8\pi h\nu^3}{c^3} \right) = \frac{2h\nu^3}{c^2} \quad (2.49)$$

This is the convention used by Rybicki and Lightman, and by Mihalas. Spitzer, on the other hand, uses the energy density formulation, as I will do. Note that  $A$  must have the same value in either formulation so it is a safer quantity to give.

Now let us attempt to deal with the matrix element

$$\langle \psi_q | -e\vec{r} \cdot \hat{\epsilon} e^{i\vec{k} \cdot \vec{r}} | \psi_k \rangle \quad (2.50)$$

The plane wave  $e^{i\vec{k} \cdot \vec{r}}$  can be expanded in a series involving spherical harmonics and spherical Bessel functions.

$$e^{i\vec{k} \cdot \vec{r}} = 4\pi \sum_{lm} i^l j_\ell(kr) Y_{lm}^*(\hat{k}) Y_{lm}(\hat{r}) \quad (2.51)$$

This forbidding expression can be used to define various multipole moments. Fortunately, it converges rapidly; indeed for most purposes only the first term is necessary. Taking only the first term leads to the dipole approximation. An alternative way to write the expansion of  $e^{i\vec{k} \cdot \vec{r}}$  is

$$e^{i\vec{k} \cdot \vec{r}} = \sum_{n=0}^{\infty} \frac{(i)^n}{n!} (\vec{k} \cdot \vec{r})^n = 1 + i\vec{k} \cdot \vec{r} + \dots \quad (2.51b)$$

## 2.4 The Dipole Approximation

Keeping only the first term in the expansion (2.51b) gives us a matrix element.

$$\langle \psi_q | -e\vec{r} \cdot \hat{\epsilon} | \psi_k \rangle \quad (2.52)$$

and we define the dipole moment operator  $\vec{\mu} = -e\vec{r}$ , which classically is just the separation between the positive and negative charges. For a very useful form of the operator, see Appendix A.5.

We also define the matrix element of the component of  $\vec{\mu}$  along the polarization vector. Suppose  $\hat{\epsilon} = \hat{z}$ , then

$$\langle \psi_q | \vec{\mu} \cdot \hat{\epsilon} | \psi_k \rangle = \langle \psi_q | \mu_z | \psi_k \rangle \quad (2.53)$$

but since we deal normally with unpolarized light, we want to average over all angles between the dipole moment and the polarization. The result is that

$$|\langle \psi_q | \mu_z | \psi_k \rangle|^2 = \frac{1}{3} |\mu_{kq}|^2 \quad (2.54)$$

where

$$|\mu_{kq}|^2 = \sum_{\text{final states}} \{ |\langle \psi_q | \mu_x | \psi_k \rangle|^2 + |\langle \psi_q | \mu_y | \psi_k \rangle|^2 + |\langle \psi_q | \mu_z | \psi_k \rangle|^2 \} \quad (2.55)$$

So that, on average, the light is polarized along a given axis  $\frac{1}{3}$  of the time. The sum is over final degenerate states (e.g., different  $m$  values). Hence,

$$B_{kq} = \frac{2\pi}{\hbar^2} \frac{|\mu_{kq}|^2}{3} = \frac{8\pi^3}{3\hbar^2} |\mu_{kq}|^2 \quad (2.56)$$

All we have to do is compute the dipole moment matrix element for the states under consideration.

## 2.5 Magnetic Dipole and Electric Quadrupole

In cases where the electric dipole moment has a vanishing matrix element, we must consider higher order terms in the plane wave expansion (equation 2.51). While the procedure is, in principle, straightforward, a different approach will illuminate some terminology. Returning to equation 2.51b and now taking the second term ( $i\vec{k} \cdot \vec{r}$ ), we are faced with the product,

$$\vec{k} \cdot \vec{r} \cdot \hat{\epsilon} \cdot \vec{p} \quad (2.57)$$

where we return to the  $\vec{A} \cdot \vec{p}$  form and consider  $\hat{\epsilon}$  to be the unit vector in the direction of  $\vec{A}$ . Now the  $\vec{r} \cdot \hat{\epsilon}$  is a dyad or tensor. Through some identities and manipulation, one can recast this to

$$\frac{1}{2} (\hat{\epsilon} \cdot \vec{p} \cdot \vec{k} \cdot \vec{r} + \hat{\epsilon} \cdot \vec{r} \cdot \vec{p} \cdot \vec{k}) + \frac{1}{2} (\vec{k} \cdot \hat{\epsilon}) \cdot (\vec{r} \cdot \vec{p}). \quad (2.58)$$

The first term is called an electric quadrupole. The second term is proportional to  $\vec{B} \cdot \vec{L}$  since  $\vec{r} \cdot \vec{p} = \vec{L}$  and  $\vec{B}$  is perpendicular to  $\vec{k}$  and  $\hat{\epsilon}$ . The second term is called a magnetic dipole term since

the magnetic dipole is proportional to the angular momentum and magnetic fields will interact with magnetic dipoles.

While the expansion in terms of partial waves (equation 2.51) is more attractive, the form involving dyadics was the historical route and Cartesian coordinates were used to analyze the quadrupole moment, leading to much confusion and awkwardness.

The form above for the magnetic dipole is incomplete because of the existence of spin (a relativistic effect and thus outside our current formalism). We need to add another term to the Hamiltonian:

$$\frac{ge}{2mc} \vec{S} \cdot \vec{B}(\vec{r}, t) \quad (2.59)$$

where  $ge/2mc$  is the magnetic moment of spin  $\vec{S}$ , and  $g = 2$  for spin. This term is of the same order of magnitude as the magnetic dipole term above and represents the interaction of the magnetic field with the *spin* magnetic moment, just as the above expression represents the interaction with the orbital angular momentum.

While we could go on considering higher and higher terms, in practice the electric quadrupole and magnetic dipole are the highest terms of much relevance in astrophysics.

## 2.6 Symmetry Operations and Some Selection Rules

Selection rules are ultimately a consequence of symmetry operations. The symmetry operations of particular relevance are inversion and rotation. Let us first consider inversion, by which we mean a reflection through the origin of all the coordinates:

$$x, y, z \rightarrow -x, -y, -z, \quad \text{or} \quad \vec{r} \rightarrow -\vec{r} \quad (2.60)$$

or, in spherical coordinates,

$$r \rightarrow -r; \quad \theta \rightarrow \pi - \theta; \quad \phi \rightarrow \phi + \pi \quad (2.61)$$

The operator that performs this transformation (denoted  $\iota$ ) is a unitary operator. Two subsequent inversions return the system to the initial state, so  $\iota^2 = \mathbf{1}$ , where  $\mathbf{1}$  is the unity operator. Consequently, we can arrange for the eigenvalues of the inversion operator to be  $\pm 1$ . These eigenvalues are called the parity. For all central force problems, the Hamiltonian depends only on the magnitude of  $\vec{r}$  so it is unchanged by inversion. The inversion operator commutes with  $H$  and the eigenfunctions of  $H$  can also be eigenfunctions of  $\iota$  and parity is a good quantum number. Each state can be assigned a definite parity. We can use equation 2.61 to determine the parity of the spherical harmonics.

$$\iota Y_{\ell m}(\theta, \varphi) = Y_{\ell m}(\pi - \theta, \phi + \pi) = (-1)^\ell Y_{\ell m}(\theta, \varphi) \quad (2.62)$$

as can be verified by examining the detailed forms of  $Y_{\ell m}(\theta, \varphi)$ . So the parity of the  $\ell$ th spherical harmonic is  $(-1)^\ell$ .

Now let us consider the action of the inversion operator on another operator. If it leaves the operator unchanged, we say that the operator is even; if the operator changes under inversion, we say it is odd. For example

$$\iota \vec{\mu} = \iota(-e\vec{r}) = -e\vec{r} = -e(-\vec{r}) = -\vec{\mu} \quad (2.63)$$

Table 2.1

Parity Behavior for Multipole Operators

$k$	pole	Name	Parity	Parity
			Electric	Magnetic
1	$2^1$	dipole	odd	even
2	$2^2$	quadrupole	even	odd
3	$2^3$	octupole	odd	even

Table 2.2

Allowed Transition Multipoles

	Electric	Magnetic	Allowed $k$
Parity	must change	cannot change	$k = 1, 3, \dots$
Parity	cannot change	must change	$k = 2, 4, \dots$

so the electric dipole operator is odd. In contrast,  $\imath\vec{p} = -\vec{p}$ , so the magnetic dipole operator is even:

$$\imath\vec{\mu}_M = \imath(\vec{r}\mathbf{x}\vec{p}) = (-\vec{r})\mathbf{x}(-\vec{p}) = \vec{\mu}_M \quad (2.64)$$

and the electric quadrupole operator is also even. In general, the pole is given by  $2^k$ .

Now let us consider static properties of systems, i.e., expectation values. If the operator is odd, and the state has definite parity, then the expectation value of that operator must vanish. To see this, we write out the full expression for the expectation value of, for example,  $x$ .

$$\langle q|x|q\rangle \equiv \int_{-\infty}^{\infty} \psi_q^+(x)x\psi_q(x)dx = \int_{-\infty}^{\infty} |\psi_q|^2 x dx. \quad (2.65)$$

If  $\psi_q(x)$  has definite parity,  $|\psi_q(-x)|^2 = |\psi_q(x)|^2$  so the integrand is odd and the integral must be zero. Consequently, if parity is a good quantum number, a system can have only even electric and odd magnetic multipole moments.

The fact that the expectation value of an operator is zero does not mean that its transition matrix elements are zero. The requirement for a non-zero value is that the whole integrand must be even under inversion, so an operator which has odd parity can induce transitions only between states with *opposite* parity. So, if parity is a good quantum number, we have the following rules:

Now consider rotational symmetry. For central force problems, the Hamiltonian is unchanged by a rotation of the coordinate system. Then the total angular momentum and its projection onto one axis are the good quantum numbers. We will use  $J$  and  $M$  to represent these, so a state can be specified by the values of  $J$ ,  $M$ , and, if necessary other quantum numbers that we will represent collectively by  $\gamma$ . The state is then specified by  $|\gamma J M\rangle$ .

It is useful to characterize operators according to their behavior under rotation of the coordinate system. This leads to the concept of spherical tensors or irreducible tensors. This is not as scary as it sounds because it turns out that our old friends, the spherical harmonics, are such spherical tensors. The selection rules arising from rotational symmetries all involve conservation of angular momentum. Since angular momentum is quantized, this conservation reduces to a kind of elaborate bookkeeping. This bookkeeping is greatly facilitated by the use of Clebsch-Gordon coefficients

(Appendix A1) and their symmetrical forms, the 3-J symbols (Appendix A2). These elements are brought together in the Wigner-Eckart theorem (Appendix A4) to show that matrix elements of an arbitrary tensor operator  $T_{kq}$  can be expressed in a simpler form in which all the dependence on the azimuthal quantum numbers is carried by a 3-J symbol.

The properties of the 3-J symbols give us two very general selection rules for radiative transitions. For an initial state  $|J' M'\rangle$  and a spherical multipole of order  $k$ , the allowed transitions must satisfy these two inequalities:

$$|J - J'| \leq k \leq |J + J'|; \quad M - M' = q \quad (2.66)$$

The first rule (also known as the triangle relation and denoted  $\Delta(J_1, J_2, J_3)$  in general) can be written as  $k = |J - J'|, |J = J' + 1|, \dots, |J + J'|$ . For dipole transitions,  $k = 1$ , we have  $\Delta J = 0, \pm 1$  by the first inequality, and  $J = 0 \rightarrow J = 0$  is not allowed by the second inequality. Physically these make sense because the photon carries one unit ( $\hbar$ ) of spin angular momentum and  $k - 1$  units of orbital angular momentum, so only  $\sqrt{\hbar^2 k(k+1)}$  units of angular momentum to transfer to the atom or molecule.

## 2.7 Natural Line Broadening

From our discussion of Einstein coefficients, we know that systems may change to lower energy states by spontaneous transitions. This means that the lifetime in excited states is finite. Through the energy-time version of the uncertainty principle, this causes a spread in the energy of the excited state:

$$\Delta E \sim \hbar/\tau, \quad (2.67)$$

where  $\tau$  is the lifetime of the excited state. The lifetime is the inverse of the total rate of decay of state  $j$ :

$$\frac{1}{\tau_j} = \sum_k A_{jk} \quad (2.68)$$

Thus, the more rapid the decay, the more the energy state is uncertain or spread out. Strictly speaking, excited states are not true eigenstates. As long as the lifetime is long compared to other relevant times, we can ignore this issue, but it does produce a finite spread in photon frequencies referred to as the natural line width. Neglecting stimulated transitions, the natural width will be roughly

$$\Delta\nu_{ji} \sim \frac{\Delta E_{ji}}{h} = \frac{1}{2\pi\tau_j} = \frac{1}{2\pi} \left( \sum_k A_{jk} + \sum_k A_{ik} \right) \quad (2.69)$$

The detailed shape of the line is less obvious, but a classical argument suffices to get the right form. We can consider the radiating system to be a harmonically oscillating charge that also has a damping term. The damping term arises from the finite lifetime: the classical oscillator will have some friction; the quantum system will decay. Since the classical model has considerable utility, we will recall its features briefly; it is sometimes called the Lorentz model.

Consider a charged mass on a spring. A damped harmonic oscillator with a charge will emit electromagnetic waves. We consider here an undriven, damped harmonic oscillator, which satisfies the following equation:

$$\ddot{x} + \Delta\omega\dot{x} + \omega_0^2x = 0 \quad (2.70)$$

By guessing  $x = x_0e^{-i\omega t}$ , we obtain  $-\omega^2x_0 = \Delta\omega(-i\omega)x + \omega_0^2x = 0$ , where  $\omega_0 = \sqrt{k/m}$  and  $\Delta\omega = \gamma/m$ . If  $\Delta\omega \ll \omega_0$ ,

$$x = x_0e^{-\Delta\omega t/2}\cos(\omega t). \quad (2.71)$$

Fourier analysis leads to a spectrum of

$$|x(\omega)|^2 = \left(\frac{x_0}{4\pi}\right)^2 \frac{1}{(\omega - \omega_0)^2 + (\frac{\Delta\omega}{2})^2} \quad (2.72)$$

With the radiated power of a dipole as

$$P = \frac{1}{3} \frac{e^2 \omega_0^4 |x(\omega)|^2}{c^3} \quad (2.73)$$

which, for  $\omega \sim \omega_0$ , can be written as

$$P \propto \frac{1}{(\omega - \omega_0)^2 + (\frac{\Delta\omega}{2})^2} \quad (2.74)$$

If plotted, this function has a shape called a Lorentzian with a width of  $\Delta\omega$ , or in frequency space,  $\Delta\nu = \Delta\omega/(2\pi)$ . The connection to the argument from the uncertainty principle is that  $\Delta\omega \sim \sum A$ .

Other line broadening mechanisms may come into play. If collisions cause transitions from either of the states involved in the radiative transition on a relevant timescale, they also affect the lifetime. The shape is still a Lorentzian, but  $\Delta\omega$  is now the sum of both radiative and collisional rates. This is called pressure broadening and it is important in planetary atmospheres. In dense, ionized gases, the electric fields on ions can create a Stark effect, which broadens the lines in a somewhat different way. Finally, the spread of velocities in a Maxwellian velocity distribution (equation 0.55) and the Doppler effect causes a broadening of lines. For slow transitions (molecular vibration and rotation) in dilute media, this is often the dominant effect.

## 2.8 Summary

Because the previous sections of this chapter may have been rather arcane in places, it may be helpful to summarize the important results here. The full interaction of a charge, or a charge distribution, with an electromagnetic field is very complex. If we ignore 2-photon processes, we can deal with a matrix element of the following form:

$$\langle \psi_q | -e\vec{r} \cdot \hat{\epsilon} e^{i\vec{k} \cdot \vec{r}} | \psi_k \rangle \quad (2.75)$$

where  $-e$  is the charge on the electron and  $\hat{\epsilon}$  is the unit vector in the direction of the electric field. If  $\vec{k} \cdot \vec{r}$  is small (consider when this approximation might fail), we can expand the exponential

into multiple poles. The levels of transitions of interest to us are electric dipole, magnetic dipole, and electric quadrupole. Here we give expressions for Einstein A values for each of these because the A values are free of the ambiguities in the definitions of the B values.

For the electric dipole,

$$A_{qk} = \frac{64\pi^4 \nu_{qk}^3}{3hc^3} |\mu_{qk}|^2, \quad (2.76)$$

where  $\nu_{qk} = (E_q - E_k)/h$  and  $|\mu_{qk}|^2$  is the dipole matrix element for the transition from state  $q$  to state  $k$ . Note that

$$|\mu_{qk}|^2 = (g_k/g_q) |\mu_{kq}|^2 \quad (2.77)$$

The expression for the magnetic dipole is exactly the same

$$A_{qk} = \frac{64\pi^4 \nu_{qk}^3}{3hc^3} |\mu_{qk}|^2, \quad (2.78)$$

where now  $|\mu_{qk}|^2$  is the matrix element of the *magnetic* dipole operator.

For the electric quadrupole, the equation changes to

$$A_{qk} = \frac{32\pi^6 \nu_{qk}^5}{5hc^5} |Q_{qk}|^2, \quad (2.79)$$

where  $Q_{qk}$  is the matrix element of the tensor quadrupole moment operator  $\vec{Q}$ , which can be expressed in Cartesian coordinates as  $Q_{\alpha\beta} = \sum e(3x_\alpha x_\beta - \delta_{\alpha\beta} r^2)$  where  $x_\alpha$  and  $x_\beta$  are the  $x$  coordinates of the radius vector  $\vec{r}$  between two charges  $\alpha$  and  $\beta$ ,  $\delta_{\alpha\beta}$  is the Kronecker delta, and the summation is performed over all charged particles. This is the usual formulation despite the fact that spherical coordinates would be more natural. Quadrupole transitions are notable in the rotational transitions of the H<sub>2</sub> molecule, and the classic text on this is by James and Coolidge [Ap. J. 87, 438 (1938)]. It took more than 40 years for these lines to be observed.

We have left the equations in terms of the matrix elements of operators, which is where all the physics remains to be done. Further treatment of the matrix elements can be done only in the context of specific systems, and we will return to these as they become relevant throughout the course.

For now, let's note that several other quantities are used in some cases, so it is useful to know the relationships. The first, used a lot in vibrational molecular spectroscopy, is the transition strength,

$$s_{qk} \equiv |\mu_{qk}|^2 g_q \equiv s_{kq} \quad (2.80)$$

where  $g_q$  is the degeneracy of state  $q$ . The advantage of  $s_{qk}$  is that it is symmetric by definition.

A quantity often used in electronic transitions is the oscillator strength, related to the classical Lorentz model discussed in §2.7. The oscillator strength, usually denoted  $f_{qk}$ , corrects the classical expression for the cross section for energy absorption by an oscillator [ $\sigma = \pi e^2/(m_e c)$ ] for quantum effects. The relation to the matrix element is

$$f_{qk} = \frac{8\pi^2 m_e \nu}{3hc^2} |\mu_{qk}|^2 \quad (2.81)$$

and the cross section is then

$$\sigma_{qk} = \frac{\pi e^2}{m_e c} f_{qk} \quad (2.82)$$

and the Einstein B is

$$B_{qk} = \frac{c}{h\nu} \sigma_{qk} = \frac{\pi e^2}{h\nu m_e} f_{qk} \quad (2.83)$$

One virtue of the oscillator strength is that it follows a sum rule,  $\sum_k f_{qk} = 1$  if only one electron makes a transition between orbitals. For dipole electronic transitions in the visual range, the values of oscillator strengths are often close to unity.

## Chapter 3

# Atomic Structure – Hydrogen

### 3.1 Review of the Hydrogen Atom without Spin

We will assume some familiarity with the quantum treatment of the hydrogen atom, but review it briefly as preparation for dealing with more complex systems. Its central importance in astronomy also dictates a fairly detailed discussion of its spectrum. We begin with an approximate model for the hydrogen atom: we ignore relativistic effects, including spin, and deal only with the Coulomb interaction between the nucleus and the electron. In the same approximation, our results will apply to any one-electron (hydrogenic) atom or ion (e.g.,  $\text{He}^+$ ,  $\text{Li}^{++}$ , etc.). As we know, we can reduce the two body problem to a problem involving the motion of the center of mass and a problem involving the relative motion – or, that of a single particle with mass  $\mu$  moving in a central potential.

$$\mu = \frac{m_e m_n}{m_e + m_n} \quad (3.1)$$

Where  $m_e$  is the mass of the electron and  $m_n$  is the mass of the nucleus.

$$\mu \simeq m_e \left(1 - \frac{m_e}{m_n}\right) \simeq m_e \quad (3.2)$$

We can set  $\mu = m_e$  in most cases because  $m_n \simeq 1836m_e$  for  $H$ .

We know the central force problem is separable and we know the solutions for the angular part – the spherical harmonics. The potential  $V = -Ze^2/r$ , where  $r$  is the distance between electron and nucleus, is time independent. So, we have the usual

$$\Psi(\vec{r}, t) = \psi(\vec{r}) e^{-iEt/\hbar} \quad (3.3)$$

where  $H\psi(\vec{r}) = E\psi(\vec{r})$  and  $\psi(\vec{r})$  is the solution to

$$\left( \frac{-\hbar^2}{2\mu} \nabla^2 - \frac{Ze^2}{r} - E \right) \psi(\vec{r}) = 0 \quad (3.4)$$

Anticipating future developments, we write

$$\psi(\vec{r}) \equiv \psi_{nlm}(\vec{r}) = R_{nl}(r) Y_{lm}(\theta, \varphi), \quad (3.5)$$

and using the results from the rigid rotor, we get

$$\left[ \frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d}{dr} \right) - \frac{l(l+1)}{r^2} + \frac{2\mu}{\hbar^2} \left( E_{nl} + \frac{Ze^2}{r} \right) \right] R_{nl}(r) = 0 \quad (3.6)$$

If we change coordinates to

$$\rho = \frac{2}{\hbar} \sqrt{-2\mu E_{nl}} r, \quad (3.7)$$

we find that

$$R_{nl}(\rho) = e^{-\rho/2} \rho^l L_{n+l}^{2l+1}(\rho) N_{nl}. \quad (3.8)$$

where  $L_{n+l}^{2l+1}(\rho)$  are associated Laguerre polynomials and  $2l+1$  and  $n+l$  must be integers.  $n$  is called the principal quantum number.  $N_{nl}$  is a normalization constant. The  $L_{n+l}^{2l+1}(\rho)$  require that

$$2l+1 \leq n+l \quad (3.9)$$

or

$$l \leq n-1$$

Introducing the Bohr radius  $a_0 = \frac{\hbar^2}{\mu e^2} \simeq 0.529 \text{ \AA}$ , we can rewrite (3.6) as

$$\rho = \frac{2Z}{na_0} r, \quad (3.10)$$

where we have used the fact that  $n$  was defined in the process of obtaining the associated Laguerre polynomials to be

$$n = \frac{\mu Ze^2}{\hbar \sqrt{-2\mu E_{nl}}} \quad (3.11)$$

Then the eigenvalues  $E_{nl}$  are

$$E_{nl} = -\frac{\mu Z^2 e^4}{2\hbar^2 n^2} \equiv E_n, \quad (3.12)$$

since  $l$  does not enter. The normalization constant  $N_{nl}$  is

$$N_{nl} = - \left\{ \left( \frac{2Z}{na_0} \right)^3 \frac{(n-l-1)!}{2n[(n+l)!]^3} \right\}^{\frac{1}{2}} \quad (3.13)$$

If we return to our original coordinate  $r$ , we can write the first few radial functions  $R_{nl}(r)$

$$\begin{array}{ll} n & l \\ \hline 1 & 0 \\ 2 & 0 \\ 2 & 1 \end{array} \quad \begin{aligned} R_{10}(r) &= 2 \left( \frac{Z}{a_0} \right)^{\frac{3}{2}} e^{-Zr/a_0} \\ R_{20}(r) &= \left( \frac{Z}{2a_0} \right)^{\frac{3}{2}} \left( 2 - \frac{Zr}{a_0} \right) e^{-Zr/2a_0} \\ R_{21}(r) &= \frac{1}{\sqrt{3}} \left( \frac{Z}{2a_0} \right)^{\frac{3}{2}} \left( \frac{Zr}{a_0} \right) e^{-Zr/2a_0} \end{aligned} \quad (3.14)$$

The  $R_{nl}(r)$  are orthonormal:

$$\int_0^\infty R_{nl}(r)R_{n'l'}(r)r^2dr = \delta_{nn'} \quad (3.15)$$

so the full eigenfunctions,  $\psi_{nlm}(\vec{r}) = R_{nl}(r)Y_{lm}(\theta, \varphi)$ , satisfy

$$\langle \psi_{nlm} | \psi_{n'l'm'} \rangle = \delta_{nn'}\delta_{ll'}\delta_{mm'} \quad (3.16)$$

For accurate work, we rewrite the eigenvalues:

$$E_n = R_H \frac{Z^2}{n^2} eV. \quad (3.17)$$

The Rydberg for hydrogen,  $R_H$ , is used in the equation for energy levels of the hydrogen atom. One occasionally encounters some specialized units for atomic work. The Bohr is a unit of length equal to the Bohr radius, defined as  $a_0 = \hbar^2/m_e e^2$ . A related unit is the Hartree, defined as twice the Rydberg. The most accurate values of the Bohr, the Rydberg, and the Hartree are given in the limit of infinite proton mass and have values as of 2010 of

$$\begin{aligned} 1 & \text{ Bohr} = 5.2917721092(17) \times 10^{-9} \text{ cm} \\ 1 & \text{ } R_\infty = \frac{e^4 m_e}{2\hbar^2} = 13.60569253(30) \text{ eV} \\ 1 & \text{ } H_\infty = 27.21138505(60) \text{ eV} \end{aligned} \quad (3.18)$$

To convert these to use in hydrogenic atoms or ions, one must correct for the finite mass, via  $R_H = \frac{R_\infty}{1+m_e/m_N}$ , where  $m_N$  is the mass of the nucleus.

Note that the energy levels are highly degenerate with  $g_n = n^2$  since for a given  $n$ , we have allowed  $l = 0, 1, \dots, n-1$  and for each  $l$ ,  $m = -l, -l+1, \dots, l-1, l$ , but  $l, m$  do not enter the energy. We shall soon remove some of this degeneracy.

Let us first introduce some more notation, spectroscopic notation to be precise. This notation arose in the attempt of spectroscopists to classify spectra in the absence of a theory for atomic structure. In this sense, they are relics, like spectral types of stars, from the early days. To understand the origin of this notation, we first consider the spectrum of the H atom. If we plot the energy level diagram implied by equation (3.17) we get the following.

There are an infinite number of states closer and closer to zero energy. Since electric dipole selection rules require  $\Delta l = \pm 1$ , but each  $n$  has  $l = 1, \dots, n-1$  transitions that are allowed between at least *some* of the degenerate states of every pair of principal quantum numbers.

The frequencies will be given by

$$\nu = \frac{\Delta E}{h} = \frac{E_{n_u} - E_{n_l}}{h} = \frac{R}{h} \left( \frac{1}{n_l^2} - \frac{1}{n_u^2} \right) \quad (3.19)$$

where  $R = 1$  Rydberg  $\simeq 13.6$  eV ( $R/hc = 109,677.577 \text{ cm}^{-1}$ ). For any chosen  $n_l$ , there will be a series of lines with  $n_u = n_l + 1, \dots$  whose frequencies tend toward  $\nu = R/hn_l^2$  as  $n_u$  tends to  $\infty$ . This is the series limit.

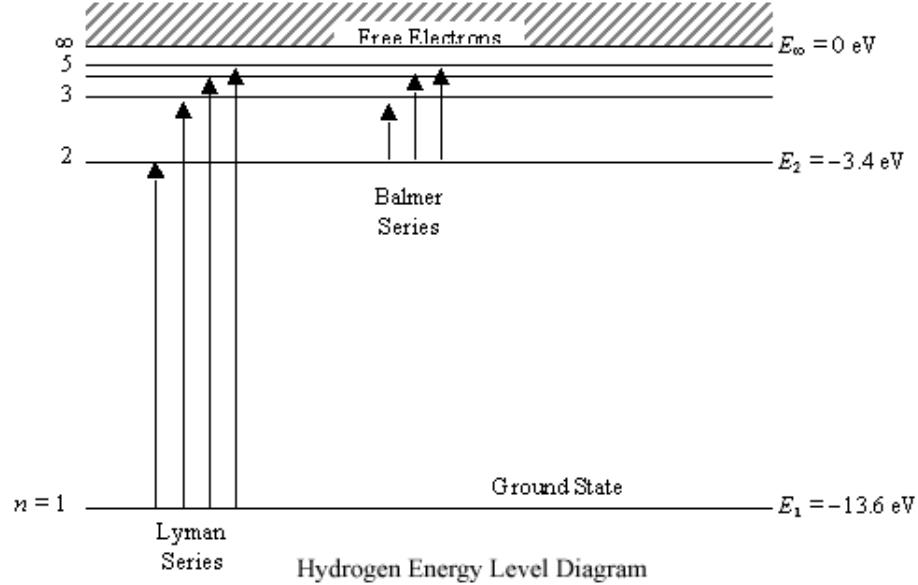


Figure 3.1: The H atom energy level diagram with a few transitions shown.

$n_l$	Name	Wavelength region	Notes
1	Lyman	Ultraviolet	$n_u = 2 \Rightarrow \text{Ly}\alpha, n_u = 3 \Rightarrow \text{Ly}\beta, \dots$
2	Balmer	optical	$n_u = 3 \Rightarrow \text{H}\alpha, n_u = 4 \Rightarrow \text{H}\beta, \dots$
3	Paschen	near-infrared	$\text{P}_\alpha$
4	Brackett	near-infrared	$\text{Br}_\alpha$
5	Pfund	mid-infrared	$\text{Pf}_\alpha$
6	Humphries	mid-infrared	$\text{Hu}_\alpha$

When the spectra of alkali atoms were observed, some similarities to the hydrogen spectrum were observed. A series analogous to the Lyman series, but displaced to longer wavelengths, was seen. This was called the *principal* series. Additional series were seen and the three most intense were called *sharp*, *diffuse*, and *fundamental* (the latter is also called the Bergmann series). For reasons that are not apparent in the context of the H atom, these series are associated with different final and initial  $l$  values. Hence, we have the association:

$l$	spectroscopic symbol		
0	$s, S$	for sharp	
1	$p, P$	for principal	
2	$d, D$	for diffuse	
3	$f, F$	for fundamental	(3.20)
4	$g, G$		
5	$h, H$		
	:		

So we could label the states of H in two ways: giving  $n, \ell, m$  or by using  $q$  to denote  $s, p, d, \dots$

$n$	$\ell$	$m$	or	$nq_m$
1	0	0	=	$1s_0$
2	1	1	=	$2p_1$
3	2	1	=	$3d_1$

(3.21)

The behavior of the radial and angular wave functions for these states is illuminating and should be studied by reference to figures in some standard book on the subject. We will characterize these states by some expectation values:

$$\langle r_{nlm} \rangle = \langle \psi_{nlm} | r | \psi_{nlm} \rangle \quad (3.22)$$

“It can be shown” that

$$\langle r_{nlm} \rangle = \frac{a_0}{2Z} [3n^2 - l(l+1)] \quad (3.23)$$

Thus the average radial distance of the electron from the nucleus increases with  $n$  for fixed  $l$  and decreases with  $l$  for fixed  $n$ .

$n$	$\ell$	$\langle r_{nlm} \rangle$
$1s$	1	$0$
$2s$	2	$0$
$2p$	2	$1$

(3.24)

The next figure shows a few sequences of lines with “Bohr” radii; the reader can judge if the distances are realistic.

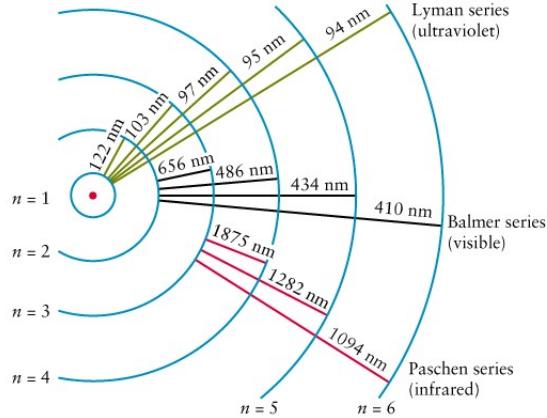


Figure 3.2: Bohr model with dubious distances, but showing more transitions.

We already know that

$$\begin{aligned} \langle L^2 \rangle_{nlm} &= l(l+1)\hbar^2 \\ \langle L_z \rangle_{nlm} &= m\hbar \end{aligned} \quad (3.25)$$

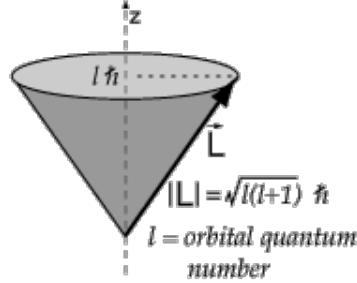


Figure 3.3: The vector model for angular momentum

since these are constants of the motion. But

$$\langle L_x \rangle_{nlm} = 0 = \langle L_y \rangle_{nlm} \quad (3.26)$$

These relations lead to the vector model in which  $L$  is represented as a vector precessing about the  $z$  axis, so that  $L_x$  and  $L_y$  average zero. For example,  $l = 2$  has 5 possible projections on the  $z$  axis

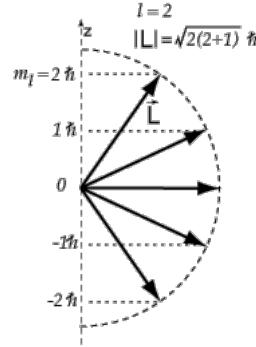


Figure 3.4: The vector model for  $l = 2$  with 5 possible projections

A useful concept at this point is the probability flux density,  $\vec{j}(\vec{r}, t)$  defined so that  $\vec{j}(\vec{r}, t) \cdot d\vec{a}$  is the probability per unit time that the particle passes through the element of surface area  $d\vec{a}$ , where

$$d\vec{a} = \hat{n}da, \quad \hat{n} \text{ is a unit vector } \perp \text{ to } d\vec{a} \quad (3.27)$$

Then for a particle of mass  $\mu$ ,

$$\vec{j}(\vec{r}, t) = \frac{-i\hbar}{2\mu} [\Psi^*(\vec{r}, t) \vec{\nabla} \Psi(\vec{r}, t) - \Psi(\vec{r}, t) \vec{\nabla} \Psi^*(\vec{r}, t)] \quad (3.28)$$

Since the two quantities in brackets are complex conjugates we get

$$\vec{j}(\vec{r}, t) = \frac{-i\hbar}{2\mu} 2i\text{Im} [\Psi^*(\vec{r}, t) \vec{\nabla} \Psi(\vec{r}, t)] \quad (3.29)$$

If we recall that

$$\Psi_{n\ell m}(\vec{r}, t) = R_{n\ell}(r)Y_{\ell m}(\theta, \varphi)e^{-iE_n t/\hbar} \quad (3.30)$$

and express  $\vec{\nabla}$  in spherical coordinates as

$$\vec{\nabla} = \hat{r}\frac{\partial}{\partial r} + \hat{\theta}\frac{1}{r}\frac{\partial}{\partial \theta} + \hat{\varphi}\frac{1}{rsin\theta}\frac{\partial}{\partial \varphi} \quad (3.31)$$

We find that only the function expressing the  $\varphi$  dependence has an imaginary part:

$$Y_{\ell m}(\theta, \varphi) = \Theta_{\ell m}(\theta)\frac{1}{\sqrt{2\pi}}e^{im\varphi}; \quad \Theta_{\ell m}(\theta) \text{ is real; } R_{n\ell}(r) \text{ is real.} \quad (3.32)$$

so

$$\begin{aligned} \vec{j}(\vec{r}, t) &= \frac{\hbar}{\mu} Im \left[ \text{Realstuff} + \frac{\hat{\varphi}im}{rsin\theta} \Psi^* \Psi \right] \\ &= \hat{\varphi} \frac{m\hbar}{\mu rsin\theta} \rho_{n\ell m}(\vec{r}) \end{aligned} \quad (3.33)$$

where  $\rho_{n\ell m} \equiv \Psi^* \Psi$ .

The only “flow” of probability is azimuthal and the flow is proportional to  $m$ . The sense of flow is increasing  $\varphi$  if  $m > 0$ .

The analogy of this probability flux to our classical notion of the electron *orbiting* around the nucleus is obvious, though caution must be exercised to avoid too literal an interpretation. We can use it however to develop the notion of an *orbital* magnetic moment caused by the “orbiting” electron.

The contribution from volume element  $dv$  to the orbital magnetic moment is

$$d\vec{M} = \frac{1}{2c} \vec{r} \mathbf{x} \left[ -e \vec{j}(\vec{r}, t) \right] dv \quad (3.34)$$

Since  $\vec{p} = -i\hbar \vec{\nabla}$ , we can write  $\vec{j}(\vec{r}, t)$  as

$$\begin{aligned} \vec{j}(\vec{r}, t) &= \frac{1}{2\mu} [\Psi^*(\vec{r}, t) \vec{p} \Psi(\vec{r}, t) - \Psi(\vec{r}, t) \vec{p} \Psi^*(\vec{r}, t)] \\ &= \frac{1}{\mu} \text{Re} [\Psi^*(\vec{r}, t) \vec{p} \Psi(\vec{r}, t)] \end{aligned} \quad (3.35)$$

and thus

$$d\vec{M} = \frac{-e}{2\mu c} \text{Re} [\Psi^*(\vec{r}, t) (\vec{r} \mathbf{x} \vec{p}) \Psi(\vec{r}, t)] dv \quad (3.36)$$

where  $\vec{r} \mathbf{x} \vec{p} = \vec{L}$  and  $\int dv$  leads to

$$\vec{M} = \frac{-e}{2\mu c} \langle \Psi(\vec{r}, t) | \vec{L} | \Psi(\vec{r}, t) \rangle \quad (3.37)$$

Now if we use  $\vec{M}_\ell$  to represent the *operator* for orbital magnetic moment, we have

$$\vec{M}_\ell = \frac{-e}{2\mu c} \vec{L} \quad (3.38)$$

A more conventional notation is to use  $\vec{\mu}_m$  for magnetic moment:

$$\vec{\mu}_M = \frac{-g_\ell \mu_B}{\hbar} \vec{L} \quad ; \quad \mu_B = \frac{e\hbar}{2m_e c} \quad ; \quad g_\ell = 1 \quad (3.39)$$

where  $\mu_B$  is the Bohr magneton with value  $0.927 \times 10^{-20}$  erg/gauss. Note that the Bohr magneton in MKS units is defined without the  $1/c$  and has the value  $0.927 \times 10^{-23}$  joules/Tesla. If you expect the missing  $c$  to cause a different factor, the discrepancy is in the definition of  $e$  in the two systems. You can use either one, but you must be consistent. Both  $\vec{\mu}_M^2$  and its projection on the  $z$  axis are constants of the motion with values

$$\begin{aligned} \langle \mu_M^2 \rangle &= (g_\ell \mu_B)^2 \ell(\ell + 1) \\ \langle \mu_M \rangle_z &= -g_\ell \mu_B m \end{aligned} \quad (3.40)$$

for the one electron atom in a  $n, \ell, m$  eigenstate.

## 3.2 Spin

In analogy with the development of an orbital magnetic moment, we will now deal with the magnetic effects of the fact that electrons have spin. The need to introduce this extra degree of freedom arose from the existence of fine structure in many spectral lines; that is, they were split into multiplets. The notion of an electron spinning on its “axis” leads naturally to the idea of a spin magnetic dipole, but if this classical notion is taken too seriously, severe conceptual difficulties arise. Essentially, spin is a relativistic effect and arises naturally in a relativistic quantum mechanics, but we will simply accept its existence and examine the consequences. Since the spin is intrinsic to the electron, it does not depend on the ordinary spatial coordinates which give the location of the electron. Instead it must depend on some “internal coordinates”. It turns out that the projection of the spin on the  $z$  axis,  $m_s$ , can take on only two values for the electron,  $\pm\hbar/2$ . So, we need a two component vector as the wave function. We introduce the operators for the square of the spin  $S^2$  and the projection on the  $z$  axis  $S_z$  such that

$$\begin{aligned} S^2 \chi_{sm_s} &= s(s+1)\hbar^2 \chi_{sm_s} \\ S_z \chi_{sm_s} &= m_s \hbar \chi_{sm_s}, \end{aligned} \quad (3.41)$$

where  $\chi_{sm_s}$  are the 2 component column vectors which are eigenfunctions of the spin operators. To get an explicit form, we adopt Pauli’s convention in which we write the spin operators as

$$\vec{S} = \frac{1}{2}\hbar(\hat{x}\sigma_x + \hat{y}\sigma_y + \hat{z}\sigma_z), \quad (3.42)$$

where

$$\sigma_x \equiv \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y \equiv \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z \equiv \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (3.43)$$

Then

$$S^2 = S_x^2 + S_y^2 + S_z^2 = \frac{3}{4}\hbar^2 \mathbf{1} \quad (3.44)$$

since  $\sigma_i^2 = \mathbf{1}$ , the unit matrix for each  $i$ .

In this convention, the eigenvalues are  $s = \frac{1}{2}$ :  $m_s = \pm\frac{1}{2}$  and the eigenfunctions are

$$\chi_{\frac{1}{2}, \frac{1}{2}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad \chi_{\frac{1}{2}, -\frac{1}{2}} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (3.45)$$

with alternate notations of

$$\begin{array}{cc} \text{“up”} & \text{“down”} \\ \alpha & \beta \\ \uparrow & \downarrow \end{array} \quad (3.46)$$

$$S_z \alpha = \frac{1}{2}\hbar \alpha; \quad S_z \beta = -\frac{1}{2}\hbar \beta \quad (3.47)$$

The  $S_i$  satisfy commutation relations exactly analogous to those of orbital angular momentum. We can define a matrix element using matrix multiplication; e.g.,

$$\langle \chi | S^2 | \chi' \rangle \equiv \chi^\dagger S^2 \chi', \quad (3.48)$$

when  $\chi^\dagger$  is a row vector - the adjoint of  $\chi$ ,  $[\chi^\dagger = (\chi^t)^*]$ . For example, if  $\chi = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  ,  $\chi^\dagger = (1 \ 0)$  .

One peculiarity of spin is that its magnetic moment is twice what one might expect:

$$\vec{\mu}_s = \frac{-g_s \mu_\beta}{\hbar} \vec{S} \quad ; \quad g_s = 2 \quad (3.49)$$

(In fact, a more precise value for  $g_s = 2.0023193043737$ , one of the most precisely known constants of nature. The corrections come from QED where  $g_s = 2(1 + \alpha/2\pi + \dots)$ .

### 3.3 The Hydrogen Atom with Spin

To incorporate the fact that the electron has spin into our model for the hydrogen atom, we must represent a wave function which depends on both spatial and spin “coordinates”. We can do this with a direct product space. This is a way of creating a higher dimensional space from two lower dimensional spaces; in this case the combined space has both ordinary spatial coordinates and spin “coordinates”, represented by spin matrices. We “multiply” the spatial wave functions by the spin wave functions:

$$\psi_{n\ell m_\ell m_s}(\vec{r}) = \psi_{n\ell m_\ell}(\vec{r}) \chi_{sm_s}, \quad (3.50)$$

with  $\psi_{n\ell m_\ell}(\vec{r})$  given by (3.4) and  $\chi_{sm_s}$  given by (3.45). The  $\psi_{n\ell m_\ell m_s}(\vec{r})$  are 2 component wave functions that are sometimes called Pauli wave functions or spin orbitals. Note that we now have four quantum numbers and the degeneracy of  $m_s = \pm 1/2$  makes the total degeneracy  $2n^2$ .

It will turn out to be useful to take linear combinations of the  $\psi_{n\ell m_\ell m_s}(\vec{r})$  that are eigenfunctions of the total angular momentum operator

$$\vec{J} = \vec{L}\mathbf{1} + \vec{S} \quad ; \quad \mathbf{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (3.51)$$

and its projection on the  $z$  axis  $J_z$ . These new wave functions are

$$\psi_{n\ell jm_j}(\vec{r}) = \sum_{m_\ell m_s} \langle \ell \frac{1}{2} m_\ell m_s | \ell \frac{1}{2} j m_j \rangle \psi_{n\ell m_\ell m_s}(\vec{r}), \quad (3.52)$$

where  $\langle \ell \frac{1}{2} m_\ell m_s | \ell \frac{1}{2} j m_j \rangle$  is a Clebsch-Gordan coefficient, which is zero unless  $\Delta(\ell, \frac{1}{2}, j)$  (triangle inequality, see §2.6) and  $m_j = m_\ell + m_s$ . As applied to this particular case

$$\begin{aligned} j &= \ell \pm 1/2 & \ell &\neq 0 \\ &= \frac{1}{2} & \ell &= 0 \end{aligned} \quad (3.53)$$

and  $m_j = -j, -j+1, \dots, 0, 1, \dots, j$ .

Explicit expressions for the components of  $\vec{J}$  are

$$\begin{aligned} J_x &= i\hbar \left( \sin \varphi \frac{\partial}{\partial \theta} + \cot \theta \cos \varphi \frac{\partial}{\partial \varphi} \right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ J_y &= -i\hbar \left( \cos \varphi \frac{\partial}{\partial \theta} - \cot \theta \sin \varphi \frac{\partial}{\partial \varphi} \right) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ J_z &= -i\hbar \frac{\partial}{\partial \varphi} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \quad (3.54)$$

Because  $\vec{J}$  is the *total* angular momentum, it commutes with the Hamiltonian in the absence of external torques, and  $j, m_j$  are good quantum numbers

$$J^2 \psi_{n\ell jm_j}(\vec{r}) = j(j+1)\hbar^2 \psi_{n\ell jm_j}(\vec{r}), \quad (3.55)$$

$$J_z \psi_{n\ell jm_j}(\vec{r}) = m_j \hbar \psi_{n\ell jm_j}(\vec{r}). \quad (3.56)$$

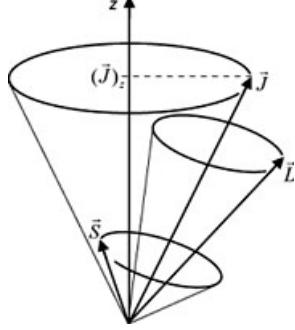


Figure 3.5: Shows  $\vec{J}$  precessing around the  $z$  axis,  $\vec{L}$  and  $\vec{S}$  precessing around  $\vec{J}$ .

We can visualize the situation with the aid of the vector model. We add  $\vec{L}$  and  $\vec{S}$  to get  $\vec{J}$ , and we can think of  $\vec{L}$  and  $\vec{S}$  precessing around  $\vec{J}$  (recall Fig 3.3) and  $\vec{J}$  precesses around  $\hat{z}$ .

Let us first consider the interaction between spin and orbital angular momentum. The spin of the electron means that it has a magnetic dipole moment (3.49), and this will have an energy in a magnetic field of  $-\vec{\mu}_s \cdot \vec{B}$ . To see where  $\vec{B}$  comes from, imagine transforming to the rest frame of the electron. From this vantage point, the nucleus is orbiting us, producing a current and hence a magnetic field. We can also estimate  $\vec{B}$  from

$$\vec{B} = -\frac{v}{c} \mathbf{x} \vec{E} = -\frac{\vec{p}}{m_e c} \mathbf{x} \frac{Ze}{r^3} \vec{r} = \frac{Ze}{m_e c r^3} \vec{L} \quad (3.57)$$

for the Coulomb potential  $Ze/r^2$ . We can see that there is an interaction potential which we can write as a perturbation, using (3.49).

$$H^{(1)} = \frac{g_s \mu_B}{\hbar} \vec{S} \cdot \frac{Ze}{m_e c r^3} \vec{L} \quad (3.58)$$

However, there is relativistic effect in transforming back to the proton's rest frame of exactly  $-\frac{1}{2} H^{(1)}$  (called Thomas precession) so we write the net spin orbit interaction as

$$H_{so} = \frac{g_s \mu_B}{2} \vec{S} \cdot \frac{Ze}{m_e c r^3} \vec{L} \quad (3.59)$$

or since  $g_s \simeq 2$ ,

$$H_{so} = \frac{\mu_B}{\hbar} \frac{Ze}{m_e c r^3} \vec{S} \cdot \vec{L} = \frac{Ze^2}{2m_e^2 c} \frac{1}{r^3} \vec{S} \cdot \vec{L} \quad (3.59b)$$

( $\mu_B = e\hbar/2m_e c$ ) We are justified in treating  $H_{so}$  as a perturbation, as we can verify from an estimate of its magnitude:

$$\left\langle \frac{1}{r^3} \right\rangle_{n\ell m} = \frac{Z^3}{a_0^3 n^3 \ell (\ell + \frac{1}{2})(\ell + 1)}, \quad (3.60)$$

which is proportional to  $\frac{1}{n^3 a_0^3}$  for hydrogen. Since  $\langle \vec{S} \cdot \vec{L} \rangle \simeq \hbar^2$ ,  $H_{so} \simeq 10^{-4}$  eV for  $Z = 1$  (but note  $H_{so} \propto Z^3$ , so it grows rapidly with  $Z$ ). Note also that  $\vec{S} \cdot \vec{L}$  does not commute with  $L_z$  or  $S_z$ ,

so  $\vec{S} \cdot \vec{L}$  is not diagonal in the representation (3.50), but it is diagonal in the representation(3.52). To see this, note that we can rewrite  $\vec{S} \cdot \vec{L}$  through a neat trick. From (3.51), we see that

$$\begin{aligned}\vec{J}^2 &= \vec{L}^2 + 2\vec{L} \cdot \vec{S} + \vec{S}^2 \\ \Rightarrow \vec{L} \cdot \vec{S} &= \frac{J^2 - L^2 - S^2}{2}\end{aligned}\tag{3.61}$$

and  $\psi_{n\ell jm_j}$  are eigenfunctions of  $L^2, S^2$ , and  $J^2$ .

So, we can write our perturbation as

$$H_{so} = \frac{Ze^2}{4m_e^2 c^2} \frac{1}{r^3} (J^2 - S^2 - L^2)\tag{3.62}$$

Since we have diagonalized our perturbation Hamiltonian<sup>1</sup> we can use non degenerate perturbation theory.

$$E_{n\ell jm_j} \simeq E_n^{(0)} + \langle \psi_{n\ell jm_j}^{(0)} | H_{so} | \psi_{n\ell jm_j}^{(0)} \rangle\tag{3.63}$$

Defining the correction term to be  $\Delta E_{so}$ , and noting that

$$(J^2 - S^2 - L^2) | \psi_{n\ell jm_j}^{(0)} \rangle = \left( j(j+1) - \frac{3}{4} - \ell(\ell+1) \right) \hbar^2 | \psi_{n\ell jm_j}^{(0)} \rangle$$

we get

$$\Delta E_{so} = \frac{Ze^2}{4m_e^2 c^2} \left\langle \frac{1}{r^3} \right\rangle \left( j(j+1) - \frac{3}{4} - \ell(\ell+1) \right)\tag{3.64}$$

and using (3.60), and  $E_n^{(0)} = -\frac{Z^2 e^2}{2a_0 n^2}$  (equation 3.17)

$$\Delta E_{so} = \frac{Z^2 |E_n^{(0)}| \alpha^2}{n} \frac{[(j(j+1) - \ell(\ell+1) - \frac{3}{4})]}{2\ell(\ell+\frac{1}{2})(\ell+1)}\tag{3.65}$$

where  $\alpha \equiv e^2/\hbar c \simeq 1/137$  is the fine structure constant. So we see that the splitting is small ( $\sim \alpha^2$ ), as promised, in comparison to the unperturbed energy  $E_n^{(0)}$ . Note that  $\Delta E_{so}$  appears to blow up for  $\ell = 0$ , but in fact equation 3.65 is only valid for  $\ell \neq 0$ . Since  $\vec{B} = 0$  for  $\ell = 0$ ,  $\Delta E_{so} = 0$  for  $\ell = 0$ . Note that  $L, S$  precess around  $J$  at  $E_{so}/\hbar \sim 10^{10}$  Hz for H.

---

<sup>1</sup> “Diagonalizing our perturbation Hamiltonian” means we have constructed a basis set and rearranged our  $H^{(1)}$  so that the basis vectors are eigenstates of the perturbation [ $J^2, S^2, L^2$  all have  $|n\ell jm_j\rangle$  as eigenstates].

Now, if one compares 3.65 to the actual hydrogen energy level diagram, they will *not* agree. That is because we have not been self-consistent. As is obvious from equations 3.57 and 3.39 the spin-orbit correction is a relativistic correction, of order  $(\frac{v}{c})^2$ , but there are other corrections of the same order. As discussed in Appendix 7 of **Bransden and Joachain**, a better treatment of hydrogen is given by Dirac's relativistic theory of the electron. Besides predicting the existence of the positron, this theory has an exact solution for a single electron, moving relativistically in a Coulomb potential. For our purposes, we only consider terms of order  $(\frac{v}{c})^2$ . (To the lowest order in  $\frac{v}{c}$ , the Dirac equation introduces no new effects). There are three terms, one of which is the spin-orbit perturbation, which we developed in another way. One term is a relativistic correction to the kinetic energy. Relativistically for the electron, but non-relativistically for the nucleus, the kinetic energy is

$$T = E - mc^2 = \sqrt{p_e^2 c^2 + m^2 c^4} - mc^2 + \frac{p_N^2}{2m_N}. \quad (3.66)$$

If  $E \simeq mc^2$ , we can expand the expression under the square root:

$$\frac{p_e^2}{2m} - \frac{1}{8} \frac{(p^2)^2}{m^3 c^2} + \dots + \frac{p_N^2}{2m_N} = \frac{p^2}{2\mu} - \frac{1}{8} \frac{(p^2)^2}{m^3 c^2} + \dots \quad (3.67)$$

Since  $\frac{p_e^2}{2m} - \frac{p_N^2}{2m_N} = \frac{p^2}{2\mu}$ .

So, we get our reduced mass as before, but with a correction term  $\left(-\frac{p^4}{8m^3 c^2}\right)$ .

There is also a relativistic correction to the potential energy, called the Darwin term. It is  $\frac{\pi\hbar^2}{2m^2 c^2} Ze^2 \delta(\vec{r})$ . Since only  $\ell = 0$  states have non-vanishing values at  $\vec{r} = 0$ , this term only affects  $\ell = 0$  states.

When these two additional terms are evaluated, one finds corrections of

$$\Delta E_1 = -E_n^{(0)} \frac{Z^2 \alpha^2}{n^2} \left[ \frac{3}{4} - \frac{n}{\ell + \frac{1}{2}} \right] \quad (\text{Relativistic Correction to Kinetic Energy}) \quad (3.68)$$

and

$$\Delta E_3 = -E_n^{(0)} \frac{Z^2 \alpha^2}{n} \quad \ell = 0 \quad (\text{Darwin Term}) \quad (3.69)$$

Combining these with the spin-orbit effect  $\Delta E_{so} \equiv \Delta E_2$ , we get

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

or

$$E_{nj} = E_n^{(0)} \left[ 1 + \frac{(Z\alpha)^2}{n^2} \left( \frac{n}{j + \frac{1}{2}} - \frac{3}{4} \right) \right] \quad (3.70)$$

For comparison, the exact solution of the Dirac equation with the Coulomb potential  $\frac{Ze^2}{r}$  is

$$E_{nj}^{\text{exact}} = mc^2 \left\{ \left[ 1 + \frac{Z\alpha}{n - j - \frac{1}{2} + \left[ (j + \frac{1}{2})^2 - Z^2 \alpha^2 \right]^{\frac{1}{2}}} \right]^{-\frac{1}{2}} - 1 \right\} \quad (3.71)$$

expanding this and keeping terms through order  $(Z\alpha)^2$  will give an identical result to what we have.

Note that both 3.70 and 3.71 depend on  $n$  and  $j$ , but *not* on  $\ell$ , the dependence on  $\ell$  from the spin-orbit effect (equation 3.65) having been cancelled by the dependence in (3.68). (This takes “a bit of algebra” to see!). So, for example,  $2s_{\frac{1}{2}}$  and  $2p_{\frac{1}{2}}$  remain degenerate. These levels are actually split when the self-energy of the electron is considered. The splitting (the Lamb shift) is only  $\nu (= \frac{\Delta E}{\hbar}) = 1057.7 \text{ MHz}$  ( $3.5 \times 10^{-2} \text{ cm}^{-1}$ ).

### 3.3.1 The Zeeman Effect

The relativistic effects (and even the Lamb shift) do not break the degeneracy associated with the projection of  $j$  on the  $z$  axis (the  $2j+1$  possible  $m$  values, see Fig 3.4). This final degeneracy remains unless an external field is applied. If the field is not too strong, the resulting splitting is called the Zeeman effect. This effect is about the only way we have to measure magnetic field strengths in astronomy, so we consider it here. Assume a magnetic field in the  $\hat{z}$  direction:

$$\vec{B} = B_z \hat{z} \quad (3.72)$$

The perturbation is

$$H^{(1)} = -\mu_z B_z \quad (-\vec{\mu} \cdot \vec{B}) \quad (3.73)$$

where

$$\vec{\mu} = \vec{\mu}_M + \vec{\mu}_s = \frac{-\mu_B}{\hbar} (g_\ell \vec{L} + g_s \vec{S}) = \frac{-\mu_B}{\hbar} (\vec{L} + 2\vec{S}) \quad (3.74)$$

with  $g_\ell = 1$  and  $g_s = 2$  (equation 3.40 and 3.49).

The  $z$  component is then

$$\mu_z = \frac{-\mu_B}{\hbar} (L_z + 2S_z) = \frac{-\mu_B}{\hbar} (J_z + S_z) \quad (3.75)$$

We need to take the matrix element of the perturbation for a state in the  $|n, \ell, j, m_j\rangle$  representation to see the effects on the energy levels, given by equation 3.70 or 3.71. This is trivial for  $J_z$  :  $J_z |n\ell jm_j\rangle = m_j \hbar |n\ell jm_j\rangle$ . But  $S_z$  is more complicated, since  $S_z$  does not commute with  $\vec{L} \cdot \vec{S}$ . For weak fields, we treat the external field as a perturbation after solving the spin orbit (internal) field effects. Using either of two approaches (see **Bransden and Joachain**, pg 215-219), one finds

$$\langle n\ell jm_j | S_z | n\ell jm_j \rangle = m_j \hbar \left[ \frac{j(j+1) + s(s+1) - \ell(\ell+1)}{2j(j+1)} \right] \quad (3.76)$$

So the energy shift (see eq 1.62) is

$$\Delta E_{mj} = g \mu_B B_z m_j \quad (3.77)$$

where  $g$  is the Landé  $g$  factor:

$$g = 1 + \frac{j(j+1) + s(s+1) - \ell(\ell+1)}{2j(j+1)} \quad (3.78)$$

or for the case of  $s = \frac{1}{2}$ ,

$$\Delta E_{mj} = \frac{2\ell+2}{2\ell+1} \mu_B B_z m_j \quad , \quad j = \ell + \frac{1}{2} \quad (\text{e.g., the } p_{3/2} \text{ state})$$

$$= \frac{2\ell}{2\ell+1} \mu_B B_z m_j \quad , \quad j = \ell - \frac{1}{2} \quad (\text{e.g., the } p_{1/2} \text{ state}) \quad (3.79)$$

The splitting is just proportional to  $B_z$  and  $M_j$ , so one gets a series of evenly spaced levels. For example,

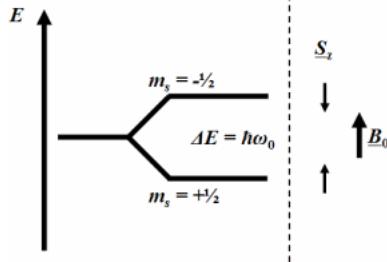


Figure 3.6: Splitting of  $p_{1/2}$  in a magnetic field.

Transitions between these states and some other states will now depend on  $m_j$ , and the  $\Delta m$  selection rules will be important, as will the polarization of the light. In fact, lines will be split into groups with different polarizations and this is how the Zeeman effect is detected.

If the magnetic field is strong enough, this perturbation will become comparable to the spin-orbit perturbation and the derivation of 3.72 – 3.79 will no longer be valid. As the field increases, first one has the Paschen-Back effect (**Bransden and Joachain**, pg. 214) and then the strong field case (**Bransden and Joachain**, pg. 210). In the latter case  $\ell$  and  $s$  get decoupled and each couples instead to the external magnetic field. At still higher field strengths, quadratic terms in  $B$  enter. These are actually relevant for some magnetic white dwarfs (e.g., AM Herculis), when  $B = 10^7 - 10^8$  Gauss (see Angel Ap. J. 216, 1 and Angel, Liebert, and Stockman, Ap. J. 292, 260).

### 3.4 The Hydrogen Atom with Nuclear Spin

Just as fine structure arises from interaction of the electron spin with a magnetic field (at least the spin-orbit part of it), so hyperfine structure arises from the interaction of the spin of the *nucleus* with a magnetic field. One way to look at this is that the nuclear spin and associated magnetic moment creates a magnetic field, causing an internal Zeeman splitting of the two spin states of the electron. It is conventional to denote the nuclear spin by  $\vec{I}$ . Then by analogy with (3.49)

$$\vec{\mu}_N = \frac{g_I \mu_N \vec{I}}{\hbar} \quad (3.80)$$

where

$$\mu_N = \frac{e\hbar}{2m_p c} = \frac{m_e}{m_p} \mu_B \quad (3.81)$$

is called the nuclear magneton, and  $g_I$  is the nuclear  $g$  factor. One tricky point is that  $g_I$  is not what usually appears in the tables of nuclear magnetic moments; rather it is  $\mu_I$ , the magnitude of  $g_I \vec{I}$ , where  $I$  is the quantum number of  $\vec{I}$  (usually just called the nuclear spin).

So one could also write

$$\vec{\mu}_N = \frac{\mu_I}{I} \frac{\vec{I}}{\hbar} \mu_N \quad (3.82)$$

Naively, one might expect a spinning charge  $Ze$  to produce a magnetic moment of  $g_I \frac{Z_e}{2m_N c} \vec{I}$ , but nuclei are not simply one big charge  $Ze$ . In particular the  $Z$  is not appropriate and  $g_I \neq 2$  in general as it is for the electron. For the proton  $\mu_I = 2.79284734 \mu_N$ , so  $g_I = \frac{\mu_I}{I} = 5.58569468 \mu_N$ , since  $I = \frac{1}{2}$ .

This particular number can be understood as an additional contribution from virtual dissociation of the proton into a neutron and a  $\pi^+$  meson (e.g. Segré's book, **Nuclei and Particles**). This also explains how the neutron (ostensibly *neutral*) has a magnetic moment ( $\mu_I = -1.9130427$ ); when  $n \rightarrow p + \pi^-$  part of the time, there is a charge to create a magnetic moment. Note the accuracy of these numbers; they can only come from spectroscopy, where very accurate frequencies can be measured. The values here come from a compilation by N. J. Stone in Atomic Data and Nuclear Data Tables 90 (2005) 75.

The magnetic dipole moment of the nucleus creates a magnetic field which can then interact with the magnetic dipole moment of the electron. The magnetic field from a point magnetic dipole can be computed from the vector potential:

$$\vec{A}(\vec{r}) = -\frac{1}{4\pi} (\vec{\mu}_N \mathbf{x} \vec{\nabla}) \frac{1}{r} \quad (3.83)$$

$$\vec{B} \equiv \vec{\nabla} \times \vec{A} = \frac{-\vec{\mu}_N}{4\pi} \nabla^2 \frac{1}{r} + \frac{1}{4\pi} \vec{\nabla} (\vec{\mu}_N \cdot \vec{\nabla}) \frac{1}{r} \quad (3.84)$$

The  $\vec{B}$  field from the nucleus acts like a Zeeman effect on the electron spin (compare to 3.73), so the perturbing Hamiltonian is

$$H_s^{(1)} = -\vec{\mu}_s \cdot \vec{B} = \frac{g_s \mu_B}{\hbar} \vec{S} \cdot \vec{B} \quad (3.85)$$

(see eq. 3.50)

There is also an interaction ( $H_L^{(1)}$ ) with the orbital angular momentum of the electron (essentially a nuclear spin-electron orbit interaction). Just as we coupled  $\vec{L}$  and  $\vec{S}$  to form  $\vec{J}$  for the electron (3.52), we now couple  $\vec{J}$  and  $\vec{I}$  to make  $\vec{F}$ , the operator for the total angular momentum including the spin of the nucleus:

$$\vec{J} + \vec{I} = \vec{F} \quad (3.86)$$

Then we construct eigenstates  $|\ell s j I F M_F\rangle$  of this total angular momentum.

A rather tedious set of calculations then produces the energy shift, the expectation value of  $H_s^{(1)} + H_L^{(1)}$  in these states:

$$\langle H^{(1)} \rangle = \frac{h\nu_0}{n^3} \left\{ \frac{F(F+1) - I(I+1) - j(j+1)}{j(j+1)(2\ell+1)} \right\} \quad (3.87)$$

where  $\nu_0 = \frac{m_e}{m_p} \alpha^2 c R(\text{cm}^{-1}) Z^3 \frac{\mu_I}{I}$  and  $R(\text{cm}^{-1})$  is the Rydberg constant in  $\text{cm}^{-1}$ , and  $\alpha$  is the fine structure constant.

Of particular interest to us is the case of the ground state of hydrogen,  ${}^2s\frac{1}{2}(n=1, \ell=0, s=\frac{1}{2}, j=\frac{1}{2})$ . Then,

$$\langle H^{(1)} \rangle = h\nu_0 \left\{ \frac{F(F+1) - I(I+1) - \frac{3}{4}}{\frac{3}{4}} \right\} \quad (3.88)$$

For normal hydrogen, with a proton nucleus,  $I = \frac{1}{2}$ , so  $F$  can be 0 or 1.  $F = 1$  will lie at higher energy and the difference will be

$$\Delta E_{hp} = \frac{4h\nu_0}{3} \left\{ (1 \cdot 2 - \frac{3}{2}) \right\} - \left\{ (0 \cdot 1 - \frac{3}{2}) \right\} = \frac{8}{3}h\nu_0 \quad (3.89)$$

If we plug the precise numbers into the equation for  $\nu_0$  we get

$$\begin{aligned} \nu_0 &= \frac{1}{1863.09} \left( \frac{1}{137.036} \right)^2 (2.99792 \times 10^{10})(109,677.575)5.5883 \\ &= 5.32911 \times 10^8 \text{Hz} \end{aligned}$$

so

$$\frac{\Delta E_{hp}}{h} = \frac{8}{3}\nu_0 = 1421.09 \times 10^6 \text{Hz} \quad \text{or} \quad \lambda \sim 21 \text{cm} \quad (3.90)$$

This is the famous 21cm line which has been measured with extreme precision in hydrogen masers to be

$$\nu_{exp} = 1420.4057518 \times 10^6 \text{Hz}$$

which differs from our calculation by  $< 0.1\%$ . The major factor in the discrepancy is the anomalous  $g$  factor of the electron (see discussion after equation 3.50).

## 3.5 The Spectrum of Hydrogen

From the previous discussion of hydrogen, we can see that hydrogen has a rich spectrum, from ultraviolet to low radio frequencies. Because of its importance in astronomy, we would like to have a fairly complete picture of the hydrogen spectrum, including the strengths of the various transitions. Let us begin with an overview with this figure.

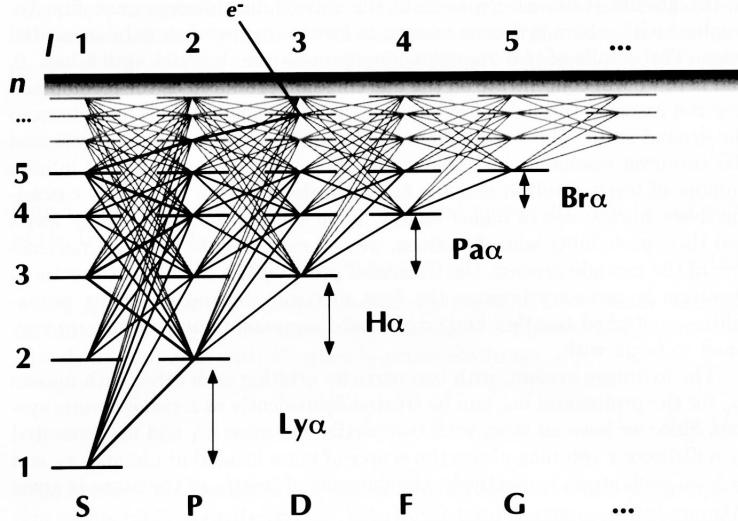


Figure 3.7: Figure from Dopita and Sutherland, **Astrophysics of the Diffuse Universe**, showing allowed transitions and the  $\Delta\ell = \pm 1$  selection rule.

This figure gives the big picture, but we must remember that levels with  $n > 1$  are split by fine structure effects, and all levels are split by hyperfine effects.

### 3.5.1 Bound-Bound ( $\Delta n \neq 0$ ) Transitions

Now let us consider the  $\Delta n \neq 0$  transitions of hydrogen – those which give rise to the various series: Lyman, Balmer, etc. We know that (see §2.8)

$$A_{ji} = \frac{64\pi^4\nu^3}{3hc^3} |\mu_{ji}|^2 = \frac{64\pi^4\nu^3}{3hc^3} |\mu_{ij}|^2 \frac{g_i}{g_j} = \frac{64\pi^4\nu^3}{3hc^3} \frac{S_{ij}}{g_j} \quad (3.91)$$

where  $S_{ij} = |\mu_{ij}|^2 g_i$ . Since  $|\mu_{ij}|^2$  includes a sum over final degenerate states, we can think of  $S_{ij}$  as also containing a sum over *initial* degenerate states so that the net result of  $\frac{S_{ij}}{g_j}$  is to *sum* over final states and *average* over initial states. This is appropriate because any given atom will be in only one of the initial states but can make a transition to any of the final states. We also have the oscillator strength  $f_{ji}$  for emission as

$$f_{ji} = \frac{8\pi^2 m_e \nu}{3he^2} |\mu_{ji}|^2 = \frac{8\pi^2 m_e \nu}{3he^2} \frac{S_{ij}}{g_j} \quad (3.92)$$

For hydrogen, we may want the oscillator strength for particular  $\ell, \ell'$  or summed over all  $\ell, \ell'$  for a given  $n, n'$  where  $n'$  is the initial state, assumed to lie above state  $n$ . We have

$$f(n', n) = \frac{1}{g_{n'}} \sum_{\ell'} g_{n' \ell'} f(n', \ell'; n, \ell) = \frac{\sum_{\ell'} g_{n' \ell'} f(n', \ell'; n, \ell)}{\sum_{\ell'} g_{n' \ell'}} \quad (3.93)$$

Let us first consider the hydrogen atom without spin. This is acceptable because the levels are barely split and the electric dipole transitions do not change spin, so the spin part basically drops out of the calculation, except for adding a degeneracy of 2 to each  $n, \ell, m_\ell$  state for the two spin states. Thus

$$g_n = 2 \sum_{\ell=0}^{n-1} (2\ell + 1) = 2n^2 \quad g_{n, \ell} = 2(2\ell + 1). \quad (3.94)$$

Now, we know that the electric dipole operator is  $\vec{\mu} = -e\vec{r}$  where  $\vec{r}$  is now the coordinate of the electron relative to the nucleus. So  $\vec{r} = x\hat{x} + y\hat{y} + z\hat{z}$  and we know that we can express  $x, y, z$  in terms of  $r$  and spherical harmonics. In fact, the transition strengths derived for the rigid rotor can be applied directly to the hydrogen atom, except that we now must deal with the radial part of the wave function also. Also, we have the selection rule  $\Delta\ell = \pm 1$  in addition to  $\Delta j = 0, \pm 1$ . If we consider the transition  $n'\ell' \rightarrow n\ell$ , we have  $\ell = \ell' + 1$  or  $\ell = \ell' - 1$ . Then from the rigid rotor results (maybe done in homework), using  $i = (n'\ell'), j = (n, \ell' - 1)$ ,

$$\left| \mu_{ji} \right|^2 = \sum_m \langle n, \ell' + 1, m | \mu | n', \ell', m' \rangle = I_r^2 \frac{(\ell' + 1)}{(2\ell' + 1)} \quad (3.95)$$

where

$$I_r = - \int_0^\infty R_{n\ell'+1}(r) e r R_{n', \ell'}(r) r^2 dr \quad (3.96)$$

In a similar way, if  $\ell = \ell' - 1$

$$\left| \mu_{ij} \right|^2 = I_r^2 \frac{\ell'}{(2\ell' + 1)} \quad (3.97)$$

or in general,

$$\left| \mu_{ij} \right|^2 = I_r^2 \frac{\max(\ell, \ell')}{(2\ell' + 1)} \quad (3.98)$$

Since  $S_{ij} = g_i |\mu_{ij}|^2$ ,

$$S(n', \ell'; n\ell) = 2(2\ell' + 1) I_r^2 \frac{\max(\ell, \ell')}{(2\ell' + 1)} = 2I_r^2 \max(\ell, \ell') \quad (3.99)$$

It is common to rewrite  $I_r$  by changing variables to  $r \rightarrow r/a_0$  and writing  $P_{n\ell}(r) = rR_{n\ell}(r)$  since then  $\int_0^\infty P_{n\ell}(r) P_{n'\ell'}(r) dr = 1$ . Then,

$$\begin{aligned} I_r^2 &= e^2 a_0^2 \left( \int_0^\infty P_{n'\ell'}(r) r P_{n\ell}(r) dr \right)^2 \\ &= e^2 a_0^2 \sigma^2 \end{aligned} \quad (3.100)$$

where  $\sigma$  is the integral. Tables of  $\sigma^2$  for hydrogen are in Green, Rush and Chandler, *Ap. J. Suppl.*, 3, 37, 1957. A general expression for the integral was given by Gordon (Ann. Phys. 2, 1031, 1929) which is also given by Mihalas, **Stellar Atmospheres** (pg. 105).

We can now give an expression for the Einstein A for  $n\ell \rightarrow n'\ell'$  using (3.91) and (3.99):

$$A(n', \ell'; n\ell) = \frac{64\pi^4\nu^3}{3hc^3} \frac{\max(\ell, \ell')}{(2\ell + 1)} e^2 a_0^2 \sigma^2 \quad (3.101)$$

If we want the oscillator strength for an upward transition, we get

$$\begin{aligned} f(n'\ell'; n\ell) &= \frac{8\pi^2 m_e \nu}{3h c^2} \frac{2e^2 a_0^2 \sigma^2 \max(\ell, \ell')}{2(2\ell' + 1)} \\ &= \frac{1}{3} \left( \frac{1}{n'^2} - \frac{1}{n^2} \right) \frac{\max(\ell, \ell')}{(2\ell' + 1)} \left( \int_0^\infty P_{n'\ell'}(r) P_{n\ell}(r) r dr \right)^2 \end{aligned} \quad (3.102)$$

using the definitions (3.19), (3.18), and (3.9). If we want the transition strength for the entire  $n' \rightarrow n$ , we must average over initial  $\ell'$  to get  $f(n', n)$  (cf. eq. 3.93). The result is

$$\begin{aligned} f(n', n) &= \frac{1}{3n'^2} \left( \frac{1}{n'^2} - \frac{1}{n^2} \right) \left[ \sum_{\ell'=1}^{n'-1} \ell' \left( \int_0^\infty P_{n'\ell'}(r) P_{n\ell'-1}(r) r dr \right)^2 \right. \\ &\quad \left. + \sum_{\ell'=0}^{n'-1} (\ell' + 1) \left( \int_0^\infty P_{n'\ell'}(r) P_{n\ell'+1}(r) r dr \right)^2 \right] \end{aligned} \quad (3.103)$$

and

$$\begin{aligned} S(n', n) &= 2e^2 a_0^2 \left[ \sum_{\ell'=1}^{n'-1} \ell' \left( \int_0^\infty P_{n'\ell'}(r) P_{n,\ell'-1}(r) r dr \right)^2 \right. \\ &\quad \left. + \sum_{\ell'=0}^{n'-1} (\ell' + 1) \left( \int_0^\infty P_{n'\ell'}(r) P_{n,\ell'+1}(r) r dr \right)^2 \right], \end{aligned} \quad (3.104)$$

where the two terms come from the two allowed transitions,  $\ell = \ell' + 1$  and  $\ell' - 1$ . Then

$$A(n, n') = \frac{64\pi^4\nu^3}{3hc^3} \frac{S(n', n)}{2n^2} \quad (3.105)$$

For example, Lyman  $\alpha$  ( $n = 2, n' = 1$ ) has an Einstein A value of  $A = 4.7 \times 10^8 \text{ s}^{-1}$ .

There is also an explicit (and horrible) form for  $f(n', n)$  but one probably prefers to use the tables of Green, Rush, and Chandler cited above.

Another common way to express the oscillator strength is in terms of a semiclassical expression due to Kramers. This is denoted  $f_K(n', n)$  and given by

$$f_K(n', n) = \frac{32}{3\pi\sqrt{3}} \left( \frac{1}{n'^2} - \frac{1}{n^2} \right)^{-3} \frac{1}{n^3 n'^5} \quad (3.106)$$

Then

$$f(n', n) = f_K(n', n) g_I(n', n) \quad (3.107)$$

where  $g_I(n', n)$ , the Gaunt factor, is a fudge factor for the oscillator strength, which is a fudge factor for the classical Lorentz equation. The Gaunt factors are often near unity and this formation is

actually useful in dealing with bound-free transitions. The Kramers form also makes it easy to see how fast the oscillator strengths decrease as you go up a *series* (constant  $n'$ , increasing  $n$ ) or to higher series (increasing  $n'$ , with  $n - n'$  constant). Think about the astrophysical consequences of this fact.

The Gaunt factor is

$$g_I(n', n) = \pi\sqrt{3} \frac{nn'}{(n - n')} \left[ \frac{n - n'}{n + n'} \right]^{2n'+2n} \Delta(n', n) \quad (3.108)$$

where  $\Delta(n', n)$  involves some hypergeometric functions.

### 3.5.2 Bound-Free Transitions

Here we are concerned with transitions from a bound state to a free electronic state; this process clearly results in a continuum since the free states can have any energy. A logical starting point would be Fermi's Golden Rule (2.16) where the density of final states would involve the density of free electron states around some mean momentum where  $\frac{pe^2}{2m} \simeq E_{ph} - |E_n|$  is the energy of the ionizing photon minus the energy needed to ionize the atom from its initial state  $|E_n\rangle$ . We will take a less direct route, following Mihalas in using a method that was introduced by Menzel and Pekeris in 1935 (MNRAS, 96, 77). We know that bound states can be represented by quantum number  $n$ , such that  $E_n = -R/n^2$  for hydrogen, and transitions occur at energies,

$$h\nu_{n'n} = R \left( \frac{1}{n'^2} - \frac{1}{n^2} \right) \quad (3.109)$$

Now suppose we represent free states by imaginary quantum numbers  $ik$ , so that

$$h\nu_{n'k} = R \left( \frac{1}{n'^2} + \frac{1}{k^2} \right) = \frac{R}{n'^2} + \frac{1}{2}mv^2 \quad (3.110)$$

where  $v$  is the velocity of the free electron. Then we use the expression for the oscillator strength between bound states (3.106 and 3.108),

$$f_{n'n} = \frac{32}{3n'^2} \left| \frac{[(n - n')/(n + n')]^{2n+2n'} \Delta(n', n)}{n^2 n'^2 (n - n') \left( \frac{1}{n'^2} - \frac{1}{n^2} \right)^3} \right| \quad (3.111)$$

which can be rewritten to

$$f_{n'k} = \frac{32}{3n'^2} \frac{\exp[-4k \arctan(n'/k)] |\Delta(n', ik)|}{k^3 n'^3 \left( \frac{1}{n'^2} + \frac{1}{k^2} \right)^{\frac{7}{2}} (1 - e^{-2\pi k})} \quad (3.112)$$

or

$$f_{n'k} = \frac{32}{3n'\sqrt{3}} \frac{1}{n'^5 k^3} \left( \frac{1}{n'^2} + \frac{1}{k^2} \right)^{-3} g_{II}(n', k), \quad (3.113)$$

$$g_{II}(n', k) = \frac{\pi\sqrt{3}kn' \exp[-4k \arctan(n'/k)] |\Delta(n', ik)|}{(k^2 + n'^2)^{\frac{1}{2}} (1 - e^{-2\pi k})} \quad (3.114)$$

in analogy with 3.107 & 3.108. An extensive numerical tabulation of  $g_{II}$  is given by Karzas and Latter (Ap. J. Suppl. 6, 167).

By analogy with (2.82), we can define a frequency-dependent cross section  $f_n$  photoionization as

$$\sigma_{n'}(\nu) = \frac{\pi e^2}{mc} f_{n'k} \frac{dk}{d\nu} \quad (3.115)$$

where  $\frac{dk}{d\nu} = -\frac{hk^3}{2R}$  (from equation 3.110). Then

$$\sigma_{n'}(\nu) = \frac{\pi e^2}{mc} \left( \frac{hk^3}{2R} \right) \left( \frac{32}{3\pi\sqrt{3}} \right) \frac{1}{k^3 n'^5} \left( \frac{R}{h\nu} \right)^3 g_{II}(n', k) \quad (3.116)$$

Now the  $k^3$  cancels so we can think of  $k$  as specifying a frequency in  $g_{II}(n', k)$ . Then

$$\sigma_{n'}(\nu) = \frac{\pi e^2}{mc} \frac{R^2}{2h^2 \nu^3} \frac{32}{3\pi\sqrt{3}} \frac{1}{n'^5} g_{II}(n', \nu) \quad (3.117)$$

or since  $R = \frac{2\pi^2 me^4}{2h^2} (Z^2)$  we get  $\frac{\pi e^2}{mc} \frac{R^2}{h^2} = \frac{4\pi^5 me^{10}}{ch^6}$ , giving

$$\sigma_{n'}(\nu) = \frac{64\pi^4}{3\sqrt{3}} \frac{me^{10}}{ch^6} \frac{1}{n'^5 \nu^3} g_{II}(n', \nu) = 2.815 \times 10^{29} \frac{g_{II}(n', \nu)}{n'^5 \nu^3} \quad (3.118)$$

The cross section for bound-free absorption out of level  $n$  jumps from zero to a larger value when  $\nu = \frac{R}{hn^2}$  (threshold) and then declines as  $\nu^{-3}$ . Near threshold the Gaunt factor is near ( $\sim 20\%$ ) to unity. Bound-free transitions dominate the opacity of A and B stars. The opacity of hydrogen will exhibit a series of sharp edges, as deeper and deeper levels can be photoionized. The exact shape will depend on the distribution of hydrogen over its energy states. For temperatures  $\sim 10^4 K$ , typical of HII regions, the edges will start to show up in the near-infrared where they take over from the free-free which is falling off exponentially with increasing  $\nu$ . The edges become more prominent as  $\nu$  increases to the Lyman edge at 912 Å. The edges are also the series limits; e.g., the Balmer edge (or jump) is the Balmer series limit  $\sim 3647$  Å, and the Paschen edge is  $\sim 8206$  Å. At wavelengths longer than the near-infrared, the bound-free contribution becomes negligible compared to the free-free(thermal-bremsstrahlung).

Another contribution to the continuous spectrum of hydrogen is the 2-photon continuum produced by the decay of the  $2^2S_{\frac{1}{2}}$  to  $1^2S_{\frac{1}{2}}$ . This transition is highly forbidden (cannot decay by emission of a single photon), but it can occur via the  $e^2 A^2$  term in equation 2.23. Since the photon energies must add up to  $E_2 - E_1 = h\nu_{Ly\alpha}$ , we have a continuum peaked around  $\nu = \frac{1}{2}\nu_{Ly\alpha}$

The radiative recombination cross sections, which are important in determining the ionization state and spectrum, can be obtained from the photoionization or bound-free cross sections by detailed balancing in analogy with the procedure used in section 2.3 to relate Einstein  $B$ 's. The resulting Milne relation is

$$\frac{\sigma_{recomb}}{\sigma_{photoion}} = \frac{\sigma_{fb}}{\sigma_{bf}} = \left( \frac{h\nu}{mc v} \right)^2 \frac{2g_n}{g_e g_+}, \quad (3.119)$$

where  $v$  is the velocity of the recombining electron,  $g_e$  is its statistical weight,  $g_+$ ,  $g_n$  are the statistical weights of the ion and neutral atom.

### 3.5.3 Free-Free Transitions

By extending the approach used for bound-free transitions, we can deal with free-free transitions by using imaginary quantum numbers for *both* final and initial states. This may seem a strange approach, but one can look at it this way. Suppose a free electron passes near a proton; the two produce a temporary dipole moment which can either emit or absorb a photon. The emission process could also be understood classically in terms of emission by an accelerated charge as the free electrons path is bent by the field of the proton, but the quantum approach is more fundamental.

By analogy with equation 3.115, we define a cross section for absorption of light at frequency  $\nu$  by a transition from initial “state”  $ik$  to find “state”  $i\ell$ , per ion and electron moving at velocity  $v$ , where  $\frac{1}{2}mv^2 = R/k^2$ .

Then  $R/k^2 + h\nu = R/\ell^2$  expresses energy conservation and we rewrite 3.115 as

$$\sigma(\nu, v) = \frac{\pi e^2}{mc} f_{k\ell} dk \left( \frac{d\ell}{d\nu} \right) \quad (3.120)$$

where  $dk$  is the band of states around  $k$  and  $d\ell (= (d\ell/d\nu)\Delta\nu)$  is band of states around  $\ell$ . The appropriate generalization of the Kramers approximation for  $f_{k\ell}$  is (since  $g_k \neq 2k^2$ )

$$f_{k\ell} = \frac{64}{3\pi\sqrt{3}} \frac{1}{g_k} \frac{1}{k^3\ell^3} \left( \frac{1}{k^2} - \frac{1}{\ell^2} \right)^{-3} g_{III}(k, \ell) \quad (3.121)$$

Where  $g_{III}(k, \ell)$  is the free-free Gaunt factor. The degeneracy  $g_k$  is now the statistical weight of the free electron

$$g_k = \frac{2m^3 4\pi v^2 dv}{h^3} = \frac{16\pi R m^2 v}{h^3 k^3} dk \quad (3.122)$$

Then

$$\sigma(\nu, v) = \left( \frac{\pi e^2}{mc} \right) \frac{64}{3\pi\sqrt{3}} \left( \frac{h^3 k^3}{16\pi R m^2 v} \right) \left( \frac{R}{h\nu} \right)^3 \frac{g_{III}(\nu, v)}{k^3 \ell^3} \frac{d\ell}{d\nu} \quad (3.123)$$

For fixed  $k$ ,  $d\ell/d\nu = h\ell^3/2R$  so

$$\sigma(\nu, v) = \left( \frac{2Rhe^2}{3\pi\sqrt{3}m^3c} \right) \frac{g_{III}(\nu, v)}{\nu^3 v} \quad (3.124)$$

Now if we have a Maxwellian distribution of velocities, we want to average  $\sigma(\nu, v)$  over that distribution, to get

$$\sigma(\nu, T) = \frac{4e^6}{3ch} \left( \frac{2\pi}{3km^3} \right)^{\frac{1}{2}} T^{-\frac{1}{2}} \nu^{-3} \bar{g}_{III}(\nu, T) \quad (3.125)$$

where  $\bar{g}_{III}(\nu, T) \equiv \int_0^\infty g_{III}(\nu, v) e^{-4} du$ ,  $u = mv^2/2kT$ . Neglecting the Gaunt factor, the free-free absorption cross section goes as  $\nu^{-3}$  and  $T^{-\frac{1}{2}}$ , as can be shown with classical calculation (e.g. Rybicki and Lightman, Chapter 5).

The free-free opacity dominates the bound-free at low frequencies, where only edges from high  $n$  states can contribute (the photon does not have enough energy to ionize lower  $n$  states). Also as the temperature rises, more and more of the hydrogen is ionized so bound-free absorption decreases, and free-free opacity becomes the dominant opacity source in O stars.

Free-free transitions are also important emission sources at low  $\nu$  from ionized regions. Thus the radio to mid-infrared emission from gas in HII regions is dominated by free-free emission.

### 3.5.4 Fine Structure Transitions ( $\Delta n = 0$ )

So far, we have neglected the fact that the energy levels are split by fine and hyperfine effects. These splittings are quite small for hydrogen so this neglect is not totally unreasonable, when considering the principal ( $\Delta n \neq 0$ ) transitions. Let us now consider transitions between these split levels (i.e.,  $\Delta n = 0$ ). These are called fine structure and hyperfine structure transitions. From eq. 3.17 and 3.70, we have

$$E_{n,j} = -\frac{R}{n^2} - \frac{Ra^2}{n^3} \left( \frac{1}{(j + \frac{1}{2})} - \frac{3}{4n} \right) \quad (3.126)$$

Where  $R = e^2/2a_0$  = Rydberg. So for fixed  $n$ , levels with  $j, j-1$  are split by

$$\Delta E = E_{n,j} - E_{n,j-1} = \frac{Ra^2}{n^3} \left( \frac{1}{j - \frac{1}{2}} - \frac{1}{j + \frac{1}{2}} \right). \quad (3.127)$$

For example, for  $n = 2$ , which we considered previously,  $j = \frac{1}{2}$  or  $\frac{3}{2}$ . The transition  $j = \frac{3}{2} \rightarrow \frac{1}{2}$  will occur at frequency

$$\nu = \frac{\Delta E}{h} = \frac{1.7510 \times 10^{10}}{n^3} \left( \frac{1}{1} - \frac{1}{2} \right) = 10.9 \times 10^9 \text{Hz} \simeq 11 \text{GHz} \quad (3.128)$$

If we consider  ${}^2P_{\frac{3}{2}} \rightarrow {}^2S_{\frac{1}{2}}$  we have  $\ell \rightarrow \ell - 1, j \rightarrow j - 1, \Delta s = 0$  so this is allowed by electric dipole. More generally for any  $(n, \ell, j) \rightarrow (n, \ell - 1, j - 1)$  we have a strength,

$$S_{j,j-1} \equiv S(n, \ell, j; n, \ell - 1, j - 1) = \frac{9n^2}{16j} (4j^2 - 1)(n^2 - \ell^2) a_0^2 e^2 \quad (3.129)$$

and Einstein A,

$$A_{j,j-1} = \frac{64\pi^4 \nu^3}{3hc^3} \frac{S_{j,j-1}}{g_j} = 7.521 \times 10^{-38} \frac{\nu^3}{g_j} \left( \frac{S_{j,j-1}}{a_0^2 e^2} \right) \text{sec}^{-1} \quad (3.130)$$

For  ${}^2P_{\frac{3}{2}} \rightarrow {}^2S_{\frac{1}{2}}$ ,  $A = 8.9 \times 10^{-7} \text{ s}^{-1}$ .

Note that this  $A$  value is *much* less than those for  $\Delta n \neq 0$ , low  $n$ . Since both are electric dipole transitions, the difference is mostly in the  $\nu^3$  factor.

For the  ${}^2P_{\frac{3}{2}} \rightarrow {}^2P_{\frac{1}{2}}$  transition in  $n = 2$ , we have  $\Delta\ell = 0$ . This will require a transition with an operator of even parity; the lowest order of these is magnetic dipole. This interaction involves the perturbation of the  $\vec{B}$  field of the electromagnetic wave and the magnetic moment of the electron  $\vec{\mu} = -(\mu_B/\hbar)(\vec{L} + 2\vec{S})$  (eq. 3.74).

As in the case of the electric dipole moment, we can write  $\vec{\mu}$  in terms of its projection on three Cartesian axis or in terms of axes representing circularly or linearly polarized light. The latter representation leaves us with 3 spherical harmonics of first order:

$$\vec{\mu} = \sum_q \mu_q \hat{e}_q = \mu \sum_q \frac{\mu_q}{\mu} \hat{e}_q \quad (3.131)$$

where  $\mu$  is the *magnitude* of the dipole moment.

The magnetic dipole operator, like the electric dipole operator, is a tensor operator of rank 1. So we will get the same selection rules on  $j$  and  $m$  from the three- $j$ -symbols:  $\Delta j = 0, \pm 1, j + j' \geq 1$ ,

and  $\Delta m = q$ . For unpolarized light, we will sum over  $q$  (equivalent to sum over  $x, y, z$  after writing in terms of spherical harmonics). Then if we compute the transition strength,  $S_{ij} = g_i |\mu_{ij}|^2$  or  $S_{ij} = \sum_{m,m'} |\langle n\ell s j | \sum_q \mu_q |n' \ell' s' j' m' \rangle|^2$ , and use the Wigner-Eckart theorem, we have

$$S_{ij} = \sum_q |\langle n\ell s j | \mu |n' \ell' s' j' \rangle|^2 \sum_{m,m'} \begin{pmatrix} j & 1 & j' \\ -m & q & m' \end{pmatrix}^2 \quad (3.132)$$

After transposing the last two columns, according to

$$\begin{pmatrix} j & 1 & j' \\ -m & q & m' \end{pmatrix} = (-1)^{j+j'+1} \begin{pmatrix} j & j' & 1 \\ -m & m' & q \end{pmatrix}, \quad (3.133)$$

we use an orthogonality relation (Sobelman- eq.4.42):

$$\sum_{mm'} \begin{pmatrix} j & j' & 1 \\ -m & m' & q \end{pmatrix}^2 = \frac{1}{2(1)+1} \delta_{11} \delta_{qq} \Delta(jj'1) \quad (3.134)$$

yielding  $\frac{1}{3}$  if  $|j-j'| \leq 1 \leq |j+j'|$  and 0 otherwise. Since  $\mu$  is independent of  $q$  in the reduced matrix element

$$\sum_q = 3 \quad (3.135)$$

and  $S_{ij} = \langle n\ell s j | \mu |n' \ell' s' j' \rangle|^2$  where  $\mu$  is the *magnitude* of  $\mu$ . The argument is general and applies to electric dipole radiation also. For the present case we have,

$$\langle n\ell s j | \mu |n' \ell' s' j' \rangle = \frac{-\mu_B}{\hbar} \langle n\ell s j | L | n' \ell' s' j' \rangle - \left( \frac{2\mu_B}{\hbar} \right) \langle n\ell s j | S | n' \ell' s' j' \rangle \quad (3.136)$$

Since  $\mu$  does not involve  $r$  for the magnetic dipole moment,  $n$  cannot change ( $\langle n|n' \rangle = \delta_{nn'}$ ) and the parity rule for *magnetic* dipole ensures that  $\Delta\ell = 0$ . So only  $j$  and  $s$  can change. If we consider  $\Delta j = \pm 1$ ,  $\Delta\ell = 0$ , then it is clear that we are flipping a spin so making transitions between fine structure components such as  $j = \ell + \frac{1}{2}$  and  $j' = \ell - \frac{1}{2}$ , and  $\Delta s = \Delta j$ . Consider  $S_{ij} \equiv S(n\ell j; n\ell j-1)$  (e.g.  ${}^2P_{\frac{3}{2}} - {}^2P_{\frac{1}{2}}$  in the  $n=2$  state of hydrogen). We find

$$S(n\ell j; n\ell j-1) = \mu_B^2 \frac{(\ell + \frac{1}{2} + j + 1)(\ell + \frac{1}{2} - j + 1)(\frac{1}{2} + j - \ell)(j + \ell - \frac{1}{2})}{4j} \quad (3.137)$$

(Sobelman eq. 9.152, pg. 226, but note the error in the third factor) Note that no radial integral is involved, so that these transitions can be calculated very simply.

Note also that  $\mu_B = \frac{1}{2}\alpha e a_0$ ; comparison to 3.100 shows that magnetic dipole transitions are down by  $\sim \alpha^2$  relative to electric dipole transitions.

In particular, for  ${}^2P_{\frac{3}{2}} - {}^2P_{\frac{1}{2}}$  in hydrogen,  $\ell = 1$ ,  $j = \frac{3}{2}$  so

$$S(21\frac{3}{2}; 21\frac{1}{2}) = \frac{4}{3} \mu_B^2 = \frac{4}{3} \frac{1}{4} \alpha^2 e^2 a_0^2 = \frac{\alpha^2}{3} e^2 a_0^2 \quad (3.138)$$

Then, since  $g_j$  is the same as in eq. 3.124 and, neglecting the Lamb shift,  $\nu$  is the same, we get

$$A({}^2P_{\frac{3}{2}} - {}^2P_{\frac{1}{2}}) \simeq 7.521 \times 10^{-38} \frac{(11 \times 10^9)^3 \alpha^2}{4} \frac{1}{3} = 4.4 \times 10^{-13} \text{s}^{-1}. \quad (3.139)$$

### 3.5.5 Hyperfine Transition

Now let us consider the hyperfine transition of hydrogen. Now we want to use a basis set of states where  $J$  and  $I$  have been added to get  $F$ . So we want the strength,

$$S(njIF; njIF') = |\langle njIF | \mu | njIF' \rangle|^2 \quad (3.140)$$

where, by analogy with the case for fine structure transitions,  $\Delta j = 0$ ,  $\Delta F = \pm 1$ . For the particular case of the ground state of hydrogen  $^2S_{\frac{1}{2}}$ ,  $\ell = 0$ , so  $\vec{\mu} = \frac{2\mu_B}{\hbar} \vec{S}$ . Life would be easy if our basis states were eigenfunctions of  $\vec{S}$ , but we are two steps removed: we have added  $\vec{L} + \vec{S} = \vec{J}$  and  $\vec{J} + \vec{I} = \vec{F}$ . The resulting tangle of angular momentum addition leads to 6j symbols. The result can be expressed as (Sobelman, pg. 229)

$$S(njIF; njIF') = g^2 \mu_B^2 (2F+1)(2F'+1) j(j+1)(2j+1) \left\{ \begin{array}{ccc} j & F & I \\ F' & j & 1 \end{array} \right\}^2, \quad (3.141)$$

where  $g$  is the Landé  $g$  factor

$$g = \left[ 1 + \frac{j(j+1) + \frac{3}{4} - \ell(\ell+1)}{2j(j+1)} \right], \quad (3.142)$$

which arises from adding  $\vec{L} + \vec{S}$  to get  $\vec{J}$  and the quantity in curly brackets is the 6j symbol which arises in  $\vec{J} + \vec{I} = \vec{F}$ . As long as we remain restricted to  $^2S_{\frac{1}{2}}$  states, a considerably simpler expression can be derived for arbitrary  $I$ . For the case of  $F' = 0$ ,  $F = 1$ ,  $j = \frac{1}{2}$ ,  $I = \frac{1}{2}$  (the 21cm line),

$$S \left( 1 \frac{1}{2} \frac{1}{2} 1; 1 \frac{1}{2} \frac{1}{2} 0 \right) = 3\mu_B^2 \quad (3.143)$$

so

$$A_{21cm} = \frac{64\pi^4\nu^3}{3hc^3} |\mu_{10}|^2 = \frac{64\pi^4\nu^3}{3hc^3} \frac{S_{ij}}{g_j} = \frac{64\pi^4\nu^3}{3hc^3} \frac{3\mu_B^2}{2(1)+1} \quad (3.144)$$

$$= \frac{64\pi^4\nu^3}{3hc^3} \mu_B^2 = 2.85 \times 10^{-15} \text{s}^{-1} \quad (3.145)$$

Note how slow this transition is; only the huge number of H atoms and the fact that this is a probabilistic lifetime makes this transition detectable. The low frequency (1420 MHz) and the fact that it is a magnetic dipole transition accounts for the 23 orders of magnitude compared to the Lyman  $\alpha$ :

$$[(\nu(21cm)/\nu(\text{Lyman } \alpha))^3 \times (1/\alpha)^2] \sim 10^{23}. \quad (3.146)$$

In a similar way, you should be able to estimate roughly the order of magnitude of other transitions by comparing to Lyman  $\alpha$ .

## Chapter 4

# Atomic Structure with Two Electrons— Helium

### 4.1 The Many Particle Problem

Since helium is the first system we will deal with that has more than one identical particle, we will first spend some time discussing the general features of the many body problem and then discuss the peculiar quantum mechanical features introduced by the indistinguishability of elementary particles. In principle, the many body problem is a straightforward extension of our work on one particle systems. As in equation 1.2 we can write the Hamiltonian as the sum of the kinetic and potential energies, where the kinetic energy depends on all the momentum coordinates and the potential energy depends on all the spatial coordinates (and perhaps on the time):

$$H = T(\vec{p}_1, \vec{p}_2, \dots, \vec{p}_N) + V(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t) \quad (4.1)$$

and

$$T = \sum_{j=1}^N -\frac{\hbar^2 \vec{\nabla}_j^2}{2m_j} \quad (4.2)$$

The solutions are functions of all the coordinates, such that

$$H\Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t) = i\hbar \frac{\partial}{\partial t} \Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t) \quad (4.3)$$

and

$$\rho(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t) = \Psi^*(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t)\Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t) \quad (4.4)$$

is the probability that, at time  $t$ , particle 1 is in volume element  $d\vec{r}_1$  about  $\vec{r}_1$ , etc. In this case, we can normalize so that

$$\int \rho(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = 1 \quad (4.5)$$

The rest of the struture goes through as before.

## 4.2 Identical Particles and Exchange Symmetry

Since all electrons are indistinguishable, all measureable quantities must be unaffected by the exchange of electrons. This is called exchange symmetry. We introduce the particle exchange operator  $P_{ij}$  that exchanges all coordinates (including spins, but we will ignore that for the moment) of particles  $i$  and  $j$ . Consider, for example, the effect of  $P_{12}$ :

$$P_{12}\Psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N, t) = \Psi(\vec{r}_2, \vec{r}_1, \dots, \vec{r}_N, t) \quad (4.6)$$

The symmetry can be expressed by the fact that  $P_{ij}$  commutes with the Hamiltonian operator:

$$[H, P_{ij}] = 0 \quad (4.7)$$

Some remarkable consequences flow from exchange symmetry. Suppose that we have some eigenfunctions of the Hamiltonian,  $\psi_E$ , such that

$$H\psi_E = E\psi_E \quad (4.8)$$

If we apply  $P_{ij}$ ,

$$P_{ij}H\psi_E = P_{ij}E\psi_E \quad (4.9)$$

and since  $H$  and  $P_{ij}$  commute,

$$HP_{ij}\psi_E = EP_{ij}\psi_E \quad (4.10)$$

so  $P_{ij}\psi_E$  is also an eigenfunction of  $H$  with eigenvalue  $E$ . So the exchange operator can generate a set of degenerate eigenfunctions. This is called **exchange degeneracy**.

There are eigenfunctions that satisfy

$$P_{ij}\psi_E = c\psi_E \quad (4.11)$$

where  $c$  is a constant. Then

$$P_{ij}P_{ij}\psi_E = c^2\psi_E. \quad (4.12)$$

But two exchanges return us to the original, so the left hand side is also equal to  $\psi_E$ , and we must have  $c^2 = 1$  or  $c = \pm 1$ . These two possibilities are the most fundamental division of particles into bosons ( $c = +1$ ) or fermions ( $c = -1$ ). As we know, all particles with integer spin are bosons, including photons, and other “force carriers”, such as  $W^\pm$ ,  $Z$ , gluons, and gravitons. Particles that make up matter (electrons, quarks, protons, neutrons) are fermions with half-integral spin. This division is an observational fact, but it also can be shown from relativistic quantum mechanics. Composite matter particles, such as pions (2 quarks) or  ${}^4\text{He}$  nuclei (4 fermions) may be bosons.

### 4.3 Two Identical Particles

Now let's focus on two identical fermions (obviously relevant to helium) but consider only one dimension for simplicity. Then

$$H = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x_1^2} + V(x_1) - \frac{\hbar^2}{2m} \frac{\partial^2}{\partial x_2^2} + V(x_2) + V'(x_1 - x_2) \quad (4.13)$$

where  $V(x_1)$  is the potential (e.g., the nuclear coulomb interaction) experienced by particle 1,  $V(x_2)$  is the potential experienced by particle 2, and  $V'(x_1 - x_2)$  is the interaction potential between the two particles. Assume that  $V'(x_1 - x_2) > 0$ , is a maximum for  $x_1 = x_2$ , and decreases as  $|V'(x_1 - x_2)|$  increases. Also assume that it is symmetric, in the sense that  $V'(x_1 - x_2) = V'(x_2 - x_1)$ . The presence of  $V'(x_1 - x_2)$  makes exact solution of the eigenfunction problem impossible, but if  $V'(x_1 - x_2)$  is small enough, we can treat it as a perturbation. Then suppose that we have solutions to the unperturbed problem:

$$H^{(0)} \psi_{n_1 n_2}^{(0)}(x_1, x_2) = E_{n_1 n_2}^{(0)} \psi_{n_1 n_2}^{(0)}(x_1, x_2) \quad (4.14)$$

where  $H^{(0)}$  is the Hamiltonian without the perturbation. It is plausible that  $\psi_{n_1 n_2}^{(0)}(x_1, x_2)$  is a direct product of the eigenfunctions for the corresponding one particle problems:

$$\psi_{n_1 n_2}^{(0)}(x_1, x_2) = \psi_{n_1}^{(0)}(x_1) \psi_{n_2}^{(0)}(x_2) \quad (4.15)$$

and  $E_{n_1 n_2}^{(0)} = E_{n_1}^{(0)} + E_{n_2}^{(0)}$ . But this form of the function does not, in general, satisfy exchange symmetry:

$$P_{12} \psi_{n_1 n_2}^{(0)}(x_1, x_2) \neq \pm 1 \psi_{n_1 n_2}^{(0)}(x_1, x_2) \quad (4.16)$$

as can be seen by taking  $n_1 \neq n_2$ .

We can however combine the two eigenfunctions to make an explicitly symmetric or antisymmetric eigenfunction. Consider the operator  $(\mathbf{1} \pm P_{12})$ . Applying this to  $\psi_{n_1 n_2}^{(0)}(x_1, x_2)$  produces

$${}^S \psi_{n_1 n_2}^{(0)} \equiv N_S \left[ \psi_{n_1}^{(0)}(x_1) \psi_{n_2}^{(0)}(x_2) + \psi_{n_2}^{(0)}(x_1) \psi_{n_1}^{(0)}(x_2) \right] \quad (4.17)$$

or

$${}^A \psi_{n_1 n_2}^{(0)} \equiv N_A \left[ \psi_{n_1}^{(0)}(x_1) \psi_{n_2}^{(0)}(x_2) - \psi_{n_2}^{(0)}(x_1) \psi_{n_1}^{(0)}(x_2) \right] \quad (4.18)$$

where  $N_S$  and  $N_A$  take care of normalization. The energy eigenvalues are unchanged. Obviously, we must use the form in equation 4.18 for two fermions. Note that if  $n_1 = n_2$ ,  ${}^A \psi_{n_1 n_2}^{(0)} = 0$ . Since  $n_1$  is shorthand for **all** the quantum numbers, including spin, the result is a statement of the Pauli exclusion principle: no two fermions can occupy the same state.

Now, let's deal explicitly with spin, so let  $n_1, n_2$  represent all quantum numbers **except** spin. Then our zero<sup>th</sup> order wave functions would have the form,

$$\psi_{n_1}^{(0)}(x_1) \psi_{n_2}^{(0)}(x_2) \alpha(1) \beta(2) \quad (4.19)$$

where  $\alpha$  and  $\beta$  are the spin functions introduced earlier (equation 3.47). Then we apply the operator,  $\frac{1}{\sqrt{2}}(\mathbf{1} - P_{12})$  to get four antisymmetric wave functions:

$$\begin{aligned}\mathcal{A}\psi_{n_1 n_2; ++}^{(0)} &\equiv \frac{1}{\sqrt{2}} [\psi_{n_1}^{(0)}(x_1)\alpha(1)\psi_{n_2}^{(0)}(x_2)\alpha(2) - \psi_{n_2}^{(0)}(x_1)\alpha(1)\psi_{n_1}^{(0)}(x_2)\alpha(2)] \\ \mathcal{A}\psi_{n_1 n_2; +-}^{(0)} &\equiv \frac{1}{\sqrt{2}} [\psi_{n_1}^{(0)}(x_1)\alpha(1)\psi_{n_2}^{(0)}(x_2)\beta(2) - \psi_{n_2}^{(0)}(x_1)\beta(1)\psi_{n_1}^{(0)}(x_2)\alpha(2)] \\ \mathcal{A}\psi_{n_1 n_2; -+}^{(0)} &\equiv \frac{1}{\sqrt{2}} [\psi_{n_1}^{(0)}(x_1)\beta(1)\psi_{n_2}^{(0)}(x_2)\alpha(2) - \psi_{n_2}^{(0)}(x_1)\alpha(1)\psi_{n_1}^{(0)}(x_2)\beta(2)] \\ \mathcal{A}\psi_{n_1 n_2; --}^{(0)} &\equiv \frac{1}{\sqrt{2}} [\psi_{n_1}^{(0)}(x_1)\beta(1)\psi_{n_2}^{(0)}(x_2)\beta(2) - \psi_{n_2}^{(0)}(x_1)\beta(1)\psi_{n_1}^{(0)}(x_2)\beta(2)]\end{aligned}\quad (4.20)$$

As long as  $H^{(0)}$  contains no spin-dependent terms, these four functions are degenerate. If  $n_1 = n_2$ , the first and last wavefunctions ( $++$  and  $--$ ) are zero (the Pauli exclusion principle again). Also, the  $+-$  and  $-+$  forms are negatives of each other if  $n_1 = n_2$ .

We could now include the effects of the perturbation, but just as for spin-orbit coupling, it is useful to change our basis set to one in which we couple two spins to form a total spin. (In our one-dimensional model, we have no orbital angular momentum.) Then we have

$$\vec{S} = \vec{S}_1 + \vec{S}_2 \quad (4.21)$$

where

$$\begin{aligned}S^2 \chi_{SM_S} &= S(S+1)\hbar^2 \chi_{SM_S} \\ S_z \chi_{SM_S} &= M_S \hbar \chi_{SM_S}\end{aligned}\quad (4.22)$$

where  $\chi_{SM_S}$  are combinations of  $\alpha$  and  $\beta$  that satisfy equation 4.22. Then we construct new zero<sup>th</sup> order eigenfunctions from

$$\psi_{n_1 n_2; SM_S}^{(0)} = \sum_{m_{s_1} m_{s_2}} a_{m_{s_1} m_{s_2}} \psi_{n_1 n_2; m_{s_1} m_{s_2}}^{(0)} \quad (4.23)$$

where the coefficients are Clebsch-Gordon coefficients,  $\langle s_1 s_2 m_{s_1} m_{s_2} | s_1 s_2 SM_S \rangle$ , which is zero unless

$$\begin{aligned}M_S &= m_{s_1} + m_{s_2} \\ |s_1 - s_2| &\leq S \leq s_1 + s_2\end{aligned}\quad (4.24)$$

We can write the new eigenfunctions in this form:

$$\psi_{n_1 n_2; 00}^{(0)} = \frac{1}{\sqrt{2}} [\psi_{n_1}^{(0)}(x_1)\psi_{n_2}^{(0)}(x_2) + \psi_{n_2}^{(0)}(x_1)\psi_{n_1}^{(0)}(x_2)] \chi_{00} \quad (4.25)$$

and

$$\psi_{n_1 n_2; 1M_S}^{(0)} = \frac{1}{\sqrt{2}} [\psi_{n_1}^{(0)}(x_1)\psi_{n_2}^{(0)}(x_2) - \psi_{n_2}^{(0)}(x_1)\psi_{n_1}^{(0)}(x_2)] \chi_{1M_S} \quad (4.26)$$

where  $\chi_{SM_S}$  are the new spin functions. Since the spatial part of equation 4.25 is symmetric, the spin part must be antisymmetric:

$$\chi_{00} \equiv \frac{1}{\sqrt{2}} [\alpha(1)\beta(2) - \beta(1)\alpha(2)]. \quad (4.27)$$

Because equation 4.26 is antisymmetric in the spatial part, it can be paired with three symmetric spin functions,  $\chi_{1M_S}$ :

$$\chi_{11} \equiv \alpha(1)\alpha(2) \quad (4.28)$$

$$\chi_{10} \equiv \frac{1}{\sqrt{2}} [\alpha(1)\beta(2) + \beta(1)\alpha(2)] \quad (4.29)$$

$$\chi_{1-1} \equiv \beta(1)\beta(2) \quad (4.30)$$

Thus we have a singlet ( $S = 0$ ) in equation 4.27 and a triplet ( $S = 1$ ) in equations 4.28-4.30.

Now consider the effect of the perturbation,  $V'(x_1 - x_2)$ . In particular, it is reasonable to assume that the unperturbed ground state of the two fermion system is that corresponding to the ground state of the single particle problem ( $n_1 = n_2 = 1$ ). In this case, equation 4.26 is zero; only a singlet state is possible. The ground state energy is then (cf. equation 1.62):

$$E_{11} = E_1^{(0)} + E_2^{(0)} + \langle \psi_{11;00}^{(0)} | V'(x_1 - x_2) | \psi_{11;00}^{(0)} \rangle \quad (4.31)$$

The matrix element is referred to as the direct integral and denoted  $J_{11}$ . Because  $V'$  does not affect spins, only the spatial part of the wave function is involved. If  $n_1 = n_2$  is plugged into equation 4.25, it is clear that we must divide it by a further  $\sqrt{2}$  to keep it normalized.

## 4.4 The Helium Atom

Now we are ready to examine the helium atom using the tools developed for the 1-D problem of two identical fermions. The issues of the wave functions will be the same, but we must generalize the Hamiltonian to apply to three dimensions and the interaction potential must be specified. Let the unperturbed Hamiltonian be

$$H^{(0)} = \sum_{i=1}^2 \left( \frac{-\hbar^2}{2m_e} \nabla_i^2 - \frac{Ze^2}{r_i} \right) \quad (4.32)$$

and the perturbation is the interaction of the two electrons:

$$H^{(1)} = \frac{e^2}{r_{12}} \quad (4.33)$$

where  $r_i$  is the distance of the  $i^{th}$  electron from the nucleus of charge  $Ze$  and  $r_{12}$  is the distance between the two electrons. The presence of  $r_{12}$  in  $H^{(1)}$  means that this is *not* a central force problem, once we turn on the perturbation.

Let us first consider the ground state of the system described by the unperturbed Hamiltonian. The solutions are those of hydrogenic atoms, so we assume that the ground state will have  $n_1 = n_2 = 1$ ;  $l_1 = l_2 = 0$ . This means that the electron spins must be antiparallel. Then we can construct the unperturbed ground state wave functions:

$$\psi_{1s,1s;00}^{(0)} = \psi_{1s}^{(0)}(\vec{r}_1) \psi_{1s}^{(0)}(\vec{r}_2) \chi_{00} \quad (4.34)$$

where

$$\chi_{00} = \frac{1}{\sqrt{2}} [\alpha(1)\beta(2) - \beta(1)\alpha(2)] \quad (4.35)$$

and  $\psi_{1s}^{(0)}(\vec{r}_1)$  is the single particle wave function with  $n_1 = 1$ ,  $l = 0$ , and  $m_l = 0$ , which we abbreviate by  $1s$ . Note that this is the special case where  $n_1 = n_2$  so only a singlet state exists.

The perturbed energy level will be

$$E_{1s,1s} = 2E_1^{(0)} + \langle \psi_{1s,1s;00}^{(0)} | \frac{e^2}{r_{12}} | \psi_{1s,1s;00}^{(0)} \rangle \quad (4.36)$$

Note that  $2E_1^{(0)} = -Z^2 e^2 / a_0 = -Z^2$  Hartrees. To zeroth order, the ground state will lie deeper than that of H by  $2Z^2$  which is  $-108.8$  eV for He. However, the perturbation is positive and will raise this level. Since we are in the ground state, we only need evaluate the matrix element in equation 4.36, which we will again call the direct integral and label  $J_{11}$ :

$$J_{11} = \int d\vec{r}_1 \int d\vec{r}_2 |\psi_{1s}^{(0)}(\vec{r}_1)|^2 \frac{e^2}{r_{12}} |\psi_{1s}^{(0)}(\vec{r}_2)|^2 \quad (4.37)$$

We have to express  $r_{12}$  in terms of  $r_1$  and  $r_2$ . We can write

$$r_{12}^2 = r_1^2 + r_2^2 - 2r_1 r_2 \cos \theta \quad (4.38)$$

Then it is possible to expand  $1/r_{12}$  in terms of Legendre polynomials (see, e.g., Jackson, Classical Electrodynamics, pg 62):

$$\frac{1}{r_{12}} = \sum_{\lambda=0}^{\infty} \frac{r_<^\lambda}{r_>^{\lambda+1}} P_\lambda(\cos \theta) \quad (4.39)$$

where  $r_<$  ( $r_>$ ) is the lesser (greater) of  $r_1$  and  $r_2$  and  $P_\lambda(\cos \theta)$  is a Legendre polynomial. Using the addition theorem for spherical harmonics (e.g., Jackson, pg. 68-69),

$$P_\lambda(\cos \theta) = \frac{4\pi}{2\lambda+1} \sum_{\mu=-\lambda}^{+\lambda} Y_{\lambda\mu}^*(\theta_1, \phi_1) Y_{\lambda\mu}(\theta_2, \phi_2) \quad (4.40)$$

we get

$$\frac{1}{r_{12}} = 4\pi \sum_{\lambda=0}^{\infty} \frac{1}{2\lambda+1} \frac{r_<^\lambda}{r_>^{\lambda+1}} \sum_{\mu=-\lambda}^{+\lambda} Y_{\lambda\mu}^*(\theta_1, \phi_1) Y_{\lambda\mu}(\theta_2, \phi_2) \quad (4.41)$$

While this hardly looks like an improvement, it now involves only variables associated with  $\vec{r}_1$  or  $\vec{r}_2$ , so it can be integrated, recalling that  $d\vec{r} = r^2 \sin \theta d\theta d\phi$ . The angular integrations can all be done in terms of 3-J symbols, leaving a messy, but doable radial integral. The net result is surprisingly simple:

$$J_{11} = \frac{5}{8} \frac{Ze^2}{a_0}; \quad \text{and thus} \quad E_{1s,1s} = -\frac{Z^2 e^2}{a_0} \left( 1 - \frac{5}{8Z} \right) \quad (4.42)$$

For Helium,  $Z = 2$  and  $E_{1s,1s} = -74.42$  eV. The substantial shift in ground state energy arises from the electron-electron repulsion. Another way to think of it is screening: the presence of one electron tends to “screen” the other electron from the full Coulomb potential of the nuclear charge. The perturbation is better dealt with in a variational method with the result:

$$E_{1s,1s} = -\frac{Z^2 e^2}{a_0} \left( 1 - \frac{5}{8Z} + \frac{25}{256Z^2} \right) = -77.51 \text{eV} \quad (4.43)$$

considerably closer to the experimental value of  $-78.62$  eV.

Now let's consider excited states. We will continue to think of the electrons as inhabiting hydrogenic atoms and then coupling them via their Coulombic interaction. Could we have two electrons excited out of the ground ( $1s$ ) state? Because of the nature of the hydrogenic energy levels, for which the first excited state lies more than halfway (about three-quarters, in fact) to the unbound continuum states, the energy of helium with two electrons in excited states exceeds the binding energy:

$$E_{2s,2s} \approx E_{1s,1s} + 2|E_{2s} - E_{1s}| = E_{1s,1s} + 2|Z^2 \frac{3}{4} 13.6\text{eV}| = E_{1s,1s} + 6 \times 13.6\text{eV} > 0 \quad (4.44)$$

Only one electron can be excited at a time if we want to retain both bound to the nucleus. Because one of the electrons must be in the  $1s$  ( $n = 1, l = 0$ ) hydrogenic state, we can write the wave functions of excited states as

$$\psi_{1s,n'l'm'_l;00}^{(0)} = \frac{1}{\sqrt{2}} \left[ \psi_{1s}^{(0)}(\vec{r}_1) \psi_{n'l'm'_l}^{(0)}(\vec{r}_2) + \psi_{n'l'm'_l}^{(0)}(\vec{r}_1) \psi_{1s}^{(0)}(\vec{r}_2) \right] \chi_{00} \quad (4.45)$$

$$\psi_{1s,n'l'm'_l;1M_S}^{(0)} = \frac{1}{\sqrt{2}} \left[ \psi_{1s}^{(0)}(\vec{r}_1) \psi_{n'l'm'_l}^{(0)}(\vec{r}_2) - \psi_{n'l'm'_l}^{(0)}(\vec{r}_1) \psi_{1s}^{(0)}(\vec{r}_2) \right] \chi_{1M_S} \quad (4.46)$$

where  $\chi_{00}$  and  $\chi_{1M_S}$  are the singlet and triplet spin functions from equations 4.27 to 4.30.

The occupation of the hydrogenic states is described as the *configuration*. For example,

Configuration	Zero <sup>th</sup> Order Energy
$1s1s$	$2E^{(0)}$
$1s2s, 1s2p$	$E_1^{(0)} + E_2^{(0)}$
$1s3s, 1s3p, 1s3d$	$E_1^{(0)} + E_3^{(0)}$

(4.47)

So, the zero<sup>th</sup> order energy of the  $1s2s$  configuration would be about  $-78 + 10 \sim -68$  eV.

Now, in excited states, another term enters the perturbation, the exchange integral, denoted  $K_{ij}$ . This arises because of the two terms in the wave functions. So the energy including electron-electron repulsion is to first order,

$$E_{1s;n'l'} \approx E_1^{(0)} + E_{n'}^{(0)} + J_{1s,n'l'} \pm K_{1s,n'l'} \quad (4.48)$$

where we take the plus sign for the singlet and the minus sign for the triplet. The direct and exchange integrals are

$$J_{nl,n'l'} \equiv \int d\vec{r}_1 \int d\vec{r}_2 |\psi_{nlm_l}^{(0)}(\vec{r}_1)|^2 \frac{e^2}{r_{12}} |\psi_{n'l'm'_l}^{(0)}(\vec{r}_2)|^2 \quad (4.49)$$

and

$$K_{nl,n'l'} \equiv \int d\vec{r}_1 \int d\vec{r}_2 \psi_{nlm_l}^{(0)*}(\vec{r}_1) \psi_{n'l'm'_l}^{(0)}(\vec{r}_1) \frac{e^2}{r_{12}} \psi_{nlm_l}^{(0)}(\vec{r}_2) \psi_{n'l'm'_l}^{(0)*}(\vec{r}_2) \quad (4.50)$$

The results of evaluating these integrals for the  $1s$ ,  $2s$  and  $1s$ ,  $2p$  configurations are

$$\begin{aligned} J_{1s,2s} &= 11.4 \text{ eV} & J_{1s,2p} &= 13.2 \text{ eV} \\ K_{1s,2s} &= 0.4 \text{ eV} & K_{1s,2p} &= 0.1 \text{ eV} \end{aligned} \quad (4.51)$$

with resulting first order energies of

$$\begin{aligned} E_{1s,2s} &= -68.0 + 11.4 \pm 0.4 = -56.2 \text{ eV for singlet; } -57.0 \text{ eV for triplet} \\ E_{1s,2p} &= -68.0 + 13.2 \pm 0.1 = -54.7 \text{ eV for singlet; } -54.9 \text{ eV for triplet} \end{aligned} \quad (4.52)$$

Note that triplet states are lower than singlet states because the electron-electron repulsion is less because they have to strictly avoid each other in the triplet state.

We need to reconsider the zero point of our energy scale. If we are talking about neutral helium, we should measure energies below that needed to remove *one* electron. That energy is 54.4 eV, so we set our zero point at  $-108.8 + 54.4 = -54.4$  eV. The first order ground state is then at  $54.4 - 78.6 = -24.2$  eV. A value of  $-24.58$  is more accurate.

The resulting energy level diagram is shown in the next figure. Note the separation into parahelium (singlet) and orthohelium (triplet). Very strong selection rules forbid transitions between them.

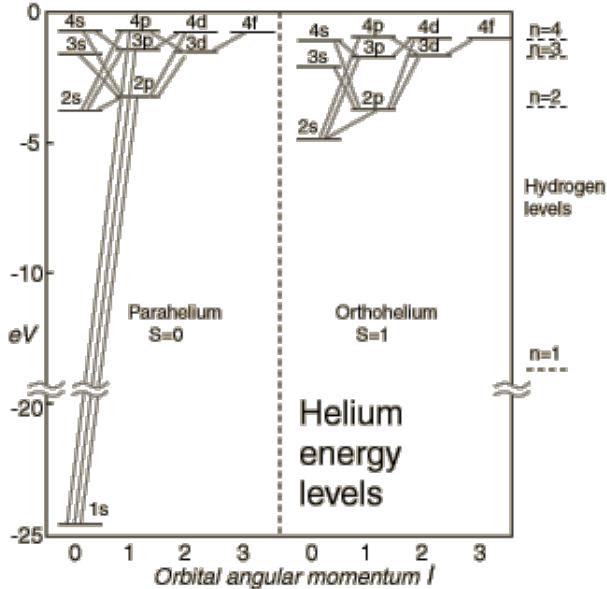


Figure 4.1: The He atom energy level diagram with a few transitions shown.

## Chapter 5

# Atomic Structure with Many Electrons

### 5.1 N-Electron Systems

Now let's return to the general many particle problem as formulated in §4.1 and §4.2. We must now generalize the methods used for the two-electron system in §4.3 and §4.4. Again we break the Hamiltonian into a zero<sup>th</sup> order Hamiltonian and a first order correction that accounts for the mutual interactions of the electrons:

$$H = H^{(0)} + H^{(1)} \quad (5.1)$$

with

$$H^{(0)} = \sum_{i=1}^N \left( \frac{-\hbar^2}{2m_e} \nabla_i^2 - \frac{Ze^2}{r_i} \right) \quad (5.2)$$

and

$$H^{(1)} = \sum_{i=1}^N \sum_{j>i}^N \frac{e^2}{r_{ij}} \quad (5.3)$$

Again, we require that  $H$  commute with all the particle exchange operators,  $P_{ij}$  since the electrons are indistinguishable.

$$[H, P_{ij}] = 0, \quad i, j = 1, 2, \dots, N; \quad i \neq j \quad (5.4)$$

Recall that the  $P_{ij}$  operators exchange **both** spatial and spin coordinates. Since electrons are fermions,

$$P_{ij}\psi_E = -\psi_E, \quad i, j = 1, 2, \dots, N; \quad i \neq j \quad (5.5)$$

where the  $\psi_E$  are eigenfunctions of  $H$ .

As our simplest approximation, we will neglect  $H^{(1)}$  so that all the electrons move independently of each other. This is called the independent particle model or the orbital approximation. Then the eigenfunctions will be combinations of direct products of the single particle (hydrogenic) eigenfunctions.

If we neglect for a moment the requirement in equation 5.5, we can construct an eigenfunction,  $\psi_\gamma$ , where  $\gamma$  represents all the relevant quantum numbers as follows:

$$\psi_\gamma^{(0)}(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = \psi_{n_1}^{(0)}(\vec{r}_1)\chi_{\frac{1}{2},m_{s_1}}(1)\psi_{n_2}^{(0)}(\vec{r}_2)\chi_{\frac{1}{2},m_{s_2}}(2) \cdots \psi_{n_N}^{(0)}(\vec{r}_N)\chi_{\frac{1}{2},m_{s_N}}(N) \quad (5.6)$$

with energy

$$E_{n_1}^{(0)} + E_{n_2}^{(0)} + \cdots + E_{n_N}^{(0)} \quad (5.7)$$

Now we must combine functions of the form of equation 5.6 to make a properly antisymmetric function. We can do this by forming a linear combination of all the degenerate functions that result from repeated pairwise particle interchanges, analogous to equation 4.20 for helium. Then

$$\psi_\gamma^{(0)} \equiv \frac{1}{\sqrt{N!}} \sum_p (-1)^p P \psi_\gamma^{(0)} \quad (5.8)$$

where the  $\psi_\gamma^{(0)}$  on the right-hand side of the equation are those of equation 5.6, for which we used the same notation. We will henceforth signify the left-hand side of equation 5.8 when we use  $\psi_\gamma^{(0)}$ . The sum in equation 5.8 is taken over all  $N!$  possible permutations of  $N$  electrons. Even permutations are equivalent to an even number of two particle interchanges, so  $(-1)^p = +1$  for even permutations and vice versa for odd permutations.  $P$  is a generalized permutation operator that can interchange electrons in all possible permutations. Thus, equation 5.8 yields an equation with all possible labelings of  $N$  electrons.

Mathematically, all these permutations are equivalent to a determinant, so the antisymmetrized wave function can be written as a **Slater determinant**:

$$\psi_\gamma^{(0)} = \frac{1}{\sqrt{N!}} \det \begin{vmatrix} \psi_{n_1}^{(0)}(\vec{r}_1)\chi_{\frac{1}{2},m_{s_1}}(1) & \psi_{n_1}^{(0)}(\vec{r}_2)\chi_{\frac{1}{2},m_{s_1}}(2) & \cdots & \psi_{n_1}^{(0)}(\vec{r}_N)\chi_{\frac{1}{2},m_{s_1}}(N) \\ \psi_{n_2}^{(0)}(\vec{r}_1)\chi_{\frac{1}{2},m_{s_2}}(1) & \psi_{n_2}^{(0)}(\vec{r}_2)\chi_{\frac{1}{2},m_{s_2}}(2) & \cdots & \psi_{n_2}^{(0)}(\vec{r}_N)\chi_{\frac{1}{2},m_{s_2}}(N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{n_N}^{(0)}(\vec{r}_1)\chi_{\frac{1}{2},m_{s_N}}(1) & \psi_{n_N}^{(0)}(\vec{r}_2)\chi_{\frac{1}{2},m_{s_N}}(2) & \cdots & \psi_{n_N}^{(0)}(\vec{r}_N)\chi_{\frac{1}{2},m_{s_N}}(N) \end{vmatrix} \quad (5.9)$$

In this form, it is clear that each electron “gets to spend some time” with each set of quantum numbers. A set of quantum numbers is associated with a single particle function and referred to as an orbital. The Pauli exclusion principle is built into equation 5.9. Suppose that two electrons have the same orbital ( $n_1 = n_2$  and  $m_{s_1} = m_{s_2}$ ); then at least two rows of the Slater determinant are equal, forcing the determinant to be zero. Electrons in identical spatial states must have their spins “paired” or anti-parallel.

Because the number of terms grows to inconvenient proportions (imagine the wave function for uranium), it is common to use a shorthand notation:

$$\psi_\gamma^{(0)}(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_N) = \left| \psi_{n_1}^{(0)}(\vec{r}_1)\chi_{\frac{1}{2},m_{s_1}}(1)\psi_{n_2}^{(0)}(\vec{r}_2)\chi_{\frac{1}{2},m_{s_2}}(2) \cdots \psi_{n_N}^{(0)}(\vec{r}_N)\chi_{\frac{1}{2},m_{s_N}}(N) \right| \quad (5.10)$$

where the vertical lines represent a Slater determinant with only the diagonal terms written out.

We can improve our Hamiltonian even before dealing with  $H^{(1)}$ . As was the case for helium, inner electrons will screen the outer electrons from the nuclear charge. We can replace the second term in equation 5.2 with a potential that corrects for screening. For example,

$$V_i(r_i) = -\frac{Ze^2}{r_i} + V_i^s(r_i) \quad (5.11)$$

As long as  $V_i^s(r_i)$  depend only on the radial coordinate ( $r_i$ ), called the **central field approximation**, we have the same equations for the angular variables that we had for hydrogen, so the orbitals' angular functions will be spherical harmonics. Of course, the radial equation has an extra term. The simplest approximation would be to assume that outermost electrons are fully screened by all the inner electrons. This might work reasonably well for elements with only one electron in the outermost orbital (alkali elements). A generalization is to imagine each electron moving in the smoothed out potential due to the “average” position of all the other electrons. If this approximation is iterated until a self consistent solution is obtained, it is called the **self-consistent field** or **SCF** approximation.

Once a self-consistent central field has been established, **non-central** forces, such as Coulomb and Exchange (electrostatic) interactions and spin-orbit, spin-spin (magnetic) interactions can be dealt with as perturbations. Before we include these effects, we will explore the use of the orbitals to understand atomic structure. We characterize the state of the atom by specifying the states of all the electrons, that is by their occupation of orbitals. Because of the Pauli principle, no more than two electrons can occupy a given spatial orbital. Thus we can “build up” the atomic state by filling the orbitals sequentially. Because the parity of an orbital is  $(-1)^l$  where  $l$  is the orbital angular momentum quantum number (equation 2.62), the parity of an  $N$ -electron atom is

$$(-1)^{l_1}(-1)^{l_2}\cdots(-1)^{l_N} = (-1)^{\sum_i^N l_i} \quad (5.12)$$

Thus, the parity is even if  $\sum_i^N l_i$  is even and odd if the sum is odd.

## 5.2 The Shell Model of the Atom

Table 5.1: Atomic Shells

$n$	Shell	$l$	Subshell	Orbitals
1	K	0	...	1s
2	L	0,1	$L_1, L_2$	2s, 2p
3	M	0,1,2	$M_1, M_2, M_3$	3s, 3p, 3d
4	N	0,1,2,3	$N_1, \dots N_4$	4s, 4p, 4d, 4f
5	O	0,1,2,3,4	$O_1, \dots O_5$	5s, 5p, 5d, 5f, 5g
6	P	0,1,2,3,4,5	$P_1, \dots P_6$	6s, 6p, 6d, 6f, 6g
7	Q	0,1,2,3,4,5,6	$Q_1, \dots Q_7$	7s, 7p, 7d, 7f, 7g

The shell model of the atom, with its historical origin and names, is still useful for understanding atomic structure. The shells are defined by the value of  $n$ , the principle quantum number. Subshells are defined by the value of orbital angular momentum of the electron ( $l$ ).

The building up, or *aufbau* principle allows us to understand the basic facts of atomic structure. We place electrons in hydrogenic states, accounting for the Pauli principle. We associate each principle quantum number with a **shell**, based on the fact that states with larger  $n$  lie on average farther from the nucleus, are more screened, and thus higher (less negative) in energy. Recall equation 3.23; the expectation value of the radius increases proportional to  $n^2$ , but there is a second term that involves  $l$ . This term will lead to some interesting effects when we get to discussing the order in which subshells (same  $n$ , different  $l$ ) will fill.

To first order, the shells fill in order of principle quantum number, but recall that low values of  $l$  correspond to highly elliptical orbits in the Bohr model, in which electrons spend more time closer to the nucleus (the origin of the second term in equation 3.23) and thus inside the screening electrons. For  $n = 4$ , this effect becomes large enough that the 4s orbital fills before the 3d orbital. Likewise the 5s orbital fills before the 4d, etc. These changes in order account for the peculiarities of the periodic table such as the transition metals and rare earths.

Two rules that work well are called **Madelung's Rules** (also known as **Klechkowski's Rules**):

1. Orbitals are filled in the order of increasing  $n + l$ ;
2. Where two orbitals have the same value of  $n + l$ , they are filled in order of increasing  $n$ .

These rules are illustrated in figure 5.1.

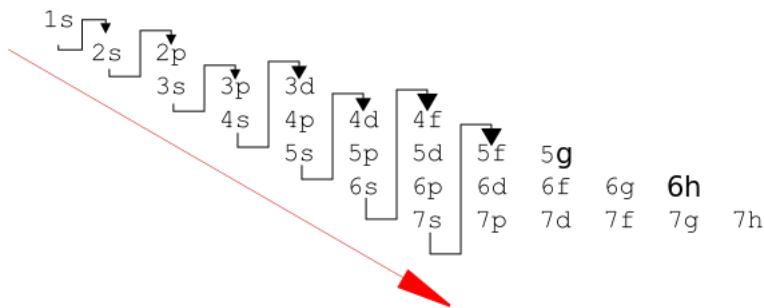


Figure 5.1: Madelung's Rule, also called Klechkowski's Rule. This gives the order of filling of atomic orbitals.

Note that Madelung's rules have exceptions, especially when ions are involved, but they are generally good guides. The net result of using these rules (with known exceptions) and arranging in order of number of electrons in outer shells (valence electrons) is the periodic table of the elements (Fig. 5.2). The detailed filling of the shells and orbitals is shown for the first two periods of the table in table 5.2.

# The Periodic Table

Figure 5.2: The Periodic Table of the Elements

While we use the hydrogenic orbitals as a basis, they are modified by the higher charge of the nucleus, moving in closer to the nucleus as  $Z$  increases. Neglecting screening,  $\langle r \rangle \propto 1/Z$ . The radial probability,  $r^2 R(r)$ , for a few inert gases illustrates this point (Fig. 5.3).

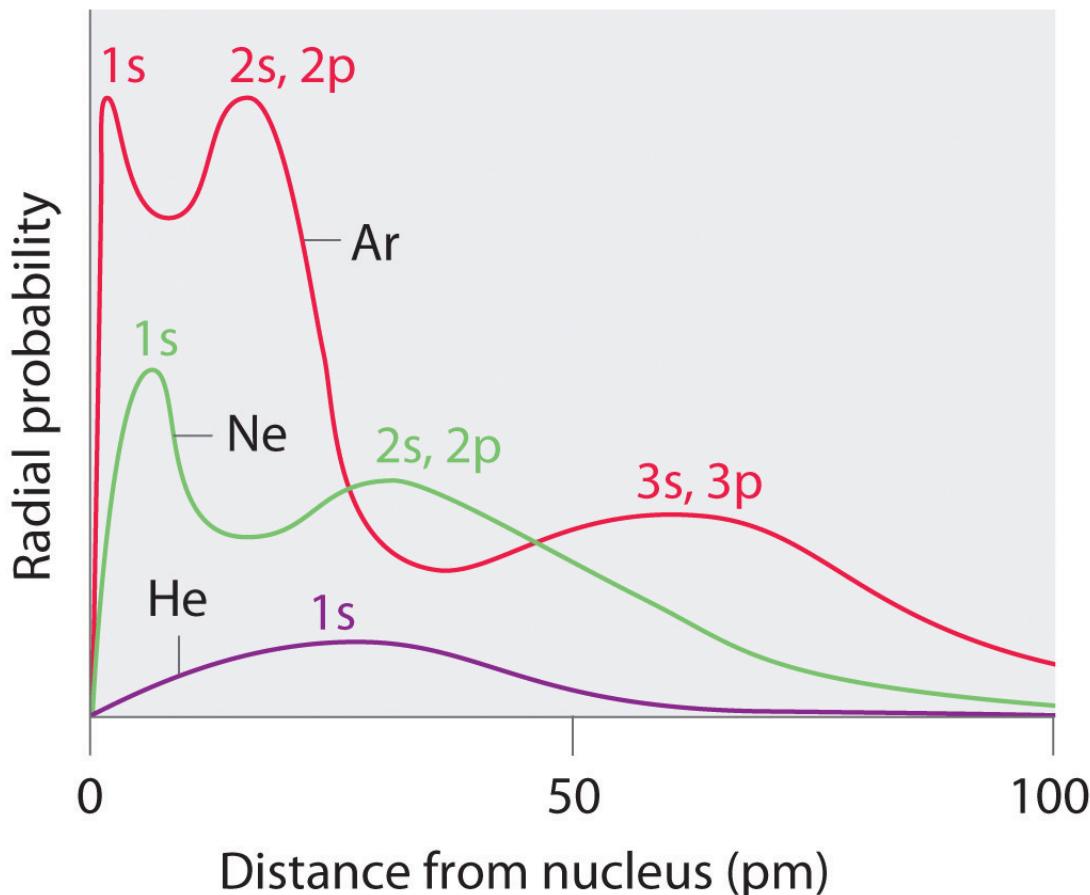


Figure 5.3: Radial probability distributions for He, Ne, and Ar. Taken from *Principles of General Chemistry*, available through the Creative Commons.

### 5.3 Angular Momentum Coupling in Atoms

Once the configuration of the atom (the filling of the orbitals) is known, the next step is to decide how any residual angular momenta will couple into orbital, spin, and total angular momentum for the atom (referred to as the “term”). We will focus on the first two periods of the table (H through Ne), as shown in the right-hand column of table 5.2. We now need to consider how the angular momenta will interact to form the ground term. In doing so, recall that filled shells have zero net orbital and spin angular momentum, and that there are  $2(2l + 1)$  possible states for a given value

of  $l$  and the number of allowed  $l$  values is equal to  $n$ . For helium, the K shell is filled, all spins are paired and the two electrons orbital angular momenta are zero, so the ground term of helium is a  $^1S_0$  signifying zero orbital angular momentum ( $S$  state) and zero total angular momentum (the zero in the subscript) and a multiplicity of 1 (a single term). Lithium must then have an electron in the L shell and it will have an unpaired spin. Since it is the 2s orbital, it has no orbital angular momentum. Thus the ground term for lithium is  $^2S_{1/2}$ , like hydrogen. Beryllium has two paired electrons in the 2s orbital, so it will have zero spin and zero orbital angular momentum; its ground term is  $^1S_0$ . Neon will fill the L shell, with all electron spins paired and equal numbers of electrons with values of +1 and -1 for  $m_l$ . So, like helium, neon will have a ground term of  $^1S_0$ . We can generalize these results to all the alkalis (like Lithium) and all the “inert gases” like helium.

Boron has one unpaired electron, so, like hydrogen, its ground term is a doublet, but since the electron is in a p state, we have  $L = l = 1$  and the ground term is  $^2P_{1/2}$ . Fluorine is missing one electron to fill the L shell, so, like boron, it has a  $^2P$  ground term. The orbital and spin angular momenta can add to either  $1/2$  or  $3/2$ . For reasons we will explore later, the  $J = 3/2$  state is slightly lower, so the ground term is  $^2P_{3/2}$ . Making analogies between electrons on the left of the table and missing electrons (or holes) on the right side is generally useful. To attack the more complex center of the periodic table, we need to have more understanding of how angular momenta will couple in multi-electron atoms.

Table 5.2: The First Two Periods with Terms

Element	K			Ground Term
	1s	2s	2p	
H	1	0	0	$^2S_{1/2}$
He	2	0	0	$^1S_0$
Li	2	1	0	$^2S_{1/2}$
Be	2	2	0	$^1S_0$
B	2	2	1	$^2P_{1/2}$
C	2	2	2	$^3P_0$
N	2	2	3	$^4S_{3/2}$
O	2	2	4	$^3P_2$
F	2	2	5	$^2P_{3/2}$
Ne	2	2	6	$^1S_0$

As we know from helium, the electrons interact via exchange forces, which cause states of different total spin  $S$  to have different energies. Thus “electrostatic” forces tend to control the total spin. The quotation marks are to remind us that we are really dealing with the consequences of the fact that electrons are fermions and try to avoid being close to each other. In atoms of modest  $Z$ , the spins couple to form a total spin,

$$S \equiv \sum_{i=1}^N \vec{S}_i \quad (5.13)$$

While it is less obvious for orbital angular momenta, the sense of rotation also affects how close the electrons can get to each other. An analogy may help: if cars on a race track are all going around

the same way, there are fewer collisions. The net result is that orbital angular momenta couple into a total  $L$ :

$$L \equiv \sum_{i=1}^N \vec{L}_i \quad (5.14)$$

As a result of these couplings, the individual angular momenta operators ( $\vec{L}_i$  and  $\vec{S}_i$ ) fail to commute with the Hamiltonian once we add these (non-central) Coulomb and exchange forces. In this approximation, however, the **total** orbital ( $\vec{L}$ ) and spin ( $\vec{S}$ ) operators are constants of the motion (commute with the Hamiltonian). The good quantum numbers are then  $L, S, M_L, M_S$ , where  $M_L$  and  $M_S$  are projections on the laboratory  $z$  axis.

If we now add magnetic perturbations (spin-orbit, spin-spin), the projections on the  $z$  axis of  $\vec{L}$  and  $\vec{S}$  are no longer good quantum numbers, but instead we couple  $\vec{L}$  and  $\vec{S}$  to make  $\vec{J}$ :

$$\vec{J} = \vec{L} + \vec{S} \quad (5.15)$$

Then,  $L, S, J, M_J$  are the appropriate quantum numbers. This form of coupling, where individual orbital and spin angular momenta first couple to make total  $\vec{L}$  and  $\vec{S}$ , then those couple more weakly to form total  $\vec{J}$ , is called **LS** or **Russell-Saunders** coupling. It is a very good approximation for light atoms, where electrostatic effects have energies around 1 eV (e.g., helium), while magnetic effects have energies around  $10^{-4}$  to  $10^{-3}$  (e.g., hydrogen and helium). Thus the energy differences for different  $J$  values are small.

For heavier elements, the situation changes. As  $Z$  increases, electrostatic terms grow more slowly than magnetic forces. Recall from hydrogen that the spin-orbit splitting grows proportional to  $Z^2|E_n^{(0)}| \sim Z^4$  (eq. 3.65). At some point, magnetic forces will overwhelm electrostatic forces, and a different coupling scheme is a better approximation. In this **j-j coupling**, each electrons  $\vec{L}_i$  and  $\vec{S}_i$  couple to form  $\vec{J}_i$  for each electron.

$$\vec{J}_i = \vec{L}_i + \vec{S}_i, \quad i = 1, 2, \dots, N \quad (5.16)$$

Then those couple more weakly to form the total  $\vec{J}$ :

$$\vec{J} = \sum_{i=1}^N \vec{J}_i \quad (5.17)$$

In this case, only  $J$  and  $M_J$  are good quantum numbers. Pure j-j coupling is rare, but it is a good approximation in excited states of heavy elements. More generally, the coupling is intermediate between LS and j-j coupling. This fact has consequences for spectra (§5.5).

Now, let's get back to the terms for the first two periods. Since these are elements of low  $Z$ , we will assume LS coupling dominates. Since closed shells must have  $L = 0$  and  $S = 0$ , we deal only with valence electrons. Guided by our experience with helium, when considering excited states, we will consider only states where a single electron is excited to a higher orbital. We have already dealt with the situation with a single valence electron (or hole). Now consider two valence electrons, but first assume that they are in different orbitals (different  $n, l$ ). These are called “non-equivalent” electrons. The possible values of  $L$  are found by adding the individual  $l$  values of the orbitals. For example, a 1s2p configuration could have only  $L = 1$ ; 2p3d could have  $L = 1, 2, 3$  because of the

triangle inequality:  $|l_1 - l_2| \leq L \leq |l_1 + l_2|$ . The same inequality for spins limits  $S$  to 0 or 1. The allowed values for three non-equivalent electrons can be worked out by adding the  $l$  and  $s$  of the third electron to the  $L$  and  $S$  of the first two. Table 5.3 shows the allowed terms for two **non-equivalent** electrons in p orbitals.

Table 5.3: Allowed Terms and Sublevels for Two Non-Equivalent p Electrons

L	S	J	Term	Sublevels
0	0	0	$^1S$	$^1S_0$
1	0	1	$^1P$	$^1P_1$
2	0	2	$^1D$	$^1D_2$
0	1	1	$^3S$	$^3S_1$
1	1	2,1,0	$^3P$	$^3P_2, ^3P_1, ^3P_0$
2	1	3,2,1	$^3D$	$^3D_3, ^3D_2, ^3D_1$

Remember that table 5.3 is valid only if the electrons are in different orbitals; if two electrons have the same  $l$ , they must have different  $n$ . Thus both  $^1S$  and  $^3S$  are allowed **only** if  $n_1 \neq n_2$ . The non-equivalence of the two electrons is indicated by writing the configuration as ss, not as s<sup>2</sup>. So table 5.3 is useful for excited configurations, but not for the ground term of the second period, which has **equivalent** electrons. To penetrate to the interior of the period (C, O, N) with 2 to 4 **equivalent** electrons, we must consider the constraints enforced by the Pauli principle. Let us begin with two equivalent p electrons ( $p^2$ ). Because  $n_1 = n_2$  and  $l_1 = l_2$ , we must have different quantum numbers for  $m_s$  or  $m_l$ . Table 5.4 gives all the possible combinations.

Table 5.4: States Allowed by the Pauli Principle for  $p^2$  Configuration

Number	$m_l$			$M_L$	$M_S$
	-1	0	+1		
1	$\uparrow\downarrow$			-2	0
2		$\uparrow\downarrow$		0	0
3			$\uparrow\downarrow$	2	0
4	$\uparrow$	$\uparrow$		-1	1
5	$\uparrow$	$\downarrow$		-1	0
6	$\downarrow$	$\uparrow$		-1	0
7	$\downarrow$	$\downarrow$		-1	-1
8	$\uparrow$		$\uparrow$	0	1
9	$\uparrow$		$\downarrow$	0	0
10	$\downarrow$		$\uparrow$	0	0
11	$\downarrow$		$\downarrow$	0	-1
12		$\uparrow$	$\uparrow$	1	1
13		$\uparrow$	$\downarrow$	1	0
14		$\downarrow$	$\uparrow$	1	0
15		$\downarrow$	$\downarrow$	1	-1

Table 5.5: First Implied Terms Table for p<sup>2</sup> Configuration

$M_L$	$M_S$		
	+1	0	-1
+2	0	1	0
+1	1	2	1
0	1	3	1
-1	1	2	1
-2	0	1	0

The next step is to count the number of microstates with given values of  $M_L$  and  $M_S$  and construct the first implied terms table. The numbers of microstates with each set of values of  $M_L$  and  $M_S$  in table 5.4 is shown in table 5.5. Next, we allocate these microstates to terms, starting with the largest possible values of  $L$  and  $S$ . From table 5.3, we find a  $^3D$  term, which requires microstates with  $L = 2$  and  $S = 1$ . If this term existed, table 5.5 would have to have non-zero entries for  $M_L = \pm 2$  and  $M_S = \pm 1$ , but those entries are zero. Thus the Pauli principle does not allow a  $^3D$  term for two equivalent p electrons. However, a microstate for  $^1D$  does exist with  $M_L = 2$ ,  $M_S = 0$ . This term requires 5 microstates ( $M_L = 2, 1, 0, -1, -2$  with  $M_S = 0$ ). Subtract these from table 5.5 to form the second implied terms table (table 5.6).

Table 5.6: Second Implied Terms Table for p<sup>2</sup> Configuration

$M_L$	$M_S$		
	+1	0	-1
+2	0	0	0
+1	1	1	1
0	1	2	1
-1	1	1	1
-2	0	0	0

Next consider  $^3P$  ( $L = 1$ ,  $S = 1$ ). We have the required microstates ( $M_L = 1, 0, -1$ , with  $M_S = 1, 0, -1$ ). Using those nine for the  $^3P$  term leaves us with only a single microstate, with  $M_L = 0$ ,  $M_S = 0$ , so the last term allowed is  $^1S$ . Because all the microstates have now been used up, we have  $^1D$ ,  $^3P$ , and  $^1S$  as the allowed terms. Similar, but tedious, procedures for other numbers of equivalent electrons leads to the following table of allowed terms (table 5.7).

We have seen that a given electron configuration could have a number of different terms, depending on how the angular momenta line up. How do we decide which has the lowest energy? Detailed calculations are quite difficult, but there are some rules that work quite well for the **ground** configuration. These are called **Hund's Rules**.

1. The terms with the highest multiplicity ( $2S + 1$ ) lie lowest in energy.
2. If several terms have the same multiplicity, the one with the largest  $L$  lies lowest in energy.

Table 5.7: Allowed Terms for Equivalent Electrons

Electron Configuration	Terms
s <sup>2</sup>	<sup>1</sup> S
p <sup>2</sup>	<sup>1</sup> S, <sup>1</sup> D, <sup>3</sup> P
p <sup>3</sup>	<sup>4</sup> S, <sup>2</sup> D, <sup>2</sup> P
p <sup>4</sup>	<sup>1</sup> S, <sup>1</sup> D, <sup>3</sup> P
p <sup>5</sup>	<sup>2</sup> P
p <sup>6</sup>	<sup>1</sup> S
d <sup>2</sup> , d <sup>8</sup>	<sup>1</sup> S, <sup>1</sup> D, <sup>1</sup> G, <sup>3</sup> P, <sup>3</sup> F
d <sup>3</sup> , d <sup>7</sup>	<sup>2</sup> P, <sup>2</sup> D(2), <sup>2</sup> F, <sup>2</sup> G, <sup>2</sup> H, <sup>4</sup> P, <sup>4</sup> F
d <sup>4</sup> , d <sup>6</sup>	<sup>1</sup> S(2), <sup>1</sup> D(2), <sup>1</sup> F, <sup>1</sup> G(2), <sup>1</sup> I, <sup>3</sup> P(2), <sup>3</sup> D, <sup>3</sup> F (2), <sup>3</sup> G, <sup>3</sup> H, <sup>5</sup> D
d <sup>5</sup>	<sup>2</sup> S, <sup>2</sup> P, <sup>2</sup> D(3), <sup>2</sup> F(2), <sup>2</sup> G(2), <sup>2</sup> H, <sup>2</sup> I, <sup>4</sup> P, <sup>4</sup> D, <sup>4</sup> F, <sup>4</sup> G, <sup>6</sup> S

Note, the (2) or (3) means that more than one term of that type is allowed.

3. If several sublevels (different  $J$  values) have the same multiplicity and  $L$ , the sublevel with the lowest  $J$  lies lowest in energy if the configuration has a shell which is less than half filled. If a shell is more than half filled, the maximum  $J$  sublevel lies lowest in energy.

The first Hund's rule is a generalization of the result we found for helium, where the triplet state is lower in energy because the electrons with parallel spin will avoid each other more and contribute less positive energy. The second Hund's rule follows from the argument just before equation 5.14: if all the electrons are "going around the same way," they can remain more separated than if some are going the "wrong" way. I know of no simple argument for rule number 3. The set of sublevels in a term is called a multiplet. If the shell is less than half full, we get "regular" multiplets; if more than half full, they are called "inverted" multiplets. Recall that the splitting between sublevels is caused by magnetic interactions, so will be substantially less than the others for light elements.

Armed with this massive set of tables and rules, we are finally ready to attack the center of the second row of the periodic table! We already argued that boron has a ground term of <sup>2</sup>P, but we can now use Hund's rule number 3 to say that the ground sublevel is <sup>2</sup>P<sub>1/2</sub>. Fluorine, with one missing electron, will also have a ground term of <sup>2</sup>P, but Hund's rule number 3 tells us that the lowest sublevel is <sup>2</sup>P<sub>3/2</sub>. Now consider carbon, with two equivalent p electrons in configuration p<sup>2</sup>. As we worked out in excruciating detail, the allowed terms are <sup>1</sup>S, <sup>1</sup>D, and <sup>3</sup>P. Hund's rule number 1 tells us that the <sup>3</sup>P is the ground term, and rule number 3 selects <sup>3</sup>P<sub>0</sub> as the lowest sublevel. For oxygen, with p<sup>4</sup>, we treat the two missing electrons as holes, so the allowed terms are the same as for carbon, but Hund's rule 3 tells us that the lowest sublevel is <sup>3</sup>P<sub>2</sub>. Now we come to nitrogen. Table 5.7 tells us that the allowed terms are <sup>2</sup>P, <sup>2</sup>D, and <sup>4</sup>S. Hund's first rule selects <sup>4</sup>S as the lowest term. Because the shell is exactly half full, Hund's rule number 3 offer no guidance, but only one  $J$  value is allowed since  $L = 0$ :  $J = 3/2$ . So, the lowest level is <sup>4</sup>S<sub>3/2</sub> and there are no other sublevels in the ground term.

Electron configurations and ground terms can be found for the other atoms online. We often deal with ions in astronomy. These can be analyzed in terms of a neutral species of lower  $Z$ . For example, He<sup>+</sup> has hydrogenic energy levels, but lying much deeper because of the extra charge. Likewise, both O<sup>+2</sup> and C<sup>+</sup> would act like neutral boron, etc. Sequences of nuclei with the same

electronic configurations are called isoelectronic sequences. Astronomers use the notation for ions of the form CI for neutral carbon, CII for singly ionized carbon, etc. though these properly apply only to the spectra of these species.

## 5.4 Structure of Energy Levels in Atoms

We are of course interested in more than the ground terms because we obtain information from transitions between states, particularly between excited states and the ground state. We start with some generalities about the structure. Generally, the energy levels have the following hierarchy of energies: different configurations have differences of several to several 10s of eVs (transitions in the ultra-violet or visible), different terms within a single configuration differ in energy by about 1 eV (transitions in the visible to near-infrared), different sublevels ( $J$  values) within a term differ by about 0.01 eV (“fine-structure” transitions in the far-infrared), and different nuclear spin states differ by about  $10^{-5}$  to  $10^{-4}$  eV (“hyperfine-structure” transitions in the radio). in the radio. These general guidelines must be taken with care. As electrons get into very excited orbitals, the energy levels get much closer together, just as for hydrogen. This is reasonable because an electron in a distant orbital sees a fully screened nucleus, which looks like a hydrogen nucleus.

The energy differences between terms are caused by the different electron interactions with different arrangements of angular momentum, as captured qualitatively by Hund’s rules. In simple form, these arise from the second term in equation 5.11, expressed there as a screening function. Calculating these energy differences is non-trivial. In ordering excited terms, Hund’s rules are often a good guideline, but are less predictive than for the ground term.

At the next level down, the sublevels within a term arise from magnetic effects, like spin-orbit and spin-spin. If we consider a single term well described by Russell-Saunders coupling, the Hamiltonian for the spin-orbit effect is simply given by

$$H_{so} = A \vec{L} \cdot \vec{S} \quad (5.18)$$

Of course, this very term makes  $\vec{L}$  and  $\vec{S}$  no longer commute with the Hamiltonian, but  $\vec{J}$  commutes, as do  $L^2$  and  $S^2$ . Following a development like that for hydrogen, we have

$$\begin{aligned} \langle H_{so} \rangle &= \langle \gamma, L, S, J, M_J | H_{so} | \gamma, L, S, J, M_J \rangle = \\ &\frac{1}{2} A \langle \gamma, L, S, J, M_J | J^2 - L^2 - S^2 | \gamma, L, S, J, M_J \rangle = \\ &\frac{1}{2} A [J(J+1) - L(L+1) - S(S+1)] \end{aligned} \quad (5.19)$$

Within a given term (fixed  $L$  and  $S$ , the energy difference between sublevels reduces to

$$E(J) - E(J-1) = AJ \quad (5.20)$$

If  $A > 0$ , higher  $J$  levels are higher, giving a regular multiplet (e.g., carbon in the ground  ${}^3P$  term). If  $A < 0$ , lower  $J$  levels are higher, giving an inverted multiplet (e.g., oxygen in the ground  ${}^3P$  term).

As with hydrogen, interactions of nuclear spin with total angular momentum lead to hyperfine splitting. We add  $\vec{J}$  and  $\vec{I}$  to get  $\vec{F}$ . Transitions between hyperfine levels in atoms other than hydrogen, deuterium, and  ${}^3\text{He}$  are not observed in astronomy, but a hyperfine transition in rubidium at

6835 MHz is an important frequency standard, and the hyperfine transition in cesium at 9192.631770 MHz is used in atomic clocks.

Finally, an external magnetic field can produce splittings, parallel to those for hydrogen.

## 5.5 Transitions and Spectra

### 5.5.1 General Considerations

The wavelengths of transitions will be determined by the energy level structure and the selection rules. The strength of transitions, hence Einstein A's, will be determined by the matrix element of the relevant operator.

We begin with the most general case, where the quantum numbers are  $\gamma$ ,  $J$ , and  $M_J$ , where  $\gamma$  represents a collection of quantum numbers specifying the configuration, for example. We will work in terms of the transition strength (equation 2.80). As long as we are dealing with unpolarized light (not dealing with Zeeman splitting), we can sum over degenerate states of  $M_J$ .

As shown for the electric dipole moment operator in appendix A.5, any operator in an expansion of the single photon electric field in spherical coordinates can be written in the form of spherical tensors,  $T_{kq}$  (appendix A.3). In the absence of external torques, such as interactions with external magnetic fields, we can apply the Wigner-Eckart theorem to remove the dependence on azimuthal quantum numbers. Thus the strength of the transition just depends on the reduced matrix element.

$$S(\gamma', J' \rightarrow \gamma, J) = \sum_q \sum_{MM'} |\langle \gamma JM | T_{kq} | \gamma' J' M' \rangle|^2 = |\langle \gamma J | T_k | \gamma' J' \rangle|^2 \quad (5.21)$$

as long as the 3-J symbol in equation A.24 is not zero. The properties of 3-J symbols then yield our most general one-photon selection rules:

$$|J - J'| \leq k \leq |J + J'| \quad (5.22)$$

and

$$\Delta M = M - M' = q \quad (5.23)$$

This is just a recapitulation of equation 2.66. When we add the parity selection rules from Table 2.2, we have all the most basic selection rules that apply to one-photon selection rules for any system in the absence of external torques.

If the states are well described by LS coupling, we may want to sum over all initial and final  $J$  as well to get the total strength for the multiplet:

$$S(\gamma', L', S' \rightarrow \gamma, L, S) = \sum_{JJ'} S(\gamma', L', S' J' \rightarrow \gamma, L, S J) \quad (5.24)$$

or even the total strength for transitions between configurations.

$$S(\gamma' \rightarrow \gamma) = \sum_{L,S} \sum_{L',S'} S(\gamma', L', S' \rightarrow \gamma, L, S) \quad (5.25)$$

### 5.5.2 Magnetic Dipole Transitions

We will consider magnetic dipole transitions first, as they are actually simpler. In the LS coupling approximation, the magnetic moment operator is

$$\vec{\mu}_M = -\frac{\mu_B}{\hbar}(\vec{L} + 2\vec{S}) \quad (5.26)$$

where  $\vec{L}$  and  $\vec{S}$  are the operators for the total orbital and spin angular momentum. Since we have LS coupling, the eigenfunctions are also eigenfunctions of  $\vec{L}$  and  $\vec{S}$ , so the matrix element will vanish unless  $L' = L$  and  $S' = S$ . The magnetic dipole operator does not affect configuration, so  $\gamma' = \gamma$  as well. So the magnetic dipole operator can only cause “fine structure” transitions between different  $J$  values of a multiplet. They are important because these are not generally caused by electric dipoles. Then (cf. Eqn. 3.137 in §3.5.4), we have

$$S(\gamma, S, L, J \rightarrow \gamma, S, L, J-1) = \mu_B^2 \frac{(L+S+J+1)(L+S-J+1)(S+J-L)(J+L-S)}{4J} \quad (5.27)$$

with selection rules

$$\Delta L = 0, \quad \Delta S = 0, \quad \Delta \gamma = 0, \quad \Delta J = \pm 1 \quad (5.28)$$

Note that no radial integral is involved, so the calculation is very easy.

### 5.5.3 Electric Pole Transitions

In contrast to the ease of calculating magnetic dipole transitions, electric dipole transitions are difficult because of the radial integrals. We will give a general treatment to illustrate the ideas.

First, let's take a closer look at  $S(\gamma', J' \rightarrow \gamma, J)$ . If the operator that effects the transition commutes with  $\vec{S}$  (the total spin operator), we also have

$$S(\gamma', J', L', S' \rightarrow \gamma, J, L, S) = |\langle \gamma J, L, S | T_k | \gamma' J', L', S' \rangle|^2 = |\langle \gamma L | T_k | \gamma' L' \rangle|^2 (2J'+1)(2J+1) \left| \begin{Bmatrix} L & J & S \\ J' & L' & k \end{Bmatrix} \right|^2 \delta_{S'S} \quad (5.29)$$

where the entity in curly brackets is a 6-J symbol. A good reference for these is by Sobelman. There is a triangle condition on the quantities with  $L$ ,  $L'$ , and  $k$ :

$$\Delta(L, L', k) \Rightarrow \Delta L = 0, \pm 1, \dots \pm k; \quad L' + L \geq k \quad (5.30)$$

and we clearly have  $\Delta S = 0$ .

So, to the extent that we have pure LS coupling and no mixing from other configurations, we can add the following selection rules for electric multipole transitions. For electric dipole ( $k = 1$ , E1), we have

$$\Delta S = 0; \quad \Delta L = 0, \pm 1; \quad L' + L \geq 1 \quad (L = 0 \not\rightarrow 0) \quad (5.31)$$

while electric quadrupole transitions ( $k = 2$ , E2) in this limit would follow

$$\Delta S = 0; \quad \Delta L = 0, \pm 1, \pm 2; \quad L' + L \geq 2 \quad (L = 0 \not\rightarrow 0, \quad L = 1 \not\rightarrow 0) \quad (5.32)$$

Since the 6-J symbols obey the sum rule,

$$\sum_{J'} (2J' + 1) \left\{ \begin{array}{ccc} L & J & S \\ J' & L' & 1 \end{array} \right\}^2 = \frac{1}{2L+1} \quad (5.33)$$

and

$$\sum_J (2J + 1) = (2L + 1)(2S + 1) \quad (5.34)$$

we have for the total strength of the multiplet,

$$S(\gamma', S, L \rightarrow \gamma, S, L) = \sum_{J' J} S(\gamma', S', L', J' \rightarrow \gamma, S, L, J) = (2S + 1) |\langle \gamma L || T_1 || \gamma' L' \rangle|^2 \quad (5.35)$$

To go further, we need to look at the operator for the electric dipole moment. The electric dipole moment operator would be the sum of the individual electron dipole moment operators:

$$\vec{\mu} = \sum_{k=1}^N \vec{\mu}_k = \sum_{k=1}^N -e\vec{r}_k \quad (5.36)$$

In the general case, the eigenfunctions will be sums of Slater determinants:

$$\psi(\gamma JM) = \sum_i c_i D(\psi^i) \quad (5.37)$$

where  $D(\psi^i)$  is a Slater determinant and  $\psi^i$  is shorthand for the  $N$  orbitals:

$$\psi^i = (\psi_1^i \psi_2^i \dots \psi_N^i) \quad (5.38)$$

in parallel with the equations 5.9 and 5.6. The individual orbitals carry the single particle quantum numbers. For example, for electron 1,

$$\psi_1^i = |n_1 l_1, m_{l_1}, m_{s_1}\rangle \quad (5.39)$$

which specify the state of a single electron. Now consider the matrix element between initial state  $i$  and final state  $f$   $[\psi(\gamma' J' M' \rightarrow \psi(\gamma JM))]$ . We are running out of indices, so we will use  $p$  and  $q$  as summation indices for the final and initial states, respectively:

$$\begin{aligned} \langle \gamma JM | \sum_{k=1}^N \vec{\mu}_k | \gamma' J' M' \rangle &= \langle \sum_p c_p^f D(\psi^p) | \sum_{k=1}^N \vec{\mu}_k | \sum_q c_q^i D'(\psi^q) \rangle = \\ \sum_p \sum_q c_p^{f*} c_q^i \langle D(\psi^p) | \sum_{k=1}^N \vec{\mu}_k | D'(\psi^q) \rangle &= \sum_p \sum_q c_p^{f*} c_q^i \sum_{k=1}^N \langle D(\psi^p) | \vec{\mu}_k | D'(\psi^q) \rangle \end{aligned} \quad (5.40)$$

Note that each term in the sum over  $k$  is the matrix element of a single particle dipole moment (or position) operator. Therefore, if  $D(\psi^p)$  and  $D'(\psi^q)$  differ by more than one set of  $\psi^i$ , the orthogonality of the single particle orbitals will make the matrix element zero. So each term in the sum over  $p$  and  $q$  will contribute only one non-zero matrix element (only that electron which changes its orbital). If we have strict LS coupling and no mixing of different configurations, there is only

one term in the sums over  $p$  and  $q$ , so only one electron can “jump” (change its orbital) at a time. Since the determinants then differ only in a single pair of orbitals, we can think of a single electron changing from orbital  $u$  to orbital  $v$ . Then one can rearrange the Slater determinant by doing  $t$  permutations to get the simple result that

$$\langle D(\psi^p) | \vec{\mu}_k | D'(\psi^q) \rangle = (-1)^t \langle n_v l_v m_{l_v} m_{s_v} | -e \vec{r}_k | n_u l_u m_{l_u} m_{s_u} \rangle \quad (5.41)$$

so we recover in this limit, the single particle selection rules and transition strengths. For example, we get the  $\Delta s = 0$ ,  $\Delta l = \pm 1$  rule for the “jumping” electron, as in hydrogen (§3.5.1). These rules apply only in the limit of LS coupling. For j-j coupling, one would get  $\Delta j = 0, \pm 1$ . In these simple limits, one might even calculate the value of the matrix element. The angular part will be the same as for hydrogen, but the radial part will need to use some appropriate radial function. Many situations are not perfectly in either limit of coupling, and many states involve mixings of different configurations, so some important transitions in astronomy are due to this hidden configuration interaction. A famous example is that of Bowen fluorescence in OIII.

The LS selection rules,  $\Delta L = \pm 1$  and  $\Delta S = 0$  for electric dipole transition will not apply in j-j coupling and are only approximate even in atoms close to LS coupling. This is the origin of the weak transitions between spin states (called “intercombination lines”) in helium. These become much stronger as  $Z$  increases. For pure j-j coupling, the electric dipole selection rules would be  $\Delta J = 0, \pm 1$ ,  $\Delta j_i = 0, \pm 1$ , and parity must change.

## 5.6 Elements of Astrophysical Interest

The ideas in this chapter are best illustrated with some actual examples, and we might as well focus on the most common elements. For now, we have only carbon.

### 5.6.1 Carbon

Carbon, a product of helium burning, has a relatively large cosmic abundance. As we might expect, it is important in a wide range of astrophysical environments, ranging from QSOs to molecular clouds. The ionization potentials are given in Table 5.8.

Table 5.8: Ionization Potentials in eV of Carbon						
C	C <sup>+</sup>	C <sup>+2</sup>	C <sup>+3</sup>	C <sup>+4</sup>	C <sup>+5</sup>	C <sup>+6</sup>
11.26030	24.38332	47.8878	64.4939	392.087	489.99334	...

From CRC values on Wikipedia page.

Note the big jump in energy required to ionize the last two (1s) electrons. Because the ionization of neutral carbon takes less energy than it does to ionize hydrogen, but ionizing C<sup>+2</sup> requires substantially more energy, we expect C<sup>+</sup> to be the dominant form in HII regions, while higher ionization states will exist only in regions with higher energy photons or very hot gas. Around O stars or in planetary nebulae, where helium can be ionized, substantial C<sup>+2</sup> will exist. In regions of atomic hydrogen, C<sup>+</sup> is the main form and it is an important coolant. As one moves into denser regions of molecular hydrogen, there is a transition from C<sup>+</sup> to neutral C to CO. C<sup>+5</sup> would be hydrogenic and the equivalent transition to Lyman  $\alpha$  would occur at  $1215\text{\AA}/36 = 33.8\text{\AA}$ .

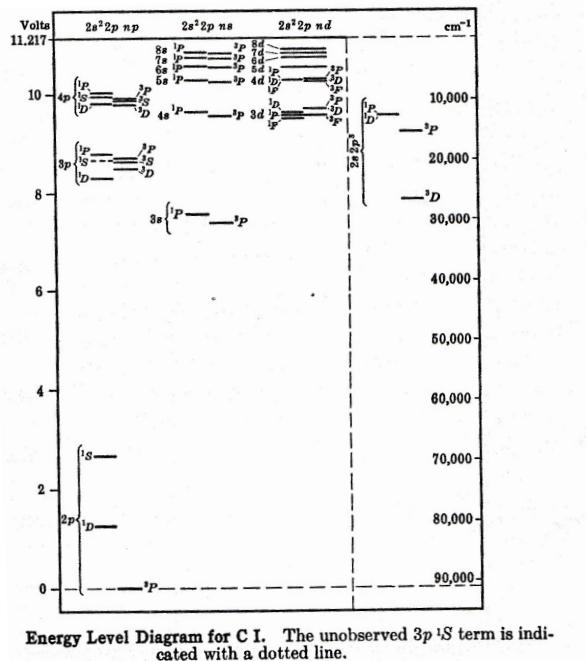
The sequence from C to C<sup>+5</sup> has configurations and ground terms in table 5.9.

Table 5.9: Configurations and Ground Terms of Carbon

C	C <sup>+</sup>	C <sup>+2</sup>	C <sup>+3</sup>	C <sup>+4</sup>	C <sup>+5</sup>
CI	CII	CIII	CIV	CV	CVI
2s <sup>2</sup> 2p <sup>2</sup>	2s <sup>2</sup> 2p	2s <sup>2</sup>	2s	1s <sup>2</sup>	1s
<sup>3</sup> P <sub>0</sub>	<sup>2</sup> P <sub>1/2</sub>	<sup>1</sup> S <sub>0</sub>	<sup>2</sup> S <sub>1/2</sub>	<sup>1</sup> S <sub>0</sub>	<sup>2</sup> S <sub>1/2</sub>

### Neutral Carbon

The energy level or **Grotian** diagram for neutral carbon is shown in figure 5.4. This is an old figure so the numbers are not very accurate, but it is the best I have found to get the big picture.



configurations do not follow Hund's rules.

The transition from the ground  $^3P_0$  term to the  $2s^22p3s\ ^3P$  term will occur at about 1700 Å, so might be seen in absorption in front of hot stars. The higher levels are unlikely to be populated in interstellar clouds, but could be in stellar atmospheres. They are separated by about  $10^4 \text{ cm}^{-1}$  or less, so lie in the far-red to infrared to radio, as  $n$  increases. As with hydrogen, transitions between very high levels tend to arise from recombination cascades, so are called carbon recombination lines. The transitions between ground terms also are separated by about  $10^4 \text{ cm}^{-1}$  and are forbidden by electric dipole. Thus CI lines are not expected to be prominent in visible spectra of stars.

The more detailed structure of the ground configuration is shown in figure 5.5. This figure shows the splitting into three terms, with order following Hund's rules and separated by about  $10^4 \text{ cm}^{-1}$ , with transitions between them occurring at wavelengths around 10,000 Å. Finally, the magnetic effects split the ground term,  $^3P$  into three sublevels, a regular multiplet by Hund's rule 3. Transitions between these sublevels occur in the submillimeter wavelength region at frequencies around 492 GHz (610 μm) and 609 GHz (370 μm). These magnetic dipole transitions are important probes of gas in the transition region between atomic and molecular hydrogen. In principle, they could provide a good temperature measurement of this gas, but optical depth effects complicate the interpretation.

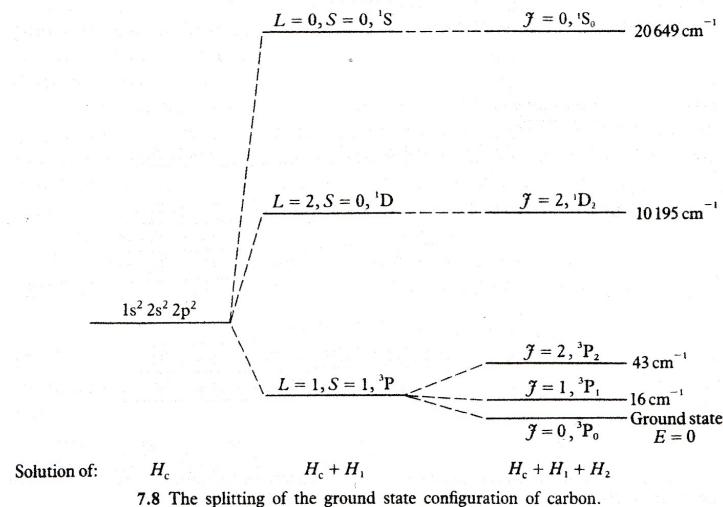


Figure 5.5: Splitting of ground configuration of carbon into terms and sublevels. Figure from Bransden and Joachain, *Physics of Atoms and Molecules*, pg. 348.

### Singly Ionized Carbon ( $\mathbf{C}^+$ )

Carbon missing one electron is isoelectronic with Boron, with a  $2s^22p$  ground configuration and  $^2P_{1/2}$  ground term. The most important transition is between the two sublevels of that term. This magnetic dipole fine structure transition from  $^2P_{3/2}$  to  $^2P_{1/2}$  occurs at 157.7 μm and is an important coolant of regions of atomic hydrogen, particularly those heated by proximity to massive stars. It has been detected in many galaxies, even to very high redshift. Other lines tend to be in the ultraviolet.

The transition from  $2s^22p\ ^2P_{1/2}$  to  $2s2p^2\ ^4P_{1/2}$  state at 2324.69 Å has been observed in absorption. Note that  $\Delta S = 2$  for this transition, so it would be forbidden under strict LS coupling.

### Doubly Ionized Carbon ( $C^{+2}$ )

Carbon missing two electrons is isoelectronic with Be and has a ground configuration of  $1s^22s^2$  and ground term of  ${}^1S_0$ . Thus, there is no fine structure in the ground term and all transitions from there are to excited configurations. Two transitions to  $1s^22s2p\ ^3P$  with three sublevels,  $J = 0, 1, 2$ , are observed in the ultraviolet around 1909 Å in planetary nebulae and in the optical from moderate redshift QSOs (e.g., Shields et al. 1981, Ap. J., 248, 569). These transitions provide instructive examples in how “forbidden” transitions can happen. All are intercombination bands because  $\Delta S \neq 0$ , so require a breakdown of strict LS coupling. The transition to the  ${}^3P_1$  state occurs by a weak electric dipole transition at 1909 Å because of the breakdown of LS coupling. The story is that the state we call  ${}^3P_1$  is really a mixture of two configurations:

$$|“2s2p\ ^3P_1”\rangle \approx 0.9993|2s2p\ ^3P_1\rangle - 0.0007|2s2p\ ^1P_1\rangle \quad (5.42)$$

because of spin-orbit interaction. Then

$$|\langle 2s^2\ ^1S_0 | \mu | “2s2p\ ^3P_1” \rangle|^2 \approx (0.0007)^2 |\langle 2s^2\ ^1S_0 | \mu | 2s2p\ ^1P_1 \rangle|^2 \quad (5.43)$$

which yields an Einstein A of  $96.8\text{ s}^{-1}$ .

The  ${}^3P_2 \rightarrow {}^1S_0$  transition cannot be electric dipole because  $\Delta J = 2$ . One’s first thought is electric quadrupole, but the two states have opposite parity, so this transition actually occurs via magnetic quadrupole. The resulting Einstein A value is only  $5.4 \times 10^{-3}\text{ s}^{-1}$  and the  ${}^3P_2$  level is thus collisionally de-excited for densities of colliders (electrons most likely) of  $4 \times 10^5\text{ cm}^{-3}$ , providing a density probe. Finally  ${}^3P_0 \rightarrow {}^1S_0$  is strictly forbidden by the  $J = 0 \nrightarrow 0$  rule.

## Chapter 6

# Introduction to Molecules in Astronomy

### 6.1 Why, Where, and What?

We study molecules because they are playing an increasingly important role in modern astronomy. They are found in comets, planetary atmospheres, relatively cool stellar atmospheres, and galaxies out to redshifts greater than 6. Molecular gas is the fuel for star formation and star formation in other galaxies is highly correlated with molecular gas. Molecules played an important role even for the first stars, but current star formation is completely dominated by molecular gas. Understanding the spectroscopy of molecules will allow us to more deeply understand their role in both theory and observations.

An updated list of the interstellar and circumstellar molecules can be found on wikipedia.<sup>1</sup> with the count as of 2013 at about 200 species. Most are small molecules with less than 5 atoms, but this is at least partially a selection effect. Larger, more complex molecules are harder to detect.

A few molecules of particular astronomical relevance include H<sub>2</sub>, CO, CS, H<sub>2</sub>O, HCO<sup>+</sup>, HCN, NH<sub>3</sub>, H<sub>2</sub>CO, CH<sub>3</sub>CN, and CH<sub>3</sub>OH.

### 6.2 The Born-Oppenheimer Equation

A straightforward attempt at an exact solution for a complex molecule would be extremely challenging. Instead, our understanding of molecules rests on many approximations, the most fundamental of which is the Born-Oppenheimer approximation. The key insight is that the forces on electrons and nuclei are comparable, but the masses of the electrons are much smaller. As a result, the electrons move much faster, and the nuclei can be considered to be stationary during the electronic motion. Then the nuclei move in the “smeared-out” potential of the electrons. Thus, we can approximately separate the problems of the electronic motion and the nuclear motion.

Before looking at the formal statement of the Born-Oppenheimer approximation, let us consider some order-of-magnitude estimates. If a typical separation between nuclei in a molecule is denoted

---

<sup>1</sup>[http://en.wikipedia.org/wiki/List\\_of\\_interstellar\\_and\\_circumstellar\\_molecules](http://en.wikipedia.org/wiki/List_of_interstellar_and_circumstellar_molecules)

by  $a$ , then the electron is confined to a similar length scale. The uncertainty principle then requires the momentum to be about  $\hbar/a$ , so that a typical electronic energy is  $E_e \sim \hbar^2/ma^2$ , where  $m$  is the electron mass. Now consider the nuclear motion, which consists of translation, vibration, and rotation. We can separate the translational degrees of freedom for an isolated system by transforming to the center of mass coordinates. Next consider vibrational motions. Both electrons and nuclei are bound by comparable forces, which we can approximate by a harmonic oscillator potential,  $E = \langle 0.5kx^2 \rangle \hbar\omega_e$ , with  $\omega_e = \sqrt{k/m}$ . The characteristic frequency of the nuclear motion will be  $\omega_N = \sqrt{k/M}$ , where  $M$  is the mass of the nucleus. Thus the energy of nuclear motion will be less than that of the electronic motion by about  $\sqrt{m/M}$ .

With typical dimensions,  $a \sim 1$  Angstrom, the electronic energy is a few eV, and vibrational energies are about 100 times less. The ratio is least for light nuclei, with H<sub>2</sub> being the molecule with the lowest ratio of electronic to vibrational energy. The energy can be thought of as an expansion in powers of  $\sqrt{m/M}$ , with the rotation energy lower by another factor (recall equation 1.44):

$$E_{rot} \sim \hbar^2 J(J+1)/2I = \hbar^2 J(J+1)/2Ma^2 \sim E_e m/M < 0.001 \text{eV}. \quad (6.1)$$

As a result, there is a hierarchy of energy levels. Each electronic state has many, more closely spaced vibrational states, and each vibrational state has many, more closely spaced, rotational states.

With these order of magnitude estimates in mind, let us proceed to the formalities of the Born-Oppenheimer approximation. The Hamiltonian of a molecule with  $n$  electrons and  $N$  nuclei can be written as  $H = T_N + T_e + V$  with  $V$  further expanded to  $V = V_{ee} + V_{NN} + V_{eN}$ , where  $T_N$  is the kinetic energy of the nuclei,  $V_{ee}$  is the interaction potential of an electron with other electrons, etc. We will use Greek indices for nuclei and roman indices for the electrons. Then

$$T_N = \sum_{\alpha}^N \frac{-\hbar^2}{2M_{\alpha}} \nabla_{\alpha}^2 \quad (6.2)$$

$$T_e = \sum_i^n \frac{-\hbar^2}{2m} \nabla_i^2 \quad (6.3)$$

$$V_{ee} = \sum_{\substack{i,j=1 \\ i>j}}^n \frac{e^2}{|\vec{r}_i - \vec{r}_j|} \quad (6.4)$$

$$V_{NN} = \sum_{\substack{\alpha,\beta=1 \\ \alpha>\beta}}^N \frac{Z_{\alpha}Z_{\beta}e^2}{|\vec{R}_{\alpha} - \vec{R}_{\beta}|} \quad (6.5)$$

$$V_{eN} = \sum_{i=1}^n \sum_{\alpha=1}^N \frac{Z_{\alpha}e^2}{|\vec{r}_i - \vec{R}_{\alpha}|} \quad (6.6)$$

We could write down all these terms for H<sub>2</sub>, but doing it for CH<sub>3</sub>OH would be quite challenging! Using our order-of-magnitude arguments, we assume that we can ignore the nuclear motion while solving the electronic motion, at least to first order. So, we solve the electronic Schrodinger equation:

$$[T_e(\vec{r}_i) + V_{ee}(\vec{r}_i) + V_{eN}(\vec{r}_i, \vec{R}_{\alpha}) + V_{NN}(\vec{R}_{\alpha})] \Phi_q(\vec{R}_{\alpha}; \vec{r}_i) = E_q(\vec{R}_{\alpha}) \Phi_q(\vec{R}_{\alpha}; \vec{r}_i) \quad (6.7)$$

The wavefunctions and eigenvalues are calculated for each  $\vec{R}_\alpha$ , which is held fixed during the calculation. The solutions,  $\Phi_q(\vec{R}_\alpha; \vec{r}_i)$  form a complete, orthonormal basis set at each  $\vec{R}_\alpha$ , so we can expand the exact wavefunction,  $\psi(\vec{R}_\alpha; \vec{r}_i)$  in terms of them:

$$\psi(\vec{R}_\alpha; \vec{r}_i) = \sum_q F_q(\vec{R}_\alpha) \Phi_q(\vec{R}_\alpha; \vec{r}_i) \quad (6.8)$$

and  $F_q(\vec{R}_\alpha)$  can be interpreted as the wavefunction for the nuclear motion when the electronic system is in state  $q$ . Then if we substitute this expansion into the *full* Schrodinger equation and project out state  $s$ , we get

$$\sum_q \langle \Phi_s | T_N + T_e + V - E | F_q \Phi_q \rangle = 0, \quad s = 1, 2, \dots \quad (6.9)$$

or using equation 6.7,

$$\sum_q \langle \Phi_s | T_N | \Phi_q F_q(\vec{R}_\alpha) \rangle + [E_s(\vec{R}_\alpha) - E] F_s(\vec{R}_\alpha) = 0, \quad s = 1, 2, \dots \quad (6.10)$$

In the general case, both  $\Phi_q$  and  $F_q$  depend on  $\vec{R}_\alpha$ , so

$$T_N(\Phi_q F_q) = \sum_\alpha -\frac{\hbar^2}{2M_\alpha} [F_q \nabla_{\vec{R}_\alpha}^2 \Phi_q + 2(\nabla_{\vec{R}_\alpha} F_q \cdot \nabla_{\vec{R}_\alpha} \Phi_q) + \Phi_q \nabla_{\vec{R}_\alpha}^2 F_q]. \quad (6.11)$$

The Born-Oppenheimer approximation consists formally of neglecting  $|\nabla_{\vec{R}_\alpha} \Phi_q|$  with respect to  $|\nabla_{\vec{R}_\alpha} F_q|$ , and thus keeping only the last term. In this case,  $\Phi_q$  comes through the nuclear Hamiltonian, and  $\langle \Phi_q | \Phi_s \rangle = \delta_{qs}$  leads to the “nuclear” Schrodinger equation:

$$[T_N + E_s(\vec{R}_\alpha) - E] F_s(\vec{R}_\alpha) = 0, \quad s = 1, 2, \dots \quad (6.12)$$

The neglected terms are not important for most molecular structure and spectroscopy problems, but they become important if collisional transitions between electronic energy states are considered.

By separating the electronic and nuclear wave functions, we now have a wave function that is a linear combination of direct product wave functions (see eq. 6.8). Usually, a single term suffices unless one needs corrections to the Born-Oppenheimer results. The “nuclear” wave function,  $F_s(\vec{R}_\alpha)$ , can be further separated into translational, vibrational, and rotational functions. This problem is straightforward, but messy, for the general case of the polyatomic molecule (see Chap. 11 of **Molecular Vibrations** by Wilson, Decius, and Cross), so we will do it only for diatomic molecules, since the principles are the same.

As with atoms, we can separate translational motion (motion of the center of mass frame) by using a relative coordinate,  $\vec{R}$ , and reduced mass  $\mu$ , defined by

$$\vec{R} = \vec{R}_a - \vec{R}_b \quad (6.13)$$

$$\mu = \frac{M_a M_b}{M_a + M_b} \quad (6.14)$$

The solutions to the translational equation in the absence of external forces are the plane waves (see §1.5). The motion in the center of mass frame (relative motion) is that of a single particle of

mass  $\mu$  a distance  $\vec{R}$  from a potential source, which now must be specified. For a diatomic molecule, the potential depends only on the magnitude of  $\vec{R}$ , assuming that there is no residual angular momentum. The correct potential is given by the solutions to the electronic equations,  $E_s(R)$ , but we will use a series of approximations.

The first and simplest approximation is essentially a delta function potential, which gives the rigid rotor approximation discussed in §1.5. The solutions are spherical harmonics  $Y_{JM}(\theta, \varphi)$  with eigenvalues  $E_r = BJ(J + 1)$  with  $B = \hbar^2/2I$  and  $I = \mu R^2$ , where we replace  $\ell$  and  $m$  by  $J$  and  $M$ . We assume, for the moment, that we can, in a generalization of the rigid rotor, continue to use these to represent the rotational  $(\theta, \varphi)$  motion. Then we can write the wave function for the nuclear motion as

$$F_s(\vec{R}) = \psi^{vib}(R)\psi^{rot}(\theta, \varphi), \quad (6.15)$$

with  $\psi^{rot}(\theta, \varphi) = Y_{JM}(\theta, \varphi)$  for the rigid rotor.

The next better approximation is to consider the bond not to be completely rigid, and to instead describe it as a spring with spring constant  $k$ . Removing the electronic energy at  $R = R_0$ , we have  $E_s(R) = E_s(R_0) + \frac{1}{2}k(R - R_0)^2$ . The solution to this problem is also well known. The wave functions are related to Hermite polynomials and the eigenvalues are  $E_v = \hbar\nu(v + \frac{1}{2})$ , where  $v$  is the vibrational quantum number and  $\nu = \frac{1}{2\pi}\sqrt{k/\mu}$ .

As a result, we factored the wave function into electronic and nuclear wave functions and further factored the nuclear wave functions into translational, vibrational, and rotational wave functions:

$$\psi(\vec{r}_i; \vec{R}_\alpha) = \Phi_q(\vec{R}_\alpha; \vec{r}_i)\psi^{trans}(\vec{R}_{cm})\psi^{vib}(R)\psi^{rot}(\theta, \varphi) \quad (6.16)$$

with corresponding energies

$$E = T_{cm} + E_s(R_0) + E_v + E_r \quad (6.17)$$

The center of mass motion [ $\psi^{trans}(\vec{R}_{cm})$ ] and  $T_{cm}$  are usually neglected unless one is dealing with collisions. The factorization is only approximate, but interactions between these different kinds of motions are usually treated as perturbations. The next improvement to the treatment is to improve the description of the vibrational motion through a more realistic approximation to the internuclear potential. It is useful to write  $E_s = E_s(\infty) + V(R)$  so that  $V(R) \rightarrow 0$  as  $R \rightarrow \infty$ . A good empirical approximation is the Morse potential:

$$V(R) = D_e[e^{-2\alpha(R-R_0)} - 2e^{-\alpha(R-R_0)}]. \quad (6.18)$$

$D_e$  is the dissociation energy of the molecule,  $D_e = E_s(\infty) - E_s(R_0)$ . If we expand the Morse potential about its minimum ( $-D_e$ ) at  $R = R_0$ , we can recover our harmonic oscillator approximation:

$$V(R) = D_e[-1 + \alpha^2(R - R_0)^2 + \dots] \quad (6.19)$$

We identify  $D_e\alpha^2 = \frac{1}{2}k$  in the harmonic oscillator potential.

The Morse potential rises more slowly as  $R \rightarrow \infty$  than it does as  $R \rightarrow 0$ , as the nuclear coulomb potential takes over, but near  $R_0$ , the harmonic oscillator potential is fairly good. Note that the

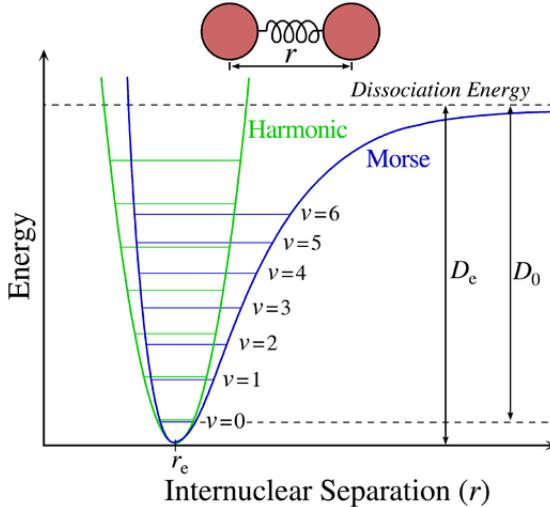


Figure 6.1: The lowest electronic energy state is shown with multiple vibrational energy levels.

zero point energy of vibrations ( $\frac{1}{2}h\nu$ ) causes a correction to the dissociation energy, which is now given by

$$D_0 = D_e - h\nu/2 \quad (6.20)$$

The next figure shows the Morse potential and the harmonic oscillator approximation, while the following figure shows the actual potentials for various states of H<sub>2</sub>, including a repulsive electronic configuration (<sup>3</sup> $\Sigma_u^+$ ).

### 6.3 Symmetries and Nomenclature

The nomenclature for electronic energy states of molecules is parallel to that of atoms, except that Greek letters are used. The key difference between an atom and a molecule is that molecules lack some symmetries of atoms. The internuclear axis of a diatomic molecule, or more generally the symmetry axis of a polyatomic molecule, selects a particular direction in space. Thus the Hamiltonian is invariant with respect to rotation about this axis, but not about the other two axes. Suppose we call the symmetry axis the *z* axis. Then we have  $[H, L_z] = 0$ , but  $[H, L_x] \neq 0$ ,  $[H, L_y] \neq 0$ ,  $[H, L^2] \neq 0$ , where *L* is the total orbital electronic angular momentum operator. We can construct simultaneous eigenfunctions of *H* and *L<sub>z</sub>*, and *M<sub>L</sub>* is a good quantum number:

$$L_z \Phi_s(\vec{R}_\alpha; \vec{r}_i) = M_L \hbar \Phi_s(\vec{R}_\alpha; \vec{r}_i) \quad M_L = 1, \pm 1, \pm 2, \dots \quad (6.21)$$

We define  $\Lambda$  to be the absolute value of  $M_L$ , the projection of the total electronic orbital angular momentumas onto the internuclear axis, and we use a Greek letter code analogous to that used for atoms. States with  $\Lambda = 0$  are  $\Sigma$  states, those with  $\Lambda = 1$  are  $\Pi$  states, those with  $\Lambda = 2$  are  $\Delta$  states, and those with  $\Lambda = 3$  are  $\Phi$  states, and so on.

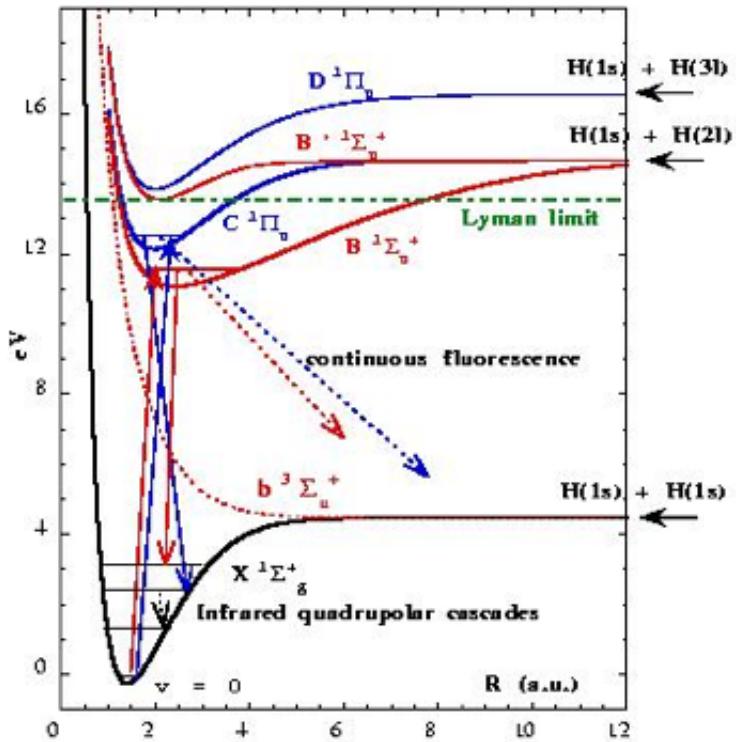


Figure 6.2: The electronic and some vibrational levels are shown for  $\text{H}_2$ .

Again by analogy with atoms, we use capital letters for quantities that apply to the whole molecule (terms) and lower case letters for quantities that apply to the individual electrons ( $\lambda = |m_\ell|$ ). So an electron with  $m_\ell = 0$  is in a  $\sigma$  orbital, one with  $m_\ell = 1$  is in a  $\pi$  orbital, etc. So, we would construct orbitals from atomic orbitals, for example, fill them with electrons, and see what is the ground term.

A second important difference between atoms and molecules is that most molecules are equivalent to closed shell atoms (noble gases), because all electron angular momenta and spins are paired. Thus most molecular ground states are denoted  ${}^1\Sigma$ , where the superscript is  $2S + 1$ . However excited states, or ground states of molecular ions or radicals (molecules with unsatisfied bonds) are often not  ${}^1\Sigma$ . Even some normal, stable molecules have ground states with non-zero electronic angular momentum. Notably  $\text{O}_2$  has a ground state of  ${}^3\Sigma$ , with unpaired electron spins.

There are two other symmetries to be dealt with. First consider a homonuclear diatomic molecule, such as  $\text{H}_2$ ,  $\text{O}_2$ ,  $\text{N}_2$ ,  $\text{C}_2$ . Then the Hamiltonian is invariant under  $\iota$ , the inversion operator that transforms  $\vec{r}_i$  into  $-\vec{r}_i$  (see Chap. 2). Since the operator also commutes with  $L_z$ , we can classify states of a given  $\Lambda$  according to their parity. States that are even under the inversion operator are

called *gerade* (denoted *g*); those that are odd are called *ungerade* (denoted *u*). This is written as a subscript, so we have  ${}^1\Sigma_g$  or  ${}^1\Sigma_u$ , for example. This notation is used only for homonuclear diatomic molecules, since heteronuclear molecules are not invariant under the inversion operator.

Every diatomic molecule does have another symmetry; they are invariant under reflection through a plane containing the internuclear axis. By convention, we take the *xz* plane. Reflection of the electrons' coordinates in this plane corresponds to  $y_i \rightarrow -y_i$ . If we call this operator  $A_y$ , then  $[A_y, H] = 0$ , but  $[A_y, L_z] = -L_z A_y$ , since  $L_z = -i\hbar(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x})$ . Thus if  $\Lambda \neq 0$ ,  $A_y \Phi(+\Lambda) = \Phi(-\Lambda)$ . So every state with  $\Lambda \neq 0$  is doubly degenerate with  $E(\Lambda) = E(-\Lambda)$ . This degeneracy can be split by interaction with rotation ( $\Lambda$  doubling). For  $\Lambda = 0$ , the states are not degenerate, so  $A_y \Phi = \pm 1\Phi$ , and we have two possible eigenvalues. We further label the states according to their behavior under  $A_y$ :  $\Sigma^+$  or  $\Sigma^-$ . The ground state of  $H_2$  is then  ${}^1\Sigma_g^+$ , and the ground state of  $O_2$  is  ${}^3\Sigma_g^-$ . Finally we note that there may be many different states with the same term (e.g.,  ${}^1\Sigma_g^+$ ) corresponding to different orbitals being filled. These are often denoted by a letter to keep them separate (A, B, ...), with X used for the ground state. Thus  $X{}^1\Sigma_g^+$  is the full name for the ground state of  $H_2$ . With this background, the notation in the figure with the energy levels for  $H_2$  should begin to make sense.

A simplified version of the hierarchy of energy levels (electronic, vibrational, rotational) is shown in figure 6.3.

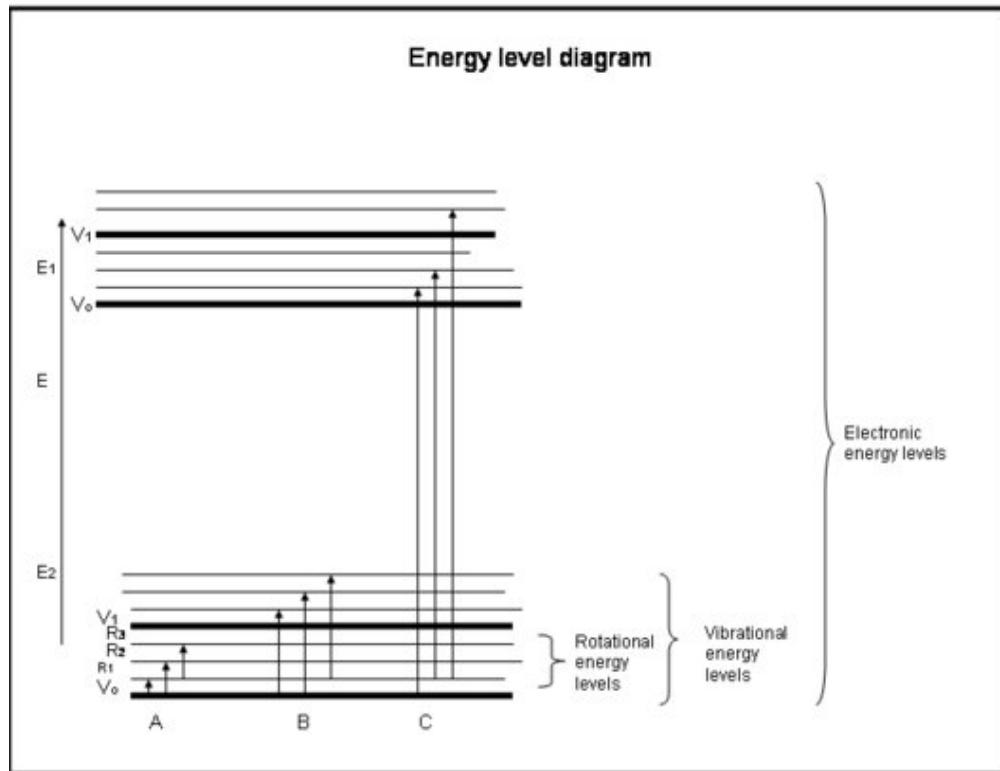


Figure 6.3: The hierarchy of energy levels is indicated schematically.

# Chapter 7

## Diatom Molecules

### 7.1 Introduction

In this chapter, we will consider generalizations of the rigid rotor and harmonic oscillator. For simplicity, we assume that the molecule has no residual *electronic* angular momentum – that is, the electronic state is a  ${}^1\Sigma$  state. We will use a coordinate  $q = R - R_0$  to represent the displacement from the equilibrium separation,  $R_0$ . Then the vibrational wave functions are  $\psi^{vib}(R) = R^{-1}\psi_v(q)$ , where the  $R^{-1}$  arises in converting from the 3-D central force problem to the 1-D problem (see, e.g., **Bransden and Joachain**, pg 237, 2.236). The internuclear potential  $V(q)$  could be the Morse potential or an expansion of the potential about  $q = 0$ , such that

$$V(q) = V(q = 0) + \frac{dV}{dq}|_{q=0}q + 1/2! \frac{d^2V}{dq^2}|_{q=0}q^2 + \dots \quad (7.1)$$

$V(q)$  is the electronic part of the energy so we ignore it here and  $\frac{dV}{dq}|_{q=0} = 0$  since  $q = 0$  is the potential minimum. Consequently, the leading term is the quadratic, leading to the harmonic oscillator wave functions:

$$\psi_v(q) = N_v(e^{-\alpha^2 q^2/2})H_v(\alpha q) \quad (7.2)$$

where  $\alpha^2 = \sqrt{\mu k}/\hbar = \mu\omega_0/\hbar$  and

$$N_v = \left( \frac{\alpha}{2^v v! \sqrt{\pi}} \right)^{0.5} \quad (7.3)$$

With these in mind, we will now consider generalizations of the rigid rotor and the harmonic oscillator approximations.

### 7.2 The Non-Rigid Rotator and the Anharmonic Oscillator

In separating the nuclear wavefunctions into rotational and vibrational functions, we have implicitly assumed that the moment of inertia is constant,  $I = \mu R_0^2$ , corresponding to an infinitely stiff spring constant. We will now relax this assumption slightly. First consider high rotational states, in which

centrifugal forces will try to stretch the bond. By equating the centrifugal force to the restoring force, one can derive a correction term, which makes the rotational energy take the form

$$E_J = BJ(J+1) - DJ^2(J+1)^2 \quad (7.4)$$

with  $D = \frac{4B^3}{\hbar^2\omega^2}$ , where  $\omega = \sqrt{k/\mu}$  is the vibrational frequency. Since we know that vibrational energies are larger than rotational by  $m/M$ , we see that  $D/B$  is like a further term in the expansion in terms of  $\sqrt{m/M}$ . So centrifugal distortion is small, but it grows proportional to  $J^4$  so can become quite important in high  $J$  levels and we can measure frequencies in the radio with extreme precision, so this term is important.

We can also consider higher order terms in the expansion of the potential (eq. 7.1). The cubic and quartic terms are often considered. Classically, these introduce overtones on top of the sinusoidal solutions to the harmonic oscillator. The quantum mechanical perturbations can be expressed as

$$E_v = \hbar\omega_e(v + \frac{1}{2}) - \hbar\omega_ex_e(v + \frac{1}{2})^2 + \hbar\omega_ex_e(v + \frac{1}{2})^3 + \dots \quad (7.5)$$

where  $\omega_e$  is the harmonic value and the  $x_e, y_e, \dots$  can be related to the amplitude of the higher order terms in the potential:

$$V^1(q) = gq^3 - jq^4 \quad (7.6)$$

where the symbols, signs, and coefficients follow the convention of Herzberg's classic book, **Spectra of Diatomic Molecules**, hereafter referred to as **DM**. The cubic term enters only in second order perturbation theory, while the quartic has the symmetry to enter in first order, so they must be considered on equal footing. The result is

$$\omega_ex_e = (2\pi c) \left( \frac{3h^2}{32\pi^4\omega_e^2\mu^2c^2} \right) \left( \frac{5g^2h}{8\pi^2\omega_e^2\mu c} - j \right) \quad (7.7)$$

The  $2\pi c$  factor converts the expression in **DM** to rad s<sup>-1</sup> from cm<sup>-1</sup>. Vibrational spectroscopy often uses the frequency unit of cm<sup>-1</sup>, in which  $\omega_e = \frac{1}{2\pi c}\sqrt{k/\mu}$ .

The correction term,  $\omega_ex_e$ , is invariably positive, so with the explicit minus sign in equation 7.5, the effect of anharmonicity is to *lower* the energy levels from those calculated for the harmonic oscillator. The correction clearly grows with  $v$  so higher levels will be lowered more.

If we now compute the frequencies of the  $\Delta v = \pm 1$  transitions, we find (using cm<sup>-1</sup>),

$$\nu(\text{cm}^{-1}) = \frac{1}{hc} (E_{v+1} - E_v) = \frac{1}{2\pi c} (\omega_e - \omega_ex_e 2(v+1)) \quad (7.8)$$

With this correction, the transitions between adjacent  $v$  are no longer all at the same frequency, so we can measure the different transitions separately. The populations in these levels can provide a probe of temperature in gas with  $T_K \sim 1000$  to 3000 K.

Finally, we consider a small breakdown in the Born-Oppenheimer approximation by considering a coupling between rotation and vibration. In particular, the assumption that  $R = R_0$  cannot be correct if the molecule is vibrating. Since the vibrational motion is much faster than the rotational motion, we can average over the vibrational cycle. Because of the anharmonicity, the molecule spends more time at  $R > R_0$  than at  $R < R_0$ , making the moment of inertia higher and the rotation constant smaller than the rigid rotor values. The effect is larger in higher vibrational states, so the rotation constant now depends on the vibrational state:

$$B_v = \frac{\hbar^2}{2I_0} - \alpha_e(v + \frac{1}{2}) \quad (7.9)$$

Note that there is a correction even for  $v = 0$  because of zero-point vibrations. The difference in  $B$  for different vibrational states will have significant consequences for vibration-rotation transitions.

We can summarize this section by writing the equations for the energies (or frequencies) of the vibrating, rotating molecule, allowing for anharmonicity, centrifugal distortion, and vibration-rotation interaction. Using the notation of **DM**, the “term values” are given in  $\text{cm}^{-1}$  as

$$T = G(v) + F_v(J) \quad (7.10)$$

Then we have

$$T = \omega_e(v + \frac{1}{2}) - \omega_e x_e(v + \frac{1}{2})^2 + \dots + B_v J(J+1) - D_v J^2(J+1)^2 + \dots \quad (7.11)$$

where  $\omega_e x_e$  is given by equation 7.7 *without* the  $2\pi c$  factor. We also have

$$B_v = B_e - \alpha_e(v + \frac{1}{2}) + \dots \quad (7.12)$$

$$D_v = D_e - \beta_e(v + \frac{1}{2}) + \dots \quad (7.13)$$

$$B_e = \frac{\hbar}{4\pi c \mu R_o^2} \quad (7.14)$$

$$D_e = \frac{4B_e^3}{\omega_e^2} \quad (7.15)$$

If we use the cubic and quartic expansion,

$$\alpha_e = \frac{24B_e^3 R_o^3 g}{\omega_e^3} - \frac{6B_e^2}{\omega_e} \quad (7.16)$$

or with a Morse potential, we have

$$\alpha_e = 6 \frac{\sqrt{\omega_e x_e B_e^3}}{\omega_e} - \frac{6B_e^2}{\omega_e} \quad (7.17)$$

and

$$\beta_e = D_e \left( \frac{8\omega_e x_e}{\omega_e} - \frac{5\alpha_e}{B_e} - \frac{\alpha_e^2 \omega_e}{24B_e^3} \right) \quad (7.18)$$

These can also be expressed in other units, with fairly obvious conversions.

A more general, if less intuitive form due to Dunham is also used. One can write the energy levels is

$$E_{vJ} = \sum_{l,j} Y_{lj}(v + \frac{1}{2})^l J^v (J+1)^j \quad (7.19)$$

where the  $Y_{lj}$  are coefficients, not spherical harmonics.

## 7.3 Vibrational and Rotational Transitions

Let's start with the frequencies for purely rotational transitions (no change in  $v$ ). For electric dipole transitions, we have the rule  $\Delta J = \pm 1$ , so let's consider the transition  $J + 1 \rightarrow J$ . Then we have

$$\nu = \frac{E_{J+1,v} - E_{J,v}}{\hbar} = 2B_v(J+1) - 4D_v(J+1)^3 + \dots \quad (7.20)$$

Since these are conventionally given in Hz, MHz, or GHz, you will find values for  $B_v$  and  $D_v$  in those units. Neglecting the centrifugal distortion and higher terms, the transitions will increase in frequency linearly with  $J + 1$ , but those terms will cause them to increase slightly more slowly, and increasingly so as  $J$  increases.

Table 7.1 lists linear molecules that are commonly observed in interstellar space with values for the rotation constant ( $B$ ) and dipole moment magnitude ( $\mu$ ). Some rotational energy diagrams are shown in figure 7.1.

Now consider vibrational transitions. If we ignore for the moment the rotational levels, the frequencies for the transition from  $v'$  to  $v''$  is

$$\nu_0 = G(v') - G(v'') \quad (7.21)$$

where  $\nu_0$ ,  $G$  are in units of  $\text{cm}^{-1}$ . Now if we include the changes in rotational level, we will get

$$\nu = \nu_0 + F_{v'}(J') - F_{v''}(J'') = \nu_0 + B_{v'}J'(J'+1) - B_{v''}J''(J''+1) \quad (7.22)$$

where we have neglected centrifugal distortion for simplicity. For electric dipole selection rules ( $\Delta J = \pm 1$ ), we get two possible transitions. For  $J'' - J' = -1$ , we get the "R branch":

$$\nu_R = \nu_0 + 2B_{v'} + (3B_{v'} - B_{v''})J + (B_{v'} - B_{v''})J^2 ; \quad J = 0, 1, 2, \dots \quad (7.23)$$

For  $J'' - J' = 1$ , we get the "P branch":

$$\nu_P = \nu_0 - (B_{v'} + B_{v''})J + (B_{v'} - B_{v''})J^2 ; \quad J = 1, 2, \dots \quad (7.24)$$

Note that we have replaced  $J''$  by  $J$  in both equations for simplicity. The R branch is displaced to higher frequencies, while the P branch is displaced to lower frequencies. There is no transition at  $\nu_0$  because  $\Delta J = 0$  is not allowed by the parity selection rule.

If there were no interaction of vibration and rotation ( $\alpha_e = 0$ ),  $B_{v'} = B_{v''} \equiv B$  and the above equations would reduce to  $\nu_R = \nu_0 + 2B + 2BJ$  and  $\nu_P = \nu_0 - 2BJ$ , and we would have a series of equally spaced lines. Examination of equation 7.12 and the fact that  $\alpha_e > 0$  shows that  $B_{v'} < B_{v''}$  if  $v' > v''$  as is our convention. Then the R branch lines draw closer together as  $J$  increases while the P branch lines spread further apart as  $J$  increases. This can be seen from the equations as the  $J^2$  term works against the  $J$  term (opposite signs) in the R branch, while they have the same sign in the P branch. For large enough  $J$ , the R branch lines can actually reach a maximum frequency and then move back toward lower frequency. This reversal is called a *bandhead*, and bandheads can be important diagnostics in the atmospheres of cool stars, planets, and protostellar disks. Note that equations 7.23 and 7.24 can both be expressed in a single formula,

$$\nu = \nu_0 + (B_{v'} + B_{v''})m + (B_{v'} - B_{v''})m^2 \quad (7.25)$$

where  $m = 1, 2, \dots (J+1)$  for the R branch and  $m = -1, -2, \dots (-J)$  for the P branch.

### Linear Molecule ${}^1\Sigma$ Example Rotational Spectra

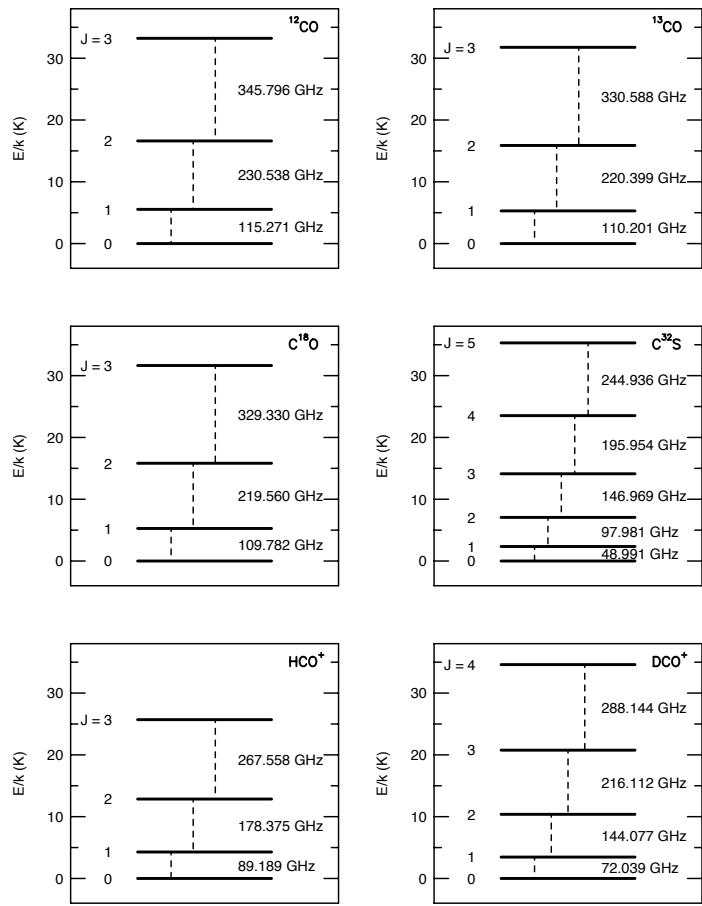


Figure 7.1: The energy level diagram of CO,  ${}^{13}\text{CO}$ ,  $\text{C}^{18}\text{O}$ , CS,  $\text{HCO}^+$ , and  $\text{DCO}^+$  for all energy levels less than 40 K above the ground state. The vertical axis is the energy above ground,  $E/k$ , in K. All energy level diagrams are plotted on the same vertical scale. Notice that heavier molecules, such as CS, have more closely spaced energy levels. Figure supplied by Yancy Shirley.

Table 7.1: Parameters of Linear Molecules in  ${}^1\Sigma$  States

Molecule	Name	B (GHz)	$\mu$ Debye <sup>a</sup>	Molecule	Name	B (GHz)	$\mu$ Debye <sup>a</sup>
CO	Carbon Monoxide	57.635968	0.11011	CS	Carbon Monosulfide	24.495562	1.957
${}^{13}\text{CO}$		55.101011	0.11046	$\text{C}^{34}\text{S}$		24.103541	1.957
${}^{17}\text{O}$		56.179990	0.11034	${}^{13}\text{CS}$		23.123856	1.957
${}^{18}\text{O}$		54.891420	0.11079	HCN	Hydrogen Cyanide	44.315975	2.984
$\text{HCO}^+$	Oxomethylum	44.5944	3.30	$\text{H}^{13}\text{CN}$		43.170137	2.984
$\text{H}^{13}\text{CO}^+$		43.37722	3.3	$\text{HC}^{15}\text{N}$		43.02769	2.984
$\text{HC}^{18}\text{O}^+$		42.58121	3.30	DCN		36.20746	2.984
$\text{DCO}^+$		36.01976	3.3	HNC	Hydrogen Isocyanide	45.33199	3.05
$\text{HOC}^+$	Hydroxymethylidinium	44.7349	4.0	$\text{HN}^{13}\text{C}$		43.54561	2.699
$\text{N}_2\text{H}^+$	Diazenylium	46.586867	3.40	$\text{H}^{15}\text{NC}$		43.02769	2.984
$\text{N}_2\text{D}^+$		38.554719	3.40	DNC		38.152998	3.050
SiO	Silicon Monoxide	21.711967	3.098	$\text{HC}_3\text{N}$	Cyanoacetylene	4.549058	3.724

a. 1 Debye =  $10^{-18}$  esu (cgs units).

All data from JPL line catalog

Table supplied by Yancy Shirley

If we include centrifugal distortion, we can improve equation 7.25 to

$$\nu = \nu_0 + (B_{v'} + B_{v''})m + (B_{v'} - B_{v''} - D_{v'} + D_{v''})m^2 \quad (7.26)$$

$$-2(D_{v'} + D_{v''})m^3 - (D_{v'} - D_{v''})m^4 \quad (7.27)$$

Generally, the last term is very small.

Now that we have established the basic structure of the rovibrational transition, let us inquire into transition strengths and selection rules. To do this, we must consider the electric dipole moment as a function of  $R$ . In almost all cases,  $\mu \rightarrow 0$  as  $R \rightarrow \infty$  (an exception occurs when the molecule dissociates into ions) and  $\mu \rightarrow 0$  as  $R \rightarrow 0$ . Thus the dipole moment must have a maximum somewhere unless it is zero everywhere. There is no particular reason why  $\mu$  should be a maximum at  $R = R_0$ .

Dipole moments are usually given in Debye, which has the value of  $1 \times 10^{-18}$  esu-cm, or statcoulomb-cm, defined by two charges of opposite charge with absolute value  $10^{-10}$  statcoulomb separated by 1 Å, or  $10^{-8}$  cm. In SI units, 1 Debye =  $3.33564 \times 10^{-30}$  Coulomb-m. Typical dipole moments for molecules are around 1 Debye, but CO has a low value around 0.1 Debye. Typical values for vibrational transitions are around 0.1 Debye. The laboratory data for vibrational spectroscopy are given in a wide range of terminologies. A description of how to relate some of these to dipole moments is given in the appendices of a paper by Evans et al. *Ap. J., 383, 674*.

The vibrational line strengths depend on the absolute value squared of the derivative of the dipole moment ( $|du/dq|^2$ ). Thus, it does not matter whether the derivative is positive or negative. Also, we may need to consider higher order terms in the expansion of  $\mu$ :

$$\mu(q) = \mu_0 + \frac{d\mu}{dq}|_0 q + \frac{1}{2!} \frac{d^2\mu}{dq^2}|_0 q^2 + \dots = \mu_0 + \mu_1 q + \mu_2 q^2 + \dots \quad (7.28)$$

This “electrical anharmonicity” will cause  $\Delta v \neq \pm 1$  transitions even if harmonic oscillator wave functions are used. If we include the anharmonicity in the wave functions, one can compute the strengths of the overtone bands,  $\Delta v = \pm 2, \pm 3, \dots$ . In general, we compute the dipole moment

matrix element for a transition between states  $i$  ( $v, J$ ) and  $j$  ( $v', J'$ ) from

$$|\mu_{ij}|^2 = \sum_{m'} [|\langle v', J' | \mu_x | v'', J'' \rangle|^2 + |\langle v', J' | \mu_y | v'', J'' \rangle|^2 + |\langle v', J' | \mu_z | v'', J'' \rangle|^2] \quad (7.29)$$

where the wave functions for the state  $|v, J\rangle$  are those given by equation 6.15. The problem is separable into two integrals in the Born-Oppenheimer Approximation. For example,

$$\begin{aligned} \langle v', J' | \mu_z | v'', J'' \rangle &= \\ \int \int \int \psi_{v'}^{vib^*}(R) \psi_{J'M'}^{rot^*}(\theta, \phi) \mu \cos(\theta) \psi_{v''}^{vib}(R) \psi_{J''M''}^{rot}(\theta, \phi) R^2 \sin(\theta) d\theta d\phi dR \\ &= \int \psi_{v'}^{vib^*}(R) \mu \psi_{v''}^{vib}(R) R^2 dR \times \\ \int \int \psi_{J'M'}^{rot^*}(\theta, \phi) \cos \theta \psi_{J''M''}^{rot}(\theta, \phi) \sin(\theta) d\theta d\phi \end{aligned} \quad (7.30)$$

The second integral is the same as that for the rigid rotor, while the first is that for the harmonic (or anharmonic) oscillator, and  $\mu$  is given by equation 7.28. Recall that

$$\psi_v^{vib}(R) = \frac{1}{R} \psi_v(q) \quad (7.31)$$

where the  $1/R$  arises in converting from the 3-D problem to the 1-D problem, and  $\psi_v(q)$  is given by equation 7.2, so that the first integral becomes

$$\int \psi_{v'}(q) \mu \psi_{v''}(q) dq \quad (7.32)$$

Now consider homonuclear molecules, such as  $H_2$ . The symmetry guarantees that the electric dipole moment will be zero for all separations,  $R$ . Thus, neither vibrational nor rotational transitions can occur by electric dipole radiation. Because these molecules are common and important in astronomy, we need to consider electric quadrupole transitions. Before doing so, we note that isotopically substituted species (isotopologues) can have small dipole moments (e.g.,  $\mu \sim 6 \times 10^{-4}$  Debye for HD).

Electrical quadrupole radiation is of course much weaker than electric dipole radiation, so the only case of astronomical importance is  $H_2$ , owing to its very high abundance in molecular clouds. It is also an important coolant in the early Universe formation of first stars. The general selection rules for n-order poles (Chap. 2) tell us that the allowed transitions are  $\Delta J = 0, \pm 2$ . The rovibrational spectra thus have three series of lines, called “O branch” ( $\Delta J = J'' - J' = 2$ ) “Q branch” ( $\Delta J = J'' - J' = 0$ ) “S branch” ( $\Delta J = J'' - J' = -2$ ). So, now in order of increasing frequency, we have O, P, Q, R, and S, making some logic out of the nomenclature.

To specify a particular rovibrational transition, we add the  $J$  value of the *lower* state, and we designate the transition by  $v' - v'' X(J'')$ , where X is one of O, P, Q, R, S. For example, 1-0 S(0) would correspond to a transition from  $v', J' = 1, 2$  to  $v'', J'' = 0, 0$ . This notation is clumsy, but too deeply entrenched to change.

Now let us return to the discussion of transition strengths. We can think of equation 7.29 as applying to the case where both initial and final electronic states are not degenerate. There has

been an implicit integration over the electronic wave functions to obtain  $\mu(q)$  in equation 7.28. It is essentially the expectation value of

$$\vec{\mu} = \sum_{\alpha} Z_{\alpha} e \vec{R}_{\alpha} - \sum_i \vec{r}_i e \quad (7.33)$$

using center of mass coordinates. We would first take the component of  $\vec{\mu}$  along the molecular axis (all other components would vanish by symmetry) and then compute its expectation value in the electronic state that the molecule is in. That is what we really mean by the molecular dipole moment.

If we sum over both final and initial states to get a line strength (as defined by equation 2.80), we get

$$S_{ij} = \sum_{\epsilon', \epsilon''} \sum_{M', M''} \sum_k |\langle v' J' | \mu_k | v'' J'' \rangle|^2 \quad (7.34)$$

where  $k$  takes on the values  $x, y, z$  (laboratory axes) and the sum over  $(\epsilon', \epsilon'')$  is only over degenerate electronic states (e.g., states that are not  $\Sigma$  states are two-fold degenerate). Again, we can factor  $S_{ij}$  into vibrational and rotational parts. These are referred to as the band strength (sometimes denoted  $S_{v''}^{v'}$ ) and the line strength factor, or Hönl-London factor ( $S_{J'', \Omega''}^{J', \Omega'}$ ), where we have included the possibility that residual angular momentum exists with projection on the molecular axis of  $\Omega$ . These are given by

$$S_{v''}^{v'} = \sum_{\epsilon', \epsilon''} \langle v' | \mu_0 | v'' \rangle|^2 = (2 - \delta_{0\Lambda})(2S + 1) \langle v' | \mu_0 | v'' \rangle|^2 \quad (7.35)$$

and

$$S_{J'', \Omega''}^{J', \Omega'} = \sum_{M', M''} \sum_k |\langle J' \Omega' M' | \mu_k / \mu_0 | J'' \Omega'' M'' \rangle|^2 \quad (7.36)$$

In these expressions, we have taken the dipole moment magnitude to be  $\mu_0$ , so that  $\mu_k / \mu_0$  is the angular part (e.g.,  $\cos(\theta)$  for the case of the  $\mu_z$ ). Except for the new wrinkle of allowing the residual angular momentum ( $\Omega$ ) to enter the Hönl-London factor, we have just repeated the development above, but in terms of line strengths. The degeneracy factors in  $S_{v''}^{v'}$  will disappear if the degeneracy is lifted by interaction of the spin and the rotation of the nuclei. Finally, we note that more accurate calculations of the wave functions and dipole moments are possible with detailed calculations.

## 7.4 Spin and Statistics

When studying multielectron atoms, we found that having identical particles introduced some strong constraints on the allowed wavefunctions. With molecules, we can also have identical nuclei, and these can have similar effects. When identical nuclei can be exchanged by a symmetry operation, some interesting consequences follow. For diatomic, homonuclear molecules, several symmetry operations will exchange the nuclei.

For example, we can rotate the molecule about an axis perpendicular to the symmetry axis. Take the molecular axis to be the  $z$  axis. Then the  $C_2$  operator (which produces a rotation by  $\pi$  radians about  $y$  axis) will exchange the two nuclei. More generally, a  $C_n$  rotation operator rotates

the system by  $2\pi/n$  about some axis. The inversion operator ( $\iota$ ) discussed in Chapter 2 will also exchange the two nuclei, as will the particle exchange operator,  $P_{12}$ .

Let us begin with the simple and astronomically important molecule,  $H_2$ , and assume that we are in the ground electronic state,  ${}^1\Sigma_g^+$ . Since the nuclei are protons, thus fermions, the total wave function must change sign under any symmetry operation that exchanges the two nuclei. so both  $C_2$  and  $\iota$ . Recalling equation 6.16, we can see that the translation wave function is unaffected by any of  $C_2$ ,  $\iota$ , or  $P_{12}$  because they occur in the center of mass frame. They also do not affect the *magnitude* of the relative coordinate, so the vibrational wavefunctions are unaffected. In the ground electronic state, the electronic wave function is symmetric (gerade, and singlet). So, the symmetry properties depend on the rotational wave functions and the (so far neglected) *nuclear* spin functions. We will deal first with the rotational wave functions. Both a  $C_2$  and an  $\iota$  cause the following changes of the angular coordinates:  $\theta \rightarrow \pi - \theta$ ;  $\phi \rightarrow \phi + \pi$  (see equation 2.61). We can see the effect on the rotational wave functions from equation 2.62, if we replace  $l$  by  $J$  and  $m$  by  $M$ :

$$C_2 Y_{JM}(\theta, \varphi) = \iota Y_{JM}(\theta, \varphi) = (-1)^J Y_{JM}(\theta, \varphi) \quad (7.37)$$

Consequently, only odd  $J$  states change sign and satisfy the Pauli Exclusion Principle. At this point, we might worry that the even  $J$  states could not exist, but we now need to consider the nuclear spin functions. Just as with two electrons, we can make three symmetric spin functions,  $\chi_{1M_S}$  and one antisymmetric spin function,  $\chi_{00}$ . So we associate the triplet nuclear spin state with the odd  $J$  rotational levels and the singlet with the even  $J$  levels.

The result is similar to what happens with two electrons in helium: we have two species of  $H_2$ : ortho- $H_2$  (odd  $J$ ) is the triplet and has a nuclear spin degeneracy of  $g_N = 3$ ; para- $H_2$  (even  $J$ ) is the singlet and has  $g_N = 1$ . These really can be considered as separate species for most purposes. Since  $H_2$  is homonuclear, the constant term in the dipole moment vanishes and rotational transitions must be by quadrupole transitions, which can only change  $J$  by 2 or 0, so cannot convert ortho to para. Similarly, collisions would have to flip a nuclear spin, which has a very low collision rate. The only way to convert the two species is chemical reactions. For example, we could break apart an para- $H_2$  and reform a ortho- $H_2$  by reaction with a different H atom. The weakness of the nuclear magnetic moment makes the rules against flipping spins much stronger than the equivalent prohibitions in atoms.

Now, what if the nuclei are bosons? Then the total wave function must *not* change sign under interchange of the nuclei. Then the even  $J$  levels will have higher degeneracy. A very peculiar result occurs for a nucleus with spin of zero, like  ${}^{16}\text{O}$ . With  $I = 0$ , the singlet state is identically zero. Thus, for  $\text{CO}_2$ , a linear molecule with oxygens at each end, the odd  $J$  states really do not exist! I used  $\text{CO}_2$  for this example because  $\text{O}_2$  has a  ${}^3\Sigma_g^-$  electronic state, so a more complicated situation.

We can generalize this discussion to two identical nuclei with any nuclear spin,  $I$ . Assuming that the rest of the wavefunction is symmetric, we consider only the rotational and nuclear spin functions. We can make a total of  $(2I+1)^2$  combinations. Of these,  $(2I+1)$  are symmetric states with aligned spins. This leaves  $(2I+1)^2 - (2I+1) = (2I+1)(2I)$  linear combinations of anti-aligned spins. Half of these will be symmetric and half will be anti-symmetric. So, we have  $(2I+1) + \frac{1}{2}(2I+1)(2I) = (2I+1)(I+1)$  symmetric states and  $\frac{1}{2}(2I+1)(2I) = (2I+1)I$  anti-symmetric states. The ratio of symmetric to anti-symmetric states is then

$$\frac{I+1}{I} \quad (7.38)$$

Note that this ratio goes to 3 for  $I = \frac{1}{2}$  and to infinity for  $I = 0$ , as we showed above.

## Chapter 8

# Polyatomic Molecules: Rotation

### 8.1 Introduction

The range of nuclear motions possible for a polyatomic molecule can be much larger than for a diatomic molecule. It is easier to deal first with rotation. The next step beyond the diatomic molecule is a linear molecule with more than two atoms.

### 8.2 Linear Molecules

In principle, *rigid* linear molecules add no new features to the rotational problem. The energies and eigenfunctions are unchanged, with only a generalization of the moment of inertia:

$$I = \frac{\frac{1}{2} \sum_j \sum_i M_i M_j R_{ij}^2}{\sum_i M_i} \quad (8.1)$$

Once we relax the assumption of rigidity, things rapidly become more complex. A linear triatomic molecule, for example, has three vibrational modes, one of which is a degenerate bending mode. The correction for interaction with vibration must be generalized to

$$B_v = B_e - \sum_i \alpha_i (v_i + \frac{1}{2}), \quad (8.2)$$

where the sum is over the vibrational modes. More subtle effects occur because of the bending modes, which essentially make the molecule non-linear. We defer discussion of these effects.

### 8.3 The Symmetric Top

With the symmetric top, we move from linear (1-dimensional) molecules to molecules that have nuclear structures in two or three dimensions. For linear molecules, we did not allow rotation *about* the nuclear axis because there was no moment of inertia about that axis. Now we deal with systems that can rotate about any axis. For an arbitrary system of masses, the rotational properties are best described by the moments of inertia about three perpendicular axes (variously labeled  $a, b, c$ )

or  $X, Y, Z$  in what follows). We can then define the ellipsoid of inertia, whose semi-major axes are the three principal moments of inertia,  $I_A$ ,  $I_B$ , and  $I_C$ , in increasing order of size. If  $I_B = I_A$ , or  $I_B = I_C$ , we have a symmetric top. A linear molecule is a special case of the symmetric top with  $I_A = 0$  and  $I_B = I_C$ . If all three moments of inertia are equal, we have the special case of a spherical top (e.g.,  $\text{CH}_4$ ). If all the moments are different, we have the general case of the asymmetric top.

A molecule is said to have an  $n$ -fold axis of symmetry if a rotation about that axis by  $2\pi/n$  leaves the distribution of nuclei in space unchanged. In the language of group theory, this is called a  $C_n$  symmetry element. If a molecule has a  $C_n$  symmetry element with  $n \geq 3$ , it is a symmetric top. Linear molecules have a  $C_\infty$  axis – the molecular axis. In practice, we will be concerned with  $C_3$  symmetries; these arise in species with nitrogen and carbon that have the forms,  $\text{NX}_3$  or  $\text{YCX}_3$ . Examples of astronomically important symmetric tops are ammonia ( $\text{NH}_3$ ) and methyl cyanide ( $\text{CH}_3\text{CN}$ ).

Molecules of the form  $\text{NX}_3$  are pyramidal. Rotation by 120 degrees around the  $a$  axis (associated with  $I_A$ ) leaves the molecule unchanged. Also, reflection through any of three vertical planes that bisect the sides of the pyramid between the H atoms leaves the molecule unchanged. Thus  $\text{NH}_3$  belongs to the point group  $C_{3v}$ . Because carbon makes tetrahedral bonds, a methyl group ( $\text{CH}_3$ ) attached to a linear chain also makes a pyramid and the molecule has  $C_{3v}$  symmetry.

Consider first the rotational energy for the general case of an asymmetric top. The classical expression for the rotational energy is

$$E = \frac{1}{2}I_A\omega_A^2 + \frac{1}{2}I_B\omega_B^2 + \frac{1}{2}I_C\omega_C^2 \quad (8.3)$$

which we can rewrite as

$$E = \frac{p_A^2}{2I_A} + \frac{p_B^2}{2I_B} + \frac{p_C^2}{2I_C} \quad (8.4)$$

where  $p_i = I_i\omega_i$  is the angular momentum.

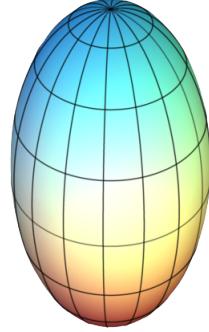


Figure 8.1: A prolate symmetric top.

For a prolate (football shaped) symmetric top,  $I_B = I_C$ , so we have

$$E = \frac{p_A^2}{2I_A} + \frac{p_B^2}{2I_B} + \frac{p_C^2}{2I_B} = \frac{p^2}{2I_B} + p_A^2 \left( \frac{1}{2I_A} - \frac{1}{2I_B} \right), \quad (8.5)$$

where  $p^2 = p_A^2 + p_B^2 + p_C^2$  is the total angular momentum. Now, we can make a transition to quantum mechanics by assuming that  $p^2 = J(J+1)\hbar^2$  and  $p_A^2 = K^2\hbar^2$ , where  $p_A$  is the projection of the

total  $p$  on the *molecular* symmetry axis. Then, as with the projection ( $M$ ) on the *laboratory*  $z$  axis, we have  $K = 0, \pm 1, \pm 2, \dots \pm J$ . With these prescriptions, we obtain

$$E_{JK} = BJ(J+1) + K^2(A - B) \quad (8.6)$$

where  $B = \hbar^2/2I_B$  and  $A = \hbar^2/2I_A$ . As with diatomic molecules,  $A$  and  $B$  can be redefined to have units of  $\text{cm}^{-1}$ , Hz, etc. Because  $I_A < I_B$ ,  $A > B$ , and the second term is positive. Consequently, each  $J$  value corresponds to a series of  $(2J+1)$  levels lying progressively higher in energy (Fig 8.3).

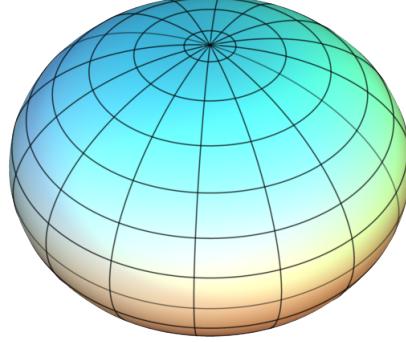


Figure 8.2: An oblate symmetric top.

The analysis for an oblate (frisbee) top ( $I_A = I_B < I_C$ ) would lead to the following energies:

$$E_{JK} = BJ(J+1) + K^2(C - B) \quad (8.7)$$

where  $C = \hbar^2/2I_C < B$ , so the energies decrease with increasing  $K$ , as shown in figure 8.3.

We have assumed without proof that  $K$  is a good quantum number – that the projection on the *molecule-fixed* axis can be a constant of the motion. This was not possible for the linear molecule, which could not rotate about its symmetry axis, but the additional freedom of the symmetric top allows it to do this. Since the energy depends on  $K^2$ , all states with  $K \neq 0$  are doubly degenerate, in addition to the  $2J+1$  degeneracy associated with  $M$ . These facts have some important consequences.

To prove that  $K$  is a good quantum number or to compute matrix elements for transitions, it is necessary to delve into the wave functions for the symmetric top. Again, we take the Born-Oppenheimer approximation and consider the rigid symmetric top. We begin with the problem of how to specify the orientation of the molecular axes with respect to axes fixed in space (the “laboratory” axes). The answer to this problem is the use of the three Euler angles  $(\alpha, \beta, \gamma)$ . In general, we need three rotations, through the three Euler angles, and an intermediate coordinate system, to rotate one coordinate system into another (see Fig. 8.4).

Now think of the  $X, Y, Z$  axes in figure 8.4 as the principal axes of the molecule. For a symmetric top, the  $Z$  axis is the axis of symmetry ( $C_3$ ). Clearly  $\gamma$  will then be associated with rotation about this axis and hence with  $K$ , while  $\alpha$  is associated with rotation about the laboratory  $z$  axis and thus with  $M$ . The second angle ( $\beta$ ) is less intuitive!

Now consider how a function of two angles, like the spherical harmonics, transforms under this rotation. We define the rotation operator  $R$  such that  $R\psi_{jm}$  is the function in the rotated coordinate

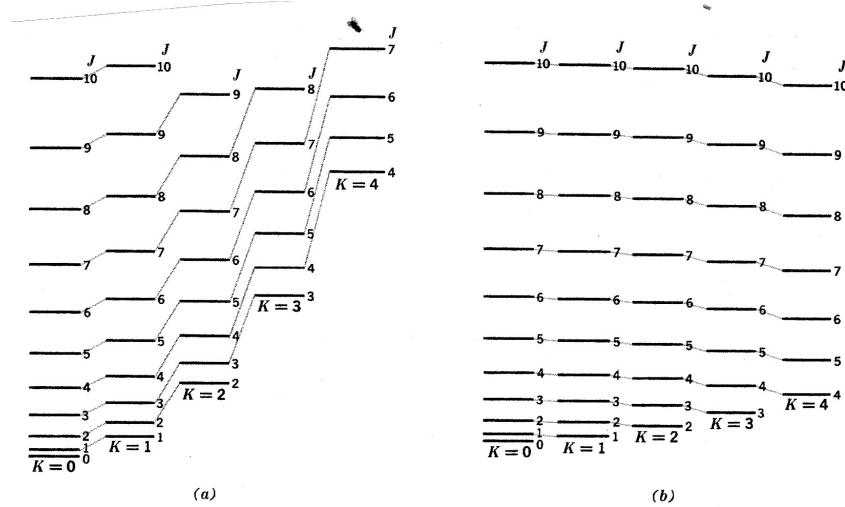


FIG. 5-11 The rotational energy-level patterns for (a) prolate and (b) oblate symmetric-top molecules. (From G. Herzberg, "Infrared and Raman Spectra," D. Van Nostrand Company, Inc., Princeton, N.J., 1945.)

Figure 8.3: The energy levels of prolate (left) and oblate (right) symmetric tops. Taken from Barrow, **Introduction to Molecular Spectroscopy**, 1962

system. For a rotation by  $\theta$  about an arbitrary axis  $\hat{n}$ ,  $R = \exp(-i\theta(\hat{n} \cdot \vec{J}))$ . We can use the Euler angles to write  $R$  as the product of three operators,  $R = R_\gamma R_\beta R_\alpha$  or

$$R = \exp(-i\gamma J_Z) \exp(-i\beta J_N) \exp(-i\alpha J_z) \quad (8.8)$$

with the angles and axes defined as in the figure. This form is awkward because it involves three different coordinate systems. It can be rewritten (see **Elementary Theory of Angular Momentum** by Rose, pg. 51) as

$$R = \exp(-i\alpha J_z) \exp(-i\beta J_y) \exp(-i\gamma J_z) \quad (8.9)$$

(This looks like an amazing sleight of hand, but the derivation is fairly simple, if not very illuminating.)

We can write the new function in terms of the old ones,

$$R\psi_{jm} = \sum_{m'} D_{m'm}^j \psi_{jm'} \quad (8.10)$$

where the coefficients are the matrix elements of  $R$ :

$$D_{m'm}^j(\alpha, \beta, \gamma) = \langle jm' | \exp(-i\alpha J_z) \exp(-i\beta J_y) \exp(-i\gamma J_z) | jm \rangle \quad (8.11)$$

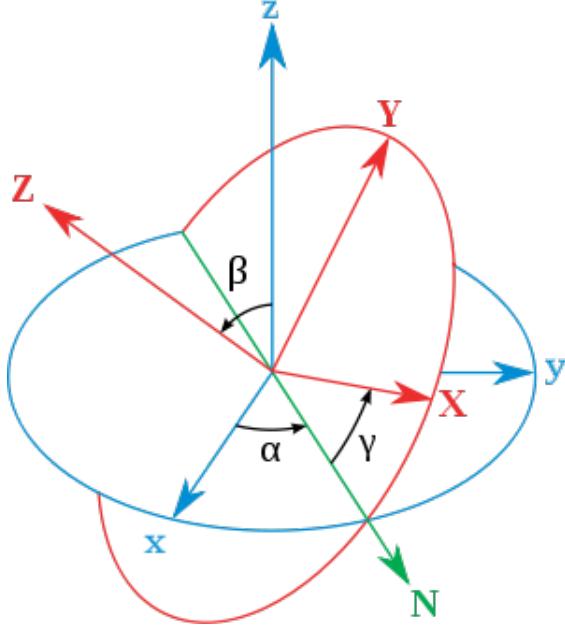


Figure 8.4: The three Euler angles. Only one axis (labeled  $N$ ) of the intermediate coordinate system is shown.

This expression can be simplified because  $|jm\rangle$  is an eigenfunction of  $J_z$ , so

$$\exp(-i\gamma J_z)|jm\rangle = \exp(-i\gamma m)|jm\rangle \quad (8.12)$$

Likewise

$$\langle jm'|\exp(-i\alpha J_z) = \exp(-i\alpha m')\langle jm'| \quad (8.13)$$

Consequently,

$$D_{m'm}^j(\alpha, \beta, \gamma) = \exp(-i\alpha m')d_{m'm}^j(\beta)\exp(-i\gamma m) \quad (8.14)$$

where  $d_{m'm}^j(\beta)$  is the reduced matrix element,

$$d_{m'm}^j(\beta) = \langle jm'|\exp(-i\beta J_y)|jm\rangle \quad (8.15)$$

This is a rather messy function of  $\beta$  involving hypergeometric functions but we never need its explicit form. The point of all this is that the Schrodinger equation for the rigid symmetric top involves the three Euler angles, and its solutions are

$$\psi_{MK}^J = \left( \frac{2J+1}{8\pi^2} \right)^{\frac{1}{2}} D_{MK}^{J*}(\alpha, \beta, \gamma) \quad (8.16)$$

which can also be written as

$$\psi_{MK}^J = \left( \frac{2J+1}{8\pi^2} \right)^{\frac{1}{2}} (-1)^{M-K} D_{-M-K}^J(\alpha, \beta, \gamma) \quad (8.17)$$

The normalization assumes that the integration is carried out over the “solid angle”,  $d\Omega = d\gamma \sin \beta d\beta d\alpha$ , and that

$$\int d\Omega \psi_{MK}^{J^*} \psi_{M'K'}^{J'} = \delta_{JJ'} \delta_{MM'} \delta_{KK'} \quad (8.18)$$

While this may seem like black magic, some motivation may be found by recalling that the spherical harmonics, which are functions of  $\theta$  and  $\phi$ , tell us how a vector is oriented relative to a coordinate system, and these are the wave functions for the 1-dimensional rigid rotor. By analogy, the  $D$  matrices are functions of the Euler angles, which specify the orientation of a whole coordinate system relative to another. The extra generality is needed because of the off-axis atoms in the non-linear molecule. In fact, the spherical harmonics are special cases of the  $D$  matrices:

$$Y_{JM}(\alpha, \beta) = \left( \frac{2J+1}{4\pi} \right)^{\frac{1}{2}} D_{M0}^{J^*}(\alpha, \beta, \gamma) = \left( \frac{2J+1}{4\pi} \right)^{\frac{1}{2}} (-1)^M D_{-M0}^J(\alpha, \beta, \gamma) \quad (8.19)$$

Referring to the figure with the Euler angles, one can see that our familiar  $\theta$  and  $\phi$  coordinates can be identified with  $\beta$  and  $\alpha$ . The new wrinkle is the last rotation about the molecular symmetry axis by  $\gamma$ .

The  $D$  matrices are not only orthonormal (eq. 8-18), but also a complete basis set, so they can be used to construct functions for more complex (asymmetric) molecules. We can also use them to compute the dipole moment matrix elements, establish selection rules, and show that, unlike linear molecules, symmetric tops can have non-vanishing expectation values for the dipole moment. To do all these things, we will use equation 8.19 to express the components of the dipole moment operator in terms of the  $D$  matrices, and then we will use the fact that the integral of three  $D$  matrices has a closed form solution.

$$\begin{aligned} \int d\Omega D_{\mu_3 m_3}^{j_3^*} D_{\mu_2 m_2}^{j_2} D_{\mu_1 m_1}^{j_1} &= (-1)^{-\mu_3 - m_3} 8\pi^2 \delta_{\mu_1 + \mu_2, \mu_3} \delta_{m_1 + m_2, m_3} \\ &\quad \begin{pmatrix} j_1 & j_2 & j_3 \\ \mu_1 & \mu_2 & -\mu_3 \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j_3 \\ m_1 & m_2 & -m_3 \end{pmatrix} \end{aligned} \quad (8.20)$$

The solution is left as an exercise for the reader, and we give only the results for the selection rules and matrix elements:

$$\begin{aligned} |\mu_{J \rightarrow J+1}|^2 &= \mu^2 \frac{(J+1)^2 - K^2}{(J+1)(2J+1)} \quad J+1 \leftarrow J, K \leftarrow K \\ |\mu_{J \rightarrow J-1}|^2 &= \mu^2 \frac{J^2 - K^2}{J(2J+1)} \quad J \rightarrow J-1, K \rightarrow K \end{aligned} \quad (8.21)$$

and the expectation value is

$$|\mu_{JJ}|^2 = \mu^2 \frac{K^2}{J(J+1)} \quad J \rightarrow J, K \rightarrow K \quad (8.22)$$

The notation is conventional in rotational spectroscopy; one writes the upper state before the lower state and uses the direction of the arrow to indicate whether the transition is absorption or emission. The expectation value in equation 8.22 can be non-zero because of the degeneracy associated with

$\pm K$ . The degenerate energy state now does not have definite parity because changing the sign of  $K$  changes the parity. One can understand the non-zero value classically because the molecule can now spin around its own symmetry axis and the dipole moment need not average to zero. With the vector model, one can see the expression 8.22 as the square of the projection of  $J(J+1)$  on the molecular  $Z$  axis. Equation 8.22 is also relevant in rovibrational transitions in allowing a Q branch in which  $J$  is unchanged. Finally, it provides a good estimate for the matrix elements of transitions between  $-K$  and  $+K$  in molecules where the degeneracy is lifted (e.g., NH<sub>3</sub> via its inversion spectrum, or slightly asymmetric tops).

The nomenclature for states is usually  $(J, K)$  [e.g., (1,1)] or  $J_K$  (e.g., 1<sub>1</sub>). The corresponding notation for transitions could be  $(J, K) \rightarrow (J', K')$  or  $J_K \rightarrow J'_{K'}$ .

Note that none of the allowed transitions change  $K$ . This can be understood physically by noting that symmetry restricts the dipole moment to lie along the symmetry axis. Since the torque about an axis is  $\vec{\mu} \times \vec{E}$ , no torque can be exerted about the symmetry axis. Because of this rule, each value of  $K$  forms a separate ladder.

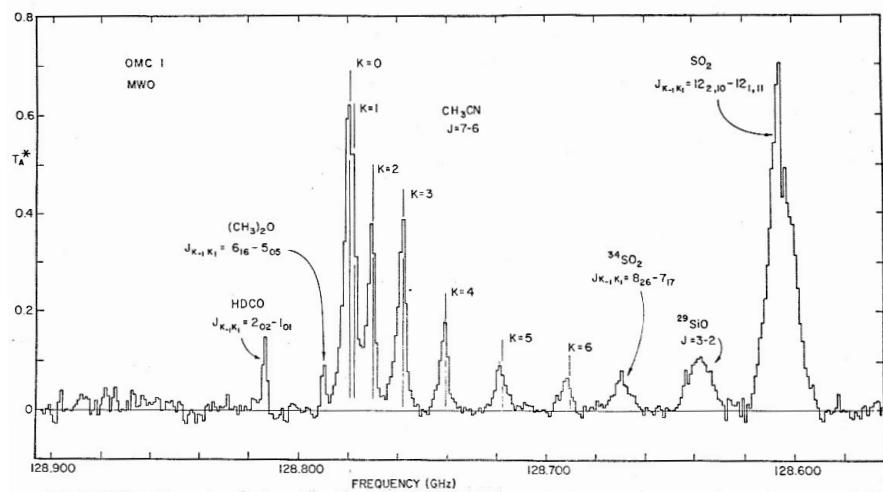


Figure 8.5: A spectrum showing different  $K$  levels for the  $J = 7 - 6$  transition of CH<sub>3</sub>CN.

We can summarize the selection rules for electric dipole transitions in symmetric tops as

$$\begin{aligned} \Delta J &= 0, \pm 1 \quad J = 0 \not\rightarrow 0 \\ \Delta K &= 0, \Delta M = 0, \pm 1 \end{aligned} \tag{8.23}$$

If  $K = 0$ ,  $\Delta J = \pm 1$ . Armed with these selection rules, and equation 8.6 for the energies, the spectrum of rotational transitions has the simple form,

$$\nu = \frac{E_{J+1,K} - E_{J,K}}{h} = 2B(J+1) \tag{8.24}$$

with  $B$  in frequency units. At this level, all transitions within a  $K$  ladder would be at the same

frequency, but centrifugal distortion will split them. Equation 8.6 can be generalized to

$$E_{JK} = BJ(J+1) + (A - B)K^2 - D_J J^2 (J+1)^2 - D_{JK} J(J+1)K^2 - D_K K^4 \quad (8.25)$$

which leads to

$$\nu_{J+1,J} = 2B(J+1)(B - D_{JK}K^2) - 4D_J(J+1)^3 \quad (8.26)$$

The  $D_{JK}$  term thus leads to series of closely spaced lines, as seen in figure 8.5.  $D_J$  is always positive, but  $D_{JK}$  may be either positive or negative. It turns out that  $D_{JK} < 0$  for  $\text{NH}_3$  and  $D_{JK} > 0$  for symmetric tops with a methyl group.

Correction for rotation-vibration interaction is usually complex, since there may be a substantial number of vibrational modes. Some of these may also result in an effective angular momentum about the symmetry axis.

## 8.4 The Asymmetric Top

If we return to the general expression for the rotational energy of a 3-dimensional object with moments of inertia,  $I_A$ ,  $I_B$ , and  $I_C$  about the principal axes ( $a$ ,  $b$ , and  $c$ ), we have the classical Hamiltonian for the rigid asymmetric top (eq. 8.3). We can switch to quantum mechanics by changing the angular momenta to operators:

$$E = \frac{L_A^2}{2I_A} + \frac{L_B^2}{2I_B} + \frac{L_C^2}{2I_C} \quad (8.27)$$

However, the eigenfunctions for the asymmetric top are not eigenfunctions of any of these angular momentum operators. Classically the projection of the angular momentum onto the principal axes is not a constant of the motion. Only the total angular momentum and its projection onto the laboratory  $z$  axis are constants of the motion, so  $J$  and  $M$  will be good quantum numbers, but these are insufficient to specify the state. Instead, we label the states with various useful pseudo quantum numbers.

We begin by defining the rotation constants (here in units of Hz):

$$A = \frac{\hbar}{8\pi^2 I_A}; \quad B = \frac{\hbar}{8\pi^2 I_B}; \quad C = \frac{\hbar}{8\pi^2 I_C}; \quad (8.28)$$

Then Ray's asymmetry parameter is

$$\kappa = \frac{2B - A - C}{A - C} \quad (8.29)$$

which goes to  $-1$  for prolate molecules ( $B = C$ ) and to  $+1$  for oblate molecules ( $A = B$ ). If the molecule is nearly symmetric, the following two parameters are useful:

$$b_p = \frac{C - B}{2A - B - C} = \frac{\kappa + 1}{\kappa - 3} \rightarrow 0 \quad \text{if prolate} \quad (8.30)$$

$$b_o = \frac{A - B}{2C - B - A} = \frac{\kappa - 1}{\kappa + 3} \rightarrow 0 \quad \text{if oblate} \quad (8.31)$$

A common specification of the energy levels arises from considering the limits as  $\kappa \rightarrow -1$  or  $+1$ . The procedure is analogous to the correlation diagrams used to discuss electronic structure.

We consider the molecule to be continuously deformable from a prolate to an oblate symmetric top (Fig. 8.6). The states are then labeled by  $J_{K_{-1}K_{+1}}$ , where  $K_{-1}$  is the  $K$  value the molecule would have if prolate ( $\kappa = -1$ ) and  $K_{+1}$  is the value it would have if oblate ( $\kappa = +1$ ). An alternative designation is  $J_\tau$ , where  $\tau = K_{-1} - K_{+1}$  runs from  $-J$  to  $+J$  and labels the various states in order of increasing energy.

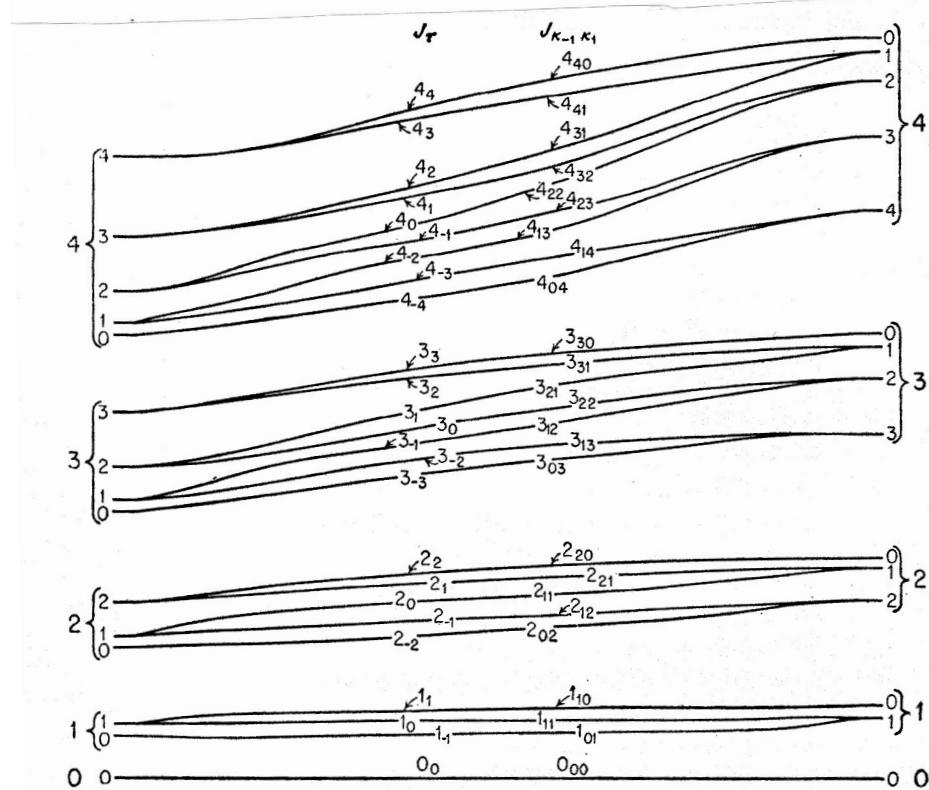


FIG. 4-1. Qualitative behavior of the asymmetric-top energy levels. The rotational constant  $B$  varies from left to right, equaling  $C$  and giving a prolate symmetric top on the left, and equaling  $A$  to give an oblate symmetric top on the right.

Figure 8.6: A correlation diagram for asymmetric top energy levels. Prolate limit is on the left, oblate limit on the right. Taken from Townes and Schalow.

For a nearly prolate top, the energies may be written in the form

$$\frac{E}{h} = \frac{B+C}{2} J(J+1) + \left( A - \frac{B+C}{2} \right) w \quad (8.32)$$

If  $b_p \rightarrow 0$ ,  $B \rightarrow C$  and  $w \rightarrow K^2$ . Consequently, it is reasonable to expand  $w$  in powers of  $b_p$ :

$$w = K^2 + c_1 b_p + c_2 b_p^2 + \dots \quad (8.33)$$

Exact expressions for  $w$ , regardless of the value of  $\kappa$  are given for low  $J$  values in **Microwave Spectroscopy** by Townes and Schawlow (pg. 85). Analogous expressions can be written for nearly oblate molecules.

Another way to look at nearly symmetric tops is to think of the  $\pm K$  degeneracy of the symmetric top being split by the asymmetry. For  $K = 1$ , the splitting is proportional to  $b$  and to  $J(J+1)$ , producing a series of doublets with increasing splitting as  $J$  increases. This phenomenon is referred to as “K-type doubling” and the levels as “K doublets.” For higher  $K$ , the splitting is less. For small  $b$ , a general formula is

$$\Delta w = \frac{b^K (J+K)!}{8^{K-1} (J-K)! [(K-1)!]^2} \quad (8.34)$$

where  $b, K$  are given by the appropriate values depending on whether the molecule is nearly prolate or nearly oblate.

The wave functions are of course rather complex, but one generally uses the symmetric top wavefunctions as the basis set:

$$\psi_{K-1, K+1, M}^J = \sum_{J'} \sum_K \sum_{M'} a_{J', K, M'} \psi_{KM'}^{J'} \quad (8.35)$$

Because  $J$  and  $M$  are still good quantum numbers, only  $J' = J$  and  $M' = M$  will enter, so that

$$\psi_{K-1, K+1, M}^J = \sum_K a_{J, K, M} \psi_{KM}^J \quad (8.36)$$

This can be further simplified because the ellipsoid of inertia has a  $C_2$  symmetry about each principal axis. Consequently (skipping the detailed argument), only odd or even  $K$  will enter the sum. A more useful form for the basis functions is then symmetrized versions:

$$\psi_{K\nu M}^J = \frac{1}{\sqrt{2}} (\psi_{KM}^J + (-1)^\nu \psi_{-KM}^J) \quad (8.37)$$

where  $\nu = 0, 1$  for  $K > 0$ , and

$$\psi_{K\nu M}^J = \psi_{0M}^J \quad (8.38)$$

for  $K = 0$ . Then one allows only  $K > 0$  in the sum

$$\psi_{K-1, K+1, M}^J = \sum_{K, \nu} a_{J, K, M} \psi_{K\nu M}^J \quad (8.39)$$

As one might expect, for nearly prolate molecules, the dominant terms will be  $K = \pm K_{-1}$ .

Symmetries may further reduce the problem, and they play a major role in determining the type of allowed transitions and the spectrum. The systematic way to deal with these symmetries is to use group theory, but we will restrict our discussion to a few simple points. The major factor in which transitions are allowed is which principal axes have non-vanishing components of the dipole

moment. In general, the dipole moment of an asymmetric molecule need not lie along any principal axis. So we have in general, for light polarized along the *laboratory z* axis:

$$\langle j | \mu_z | i \rangle = \mu_a \langle j | \cos az | i \rangle + \mu_b \langle j | \cos bz | i \rangle + \mu_c \langle j | \cos cz | i \rangle \quad (8.40)$$

where  $\cos(az)$  is the cosine of the angle between the *a* and *z* axes. Because the selection rules will differ for each component, it is more useful to consider separately the matrix elements for transitions associated with each component of  $\mu$ . Many of the simpler asymmetric tops have a dipole moment that does lie along one of the principal axes. A convenient summary of the allowed types of transitions, taken from Townes and Schawlow, pg. 94 follows.

Component	$\Delta K_\kappa$	$K_{-1}K_{+1}$
$\mu_a$	$\Delta K_{-1} = 0, 2, \dots$	ee - eo
$\mu_a$	$\Delta K_{+1} = 1, 3, \dots$	eo - oe
$\mu_b$	$\Delta K_{-1} = 1, 3, \dots$	ee - oo
$\mu_b$	$\Delta K_{+1} = 1, 3, \dots$	eo - oe
$\mu_c$	$\Delta K_{-1} = 1, 3, \dots$	ee - oe
$\mu_c$	$\Delta K_{+1} = 0, 2, \dots$	oo - eo

and of course,  $\Delta J = 0, \pm 1$  for all of these.

For nearly symmetric tops, the dipole moment matrix elements are well approximated by those of the corresponding symmetric top, with  $K_{-1} = K$  if nearly prolate, etc. A convenient table for constructing these matrix elements for  $\mu$  along any axis is given by Townes and Schawlow, pg. 96. For more asymmetric cases, one can use the numerical factors in Appendix V of Townes and Schawlow, using  $\kappa$  to interpolate. Note that the strength, rather than the dipole moment matrix element, is given and the P, Q, R branch notation is used to specify  $\Delta J$ .

## 8.5 Spin and Statistics

If identical nuclei can be interchanged by a symmetry operation, restrictions exist, which may have some interesting consequences. Identical nuclei are of course natural to symmetric tops. Because of the cosmic abundances of the elements, most astrophysically important molecules with identical nuclei involve hydrogen. The two most common symmetries are  $C_{2v}$  and  $C_{3v}$ . These are point groups (they leave a point fixed in space). The symmetry elements are

$$C_{2v} : E, C_2, 2\sigma_v \quad C_{3v} : E, C_3, 3\sigma_v \quad (8.41)$$

where *E* is the identity,  $C_n$  is rotation by  $2\pi/n$  about an axis, and  $\sigma_v$  is reflection through a plane passing through the *vertical* axis, taken to be the symmetry axis (Fig. 8.7).

These group theoretical approaches allow a classification of eigenstates into different symmetry species, analogous to ortho and para  $H_2$ . For example, molecules with  $C_2$  symmetry (e.g.,  $H_2O$ ,  $H_2CO$ ) divide into ortho and para species (more properly A and B species),  $K_{-1}$  odd is ortho and  $K_{+1}$  even is para in the case of  $H_2CO$ . For  $H_2O$ ,  $K_{-1} + K_{+1}$  even is para, while  $K_{-1} + K_{+1}$  odd is ortho. Transitions between ortho and para states are strongly forbidden, as discussed for diatomic molecules, and the statistical weights differ:  $g_I = 3$  for ortho, while  $g_I = 1$  for para.

For molecules with  $C_{3v}$  symmetry, we have symmetric tops. The eigenstates divide into A and E species as follows.

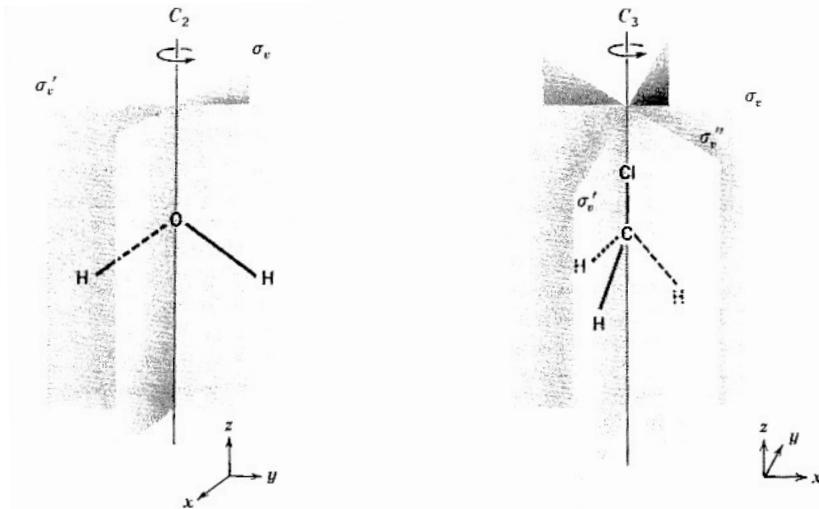


FIG. 8-2 The symmetry elements of  $\text{H}_2\text{O}$ , point group  $\text{C}_{2v}$ , and  $\text{CH}_3\text{Cl}$ , point group  $\text{C}_{3v}$ . (The coordinate systems are drawn in ways that will be convenient for later treatments.)

Figure 8.7: Symmetry operations for  $C_{2v}$  and  $C_{3v}$  species. Taken from Barrow.

Species	Allowed $K$	Notes
A	0, 3, 6, ...	$K = 3n$ , n an integer
E	1, 2, 4, 5, ...	$K \neq 3n$ , n an integer

For three fermions, the statistical weights due to nuclear spin are

Species	$g_I$	if $I = \frac{1}{2}$
A	$\frac{2I+1}{3}(4I^2 + 4I + 3)$	4
E	$\frac{2I+1}{3}(4I^2 + 4I)$	2

The total statistical weight is then

$$g_{JK} = g_J g_K g_I \quad (8.42)$$

where  $g_J = 2J + 1$  and  $g_K = 2$  for  $K \neq 0$  and 1 for  $K = 0$ . Electric dipole transitions do not change  $K$ , but collisions are also affected:  $\Delta K = 3n$  collisions are strongly favored over  $\Delta K \neq 3n$ , so the A and E species are quite separate and are often again called ortho and para, but this is not strictly correct.

The discussion above is adequate for the methyl group symmetric tops, but  $\text{NH}_3$  is more interesting. Think of the  $\text{NH}_3$  molecule as a planar molecule that can vibrate so that the N moves to either side of the plane defined by the three hydrogens. If  $\text{NH}_3$  were truly planar (like  $\text{BF}_3$  for example), we could imagine a roughly harmonic oscillator potential centered on  $z = 0$ . For  $\text{NH}_3$ , a small hump

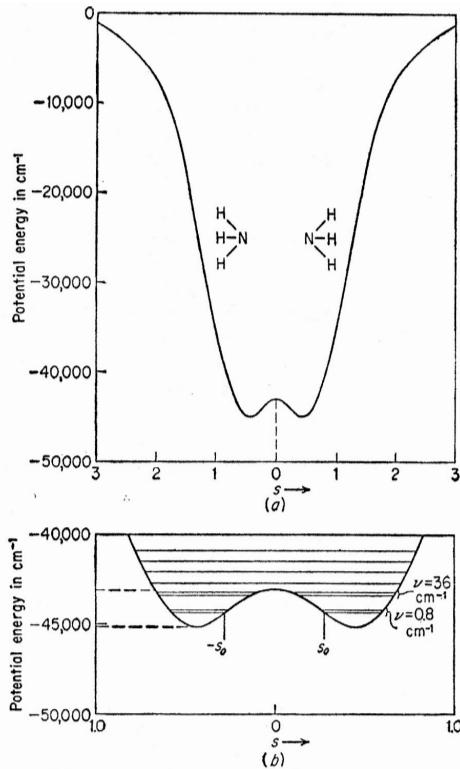


Figure 8.8: The potential for a harmonic oscillator with a quartic perturbation. Taken from Townes and Schawlow.

exists in the potential which repels the N as it tries to go through. This causes the energy levels of the oscillator to come together in pairs, as does the quartic term in the expansion in a power series. The perturbation mixes the first two harmonic oscillator eigenfunctions to produce basis functions in which the N is partly on one side and partly on the other.

$$\begin{aligned}\psi_0 &= \frac{1}{\sqrt{2}} (\psi_{left} + \psi_{right}) \\ \psi_1 &= \frac{1}{\sqrt{2}} (\psi_{left} - \psi_{right})\end{aligned}\tag{8.43}$$

Then the perturbation results in a time dependent function:

$$\Psi = \frac{1}{\sqrt{2}} (\psi_0 + \psi_1 e^{i\Delta Et/\hbar}) e^{iE_0 t/\hbar}\tag{8.44}$$

where  $\Delta E = E_1 - E_0$ . At  $t = 0$ ,  $\Psi = \frac{1}{\sqrt{2}}(\psi_0 + \psi_1)$ , which corresponds to the N being on the left side of the barrier. Then at  $t = \pi\hbar/\Delta E$ ,  $\Psi = \frac{1}{\sqrt{2}}(\psi_0 - \psi_1)$ , which corresponds to the N on the right side. Thus we have an oscillation of the N from side to side, like an umbrella turning itself inside-out.

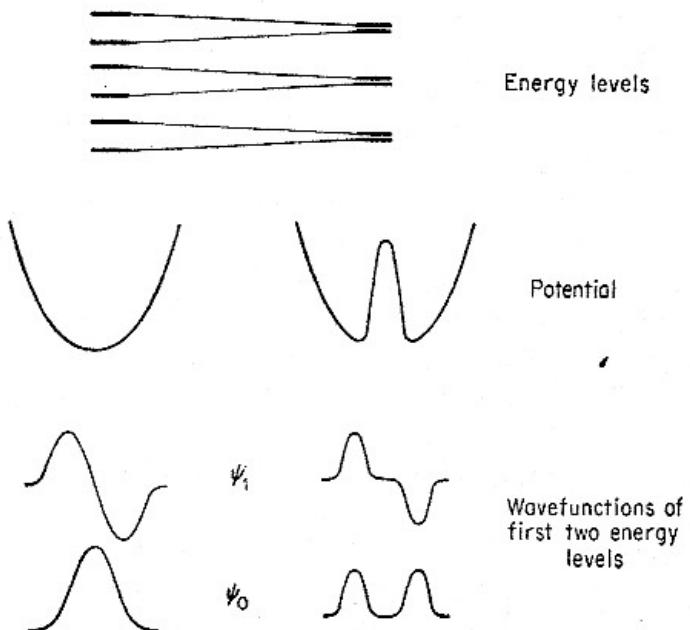


FIG. 3-8. Behavior of a vibration with introduction of a potential barrier.

Figure 8.9: The wavefunctions for a harmonic potential with a hump in the middle. Taken from Townes and Schawlow.

The higher the barrier, the longer the period. For most symmetric tops, the period exceeds the age of the Universe. Then the two levels effectively coincide and any oscillation is about a position on one side or the other. For  $\text{NH}_3$ , the barrier is small enough that the frequency is about 24 GHz, for a wavelength of about 1 cm. This transition, called an inversion transition, was how interstellar  $\text{NH}_3$  was discovered by my graduate school officemate, Al Cheung. It was the first polyatomic molecule to be discovered in space.

When the N changes sides, it effectively changes the direction of the Z axis, and therefore the sign of  $K$ , so we can also think of this as splitting the  $\pm K$  degeneracy. Detailed considerations of the restrictions imposed by the three identical H nuclei leads to the fact that the A species is favored over the E species by a factor of 2, but that for  $K = 0$ , alternate states do not exist.

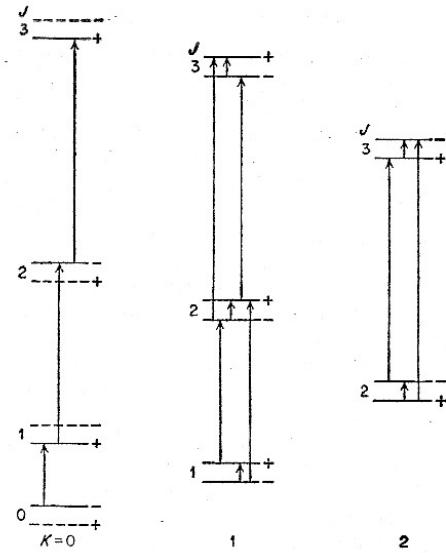


FIG. 3-9. Levels and possible transitions in the rotation-inversion spectrum of  $\text{NH}_3$ .  
The dotted levels for  $K = 0$  are forbidden by the exclusion principle.

Figure 8.10: The energy levels of  $\text{NH}_3$  with the parity of the levels indicated by  $\pm$ . Taken from Townes and Schawlow.

## Chapter 9

# Polyatomic Molecules: Vibration

### 9.1 Degrees of Freedom, Normal Modes

Consider first the general problem of how many vibrational motions are possible. For  $n$  atoms, we would need  $3n$  coordinates to describe all their locations, so the  $n$ -atom system has  $3n$  degrees of freedom. We can perform coordinate transformations analogous to those used in discussing the hydrogen atom to assign three degrees of freedom to the translational motion of the center of mass. Further, 3 can be assigned to the rotation about the three principle axes (only 2 for linear molecules). This leaves  $3n - 6$  vibrational degrees of freedom ( $3n - 5$  for linear molecules). Clearly, molecules with many atoms can have an extremely large number of vibrational modes, though some of them may be degenerate. At the other extreme, a diatomic molecule would have  $3(2) - 5 = 1$  degree of freedom, vibration along the bond, and a triatomic, linear molecule would have  $3(3) - 5 = 4$  two of which would be degenerate bending modes.

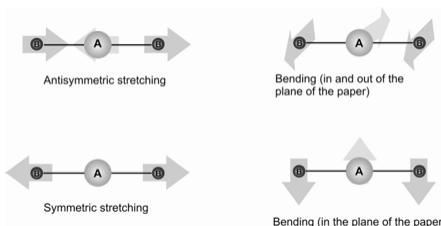


Figure 9.1: Two stretching modes (symmetric and antisymmetric) and two degenerate bending modes.

The modes are shown in figure 9.1. A nice demonstration of these modes can be found online<sup>1</sup>.

The actual separation of rotational and vibrational motions is somewhat complex in polyatomic molecules and not very enlightening. A full description can be found in Chapter 11 of the classic book, *Molecular Vibrations*, by Wilson, Decius, and Cross. Essentially, one moves to the rotating frame of reference described by the Euler angles and then treats the vibration separately by writing

<sup>1</sup>[http://chemwiki.ucdavis.edu/Physical\\_Chemistry/Spectroscopy/Vibrational\\_Spectroscopy/Vibrational\\_Modes](http://chemwiki.ucdavis.edu/Physical_Chemistry/Spectroscopy/Vibrational_Spectroscopy/Vibrational_Modes)

the Hamiltonian in terms of the Cartesian coordinates of each nucleus. In the classical mechanics limit,

$$T = \frac{1}{2} \sum_{\alpha=1}^N m_\alpha \left[ \left( \frac{d\Delta x_\alpha}{dt} \right)^2 + \left( \frac{d\Delta y_\alpha}{dt} \right)^2 + \left( \frac{d\Delta z_\alpha}{dt} \right)^2 \right] \quad (9.1)$$

where  $\Delta x_\alpha$  is the displacement of  $x$  of the  $\alpha$ th nucleus from its equilibrium position. We then define new coordinates,

$$q_1 = \sqrt{m_1} \Delta x_1; q_2 = \sqrt{m_1} \Delta y_1; q_3 = \sqrt{m_1} \Delta z_1; q_4 = \sqrt{m_2} \Delta x_2; \dots \quad (9.2)$$

allowing us to write the kinetic energy in compact form:

$$T = \frac{1}{2} \sum_{i=1}^{3N} \dot{q}_i^2; \quad \dot{q}_i = \frac{dq_i}{dt} \quad (9.3)$$

For small displacements, the potential energy can be written as a power series in  $q_i$ :

$$V = V_0 + \sum_{i=1}^{3N} \left( \frac{\partial V}{\partial q_i} \right)_0 q_i + \frac{1}{2} \sum_{i,j=1}^{3N} \left( \frac{\partial^2 V}{\partial q_i \partial q_j} \right)_0 q_i q_j + \dots \quad (9.4)$$

and the first two terms are zero, as shown for diatomic molecules in §7.1. We then have

$$V = \frac{1}{2} \sum_{i,j=1}^{3N} f_{ij} q_i q_j; \quad f_{ij} = \left( \frac{\partial^2 V}{\partial q_i \partial q_j} \right)_0 \quad (9.5)$$

The equations of motion can then be solved classically for the amplitudes ( $A_i$ ) and frequencies ( $\lambda$ ) of the normal modes ( $q_i = A_i \cos(\lambda^{0.5}t + \epsilon)$ ). Actually, only the relative amplitudes of the normal modes are determined. That is, for the  $k^{th}$  normal mode, we define a relative amplitude of the  $i^{th}$  coordinate by  $l_{ik}$ ; a convenient normalization is

$$l_{ik} = \frac{A_{ik}}{\left[ \sum_i (A_{ik})^2 \right]^{0.5}} \quad (9.6)$$

where  $A_{ik}$  is an arbitrary set of amplitudes with the right ratios. Then, we define the **normal coordinates**,  $Q_k$  by

$$Q_k = \sum_{i=1}^{3N} l_{ki}'' q_i \quad (9.7)$$

where the coefficients are related to the normalized amplitudes ( $l_{ik}$ ) by

$$\sum_{i=1}^{3N} l_{ki}'' l_{il} = \delta_{kl} \quad (9.8)$$

and the  $l_{ik}$  are the coefficients of the inverse transformation to

$$q_i = \sum_{k=1}^{3N} l_{ik} Q_k \quad (9.9)$$

Now we have “coordinates” that refer, not to the position of a single nucleus, but to the a “kind of motion” called a normal mode. The price we have paid in losing coordinates with obvious meaning is regained because now the potential and kinetic energies contain no cross terms.

$$T + V = \sum_{i=1}^{3N} \dot{Q}_k^2 + \frac{1}{2} \sum_{i=1}^{3N} \lambda_k Q_k^2 \quad (9.10)$$

Six (five for linear molecules) of these normal modes will have zero frequency and these correspond to the translational and rotational modes. We can remove those to obtain

$$H = T + V = \sum_{i=1}^{3N-6} \dot{Q}_k^2 + \frac{1}{2} \sum_{i=1}^{3N-6} \lambda_k Q_k^2 \quad (9.11)$$

where we will just use  $3N - 6$  for simplicity. Now we make the transition to quantum mechanics in the usual way.

$$\frac{-\hbar^2}{2} \sum_{k=1}^{3N-6} \frac{\partial^2 \psi_v}{\partial Q_k^2} + \frac{1}{2} \sum_{k=1}^{3N-6} \lambda_k Q_k^2 \psi_v = E_v \psi_v \quad (9.12)$$

where  $E_v$  is the vibrational energy. Now, we really reap the benefits of the obscure normal coordinates because the equation is separable. That is, we postulate a solution:

$$E_v = E(1) + E(2) + \cdots + E(3N - 6) \quad (9.13)$$

$$\psi_v = \psi(Q_1)\psi(Q_2)\cdots\psi(Q_{3N-6}) \quad (9.14)$$

and we find that the Schrodinger equation is satisfied if the following  $3N - 6$  equations are satisfied:

$$\frac{-\hbar^2}{2} \frac{d^2 \psi(Q_k)}{dQ_k^2} + \frac{1}{2} \lambda_k Q_k^2 \psi(Q_k) = E(k) \psi(Q_k); \quad k = 1, 3N - 6 \quad (9.15)$$

Thus, we have a wave function that is a product of harmonic oscillator wave functions, each corresponding to a normal mode, and an energy that is the sum of energies of  $3N - 6$  harmonic oscillators:

$$E_v = (v_1 + \frac{1}{2})h\nu_1 + (v_2 + \frac{1}{2})h\nu_2 + \cdots + (v_{3N-6} + \frac{1}{2})h\nu_{3N-6} \quad (9.16)$$

where  $\nu_1$  is the classical frequency of the vibration (here in Hz).

Notice that the zero point energies,

$$E(0) = \frac{1}{2}h \sum_{k=1}^{3N} \nu_k \quad (9.17)$$

may add to a substantial number ( $4500 \text{ cm}^{-1}$ ) in the case of  $\text{H}_2\text{O}$ . Because each normal mode produce a series of level like those of a diatomic molecule, we can specify the vibrational state by giving the quantum numbers ( $v_k$ ) for each normal mode:

$$(v_1, v_2, v_3, \dots, v_{3N-6}) \quad (9.18)$$

The three normal modes of  $\text{H}_2\text{O}$  and the energy levels are shown in figure 9.2.

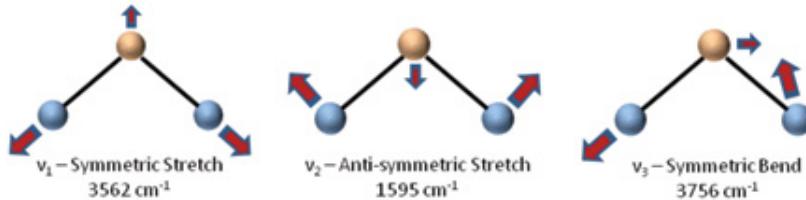


Figure 9.2: Three modes of water molecule with fundamental frequencies.

Now for a bit of nomenclature. Energy levels with all  $v_k = 0$  except one, for which  $v_i = 1$  are called **fundamental** levels. If  $v_i > 1$ , it is an **overtone** level. If two or more quantum numbers are non-zero, it is a **combination** level. Likewise a transition of the sort

$$(001) \rightarrow (000) \quad (9.19)$$

is called a fundamental transition;  $(002) \rightarrow (000)$  would be an overtone, and  $(011) \rightarrow (000)$  is a combination transition. The frequency of the transitions are given by

$$\nu = \frac{E_{upper} - E_{lower}}{h} \quad (9.20)$$

which would be  $\nu_k$  for a fundamental transition in the harmonic oscillator approximation. Of course, one can improve the accuracy by including anharmonic terms as in the case of a diatomic molecule. Also the  $\Delta v_k = \pm 1$  selection rule of the harmonic oscillator can be modified by electrical or mechanical anharmonicity. It should be noted that some normal modes will cause no change in the electric dipole moment; these are called “infrared-inactive” modes.

Detailed knowledge of the normal coordinates is not necessary to effect the separation into normal modes, which gives us the basic structure of the energy levels of individual oscillators. More powerful techniques for dealing with more complex molecules have been developed; matrix methods and group theory are useful tools.

## 9.2 Rovibrational Transitions of Polyatomic Molecules

Because this subject can get very complex, we will content ourselves with a qualitative discussion of the rovibrational spectra of polyatomic molecules.

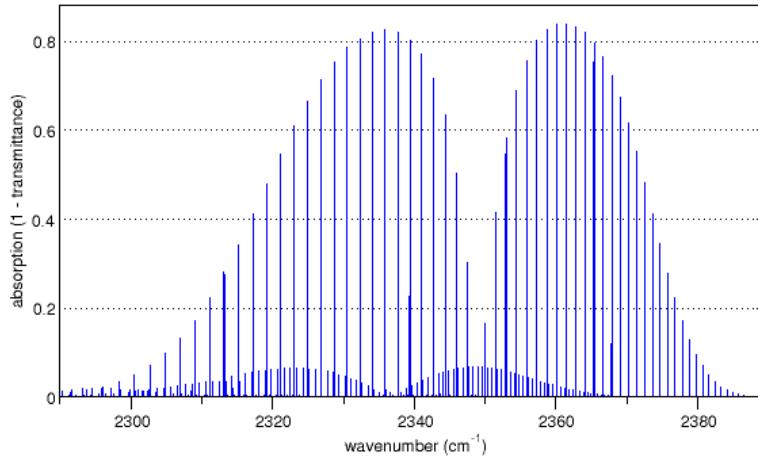


Figure 9.3: The parallel band of the linear molecule  $\text{CO}_2$ , showing P and R branches. The other, weaker bands are caused by other species.

Let us begin with linear molecules, which have two distinct types of transitions. If the vibration produces an oscillating dipole moment parallel to the molecular axis, it is called a **parallel** band. If the oscillating dipole is perpendicular to the axis, it is called a **perpendicular** band.

The parallel band is quite analogous to the bands for diatomic molecules, having both P and R branches, but no Q ( $\Delta J = 0$ ) branch.

A perpendicular band may have a Q branch because the molecule is actually no longer linear in the excited state of a bending mode. Consequently, the molecule loses the symmetry that prevents a  $\Delta J = 0$  transition, just as a symmetric top does (see equation 8.22). Because the spacing of rotational levels is nearly the same in the two vibrational states, the Q-branch lines tend to lie at nearly the same frequency ( $\nu_0$ , the band center), producing a strong absorption. Because  $B_1$  is slightly less than  $B_0$  (see equation 7.12), higher  $J$  lines do decrease slowly in frequency, causing the Q branch to have a low frequency shoulder.

Symmetric top molecules also have both parallel and perpendicular bands, where now parallel refers to the symmetry axis (e.g., the  $C_3$  axis). Then the selection rules are given by

$$\begin{aligned} \Delta J &= 0, \pm 1, \quad \Delta K = 0, \quad \text{if } K \neq 0 \\ \Delta J &= \pm 1, \quad \Delta K = 0, \quad \text{if } K = 0 \end{aligned} \tag{9.21}$$

as reflected by equations 8.21 and 8.22. We can think of each  $K$  ladder's contribution separately. The  $K = 0$  ladder will have only P and R branches. The  $K = 1$  ladder will have a Q branch and P and R branches that begin at P(2) and R(1), since the  $J = 0$  state does not exist in the

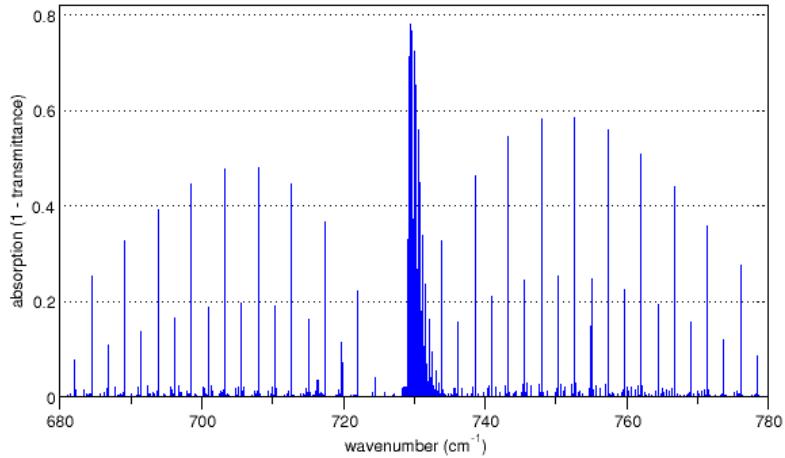


Figure 9.4: A perpendicular band of acetylene, showing a Q branch as well as P and R branches. Note the alternation of strong and weak lines, with a ratio of about 3, caused by the ortho-para effects of the identical hydrogens.

$K = 1$  ladder. Similarly, higher  $K$  ladders will have Q branches and P and R branches that start at  $P(K + 1)$  and  $R(K)$ . The result is shown below.

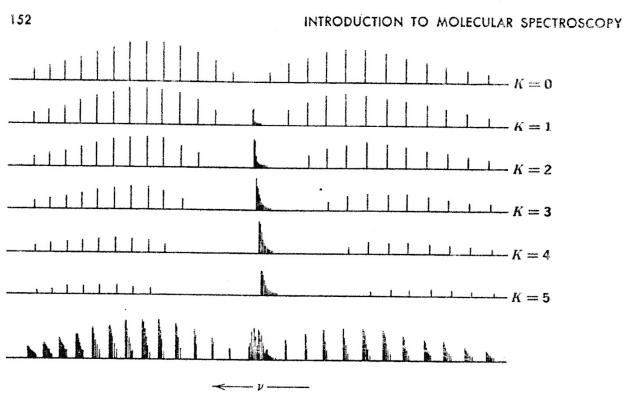


FIG. 7-11 The components of a parallel band showing the contributions from each of the  $K$  levels of the  $v = 0$  state.

Figure 9.5: A parallel band of a symmetric top, showing the sub-bands caused by different  $K$  ladders and the net result at the bottom. Figure and caption taken from Barrow, *Introduction to Molecular Spectroscopy*.

A perpendicular band of a symmetric top will have selection rules as follows:

$$\Delta K = \pm 1, \quad \Delta J = 0, \pm 1 \quad (9.22)$$

The band is composed of a sum of sub-bands, two for each initial  $K \neq 0$ , depending on whether  $\Delta K$  is  $+1$  or  $-1$ . Each sub-band consists of P, Q, and R sub-branches centered at a frequency displaced from the band center ( $\nu_0$ ) by an amount,

$$\begin{aligned} \Delta\nu &= (A - B) + 2(A - B)K \quad \text{for } \Delta K = +1, \quad K = 0, 1, 2, \dots \\ \Delta\nu &= (A - B) - 2(A - B)K \quad \text{for } \Delta K = -1, \quad K = 1, 2, \dots \end{aligned} \quad (9.23)$$

[

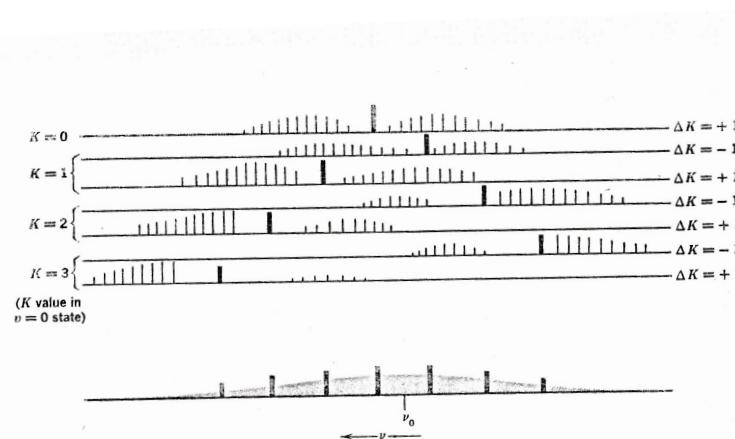


FIG. 7-14 The components of a perpendicular band showing the contributions from each of the  $K$  levels of the  $v = 0$  state.

Figure 9.6: A perpendicular mode of a symmetric top, showing the sub-band structure, with two displaced sub-bands for each  $K \neq 0$ . Figure and caption taken from Barrow, *Introduction to Molecular Spectroscopy*.

The net result may be a series of Q branches superimposed on an unresolved background of many lines. The spacing of the Q branches will be  $2(A - B)$ . An example is shown below.

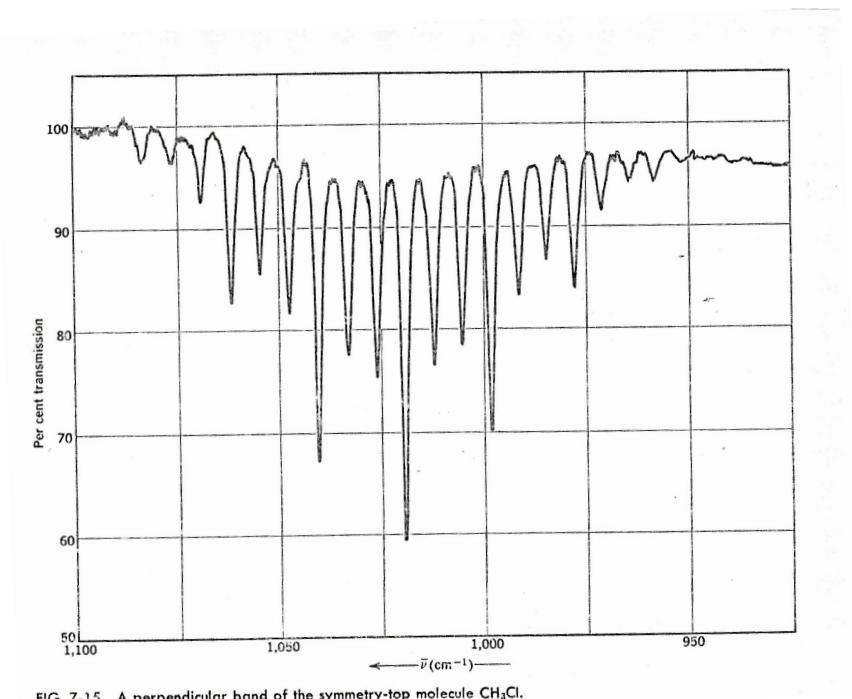


FIG. 7-15 A perpendicular band of the symmetry-top molecule CH<sub>3</sub>Cl.

Figure 9.7: The resulting spectrum of a perpendicular band, in this case of CH<sub>3</sub>Cl. Figure and caption taken from Barrow, *Introduction to Molecular Spectroscopy*.

## Chapter 10

# Including Other Angular Momenta

### 10.1 Molecules with Residual Electronic Angular Momentum

In our analysis of molecular rotation, we have so far assumed that only the relative motion of the nuclei contribute angular momentum. That was not unreasonable because the ground electronic state of most molecules is a  ${}^1\Sigma$ , indicative of zero spin and zero orbital angular momentum from the electrons. The predominance of this ground state results from the tendency of atoms to form bonds that fill all the available orbitals, as discussed in Chapter 6. If all the allowed bond are not satisfied, there will be unpaired electrons, so neither the spin nor the electronic angular momenta can cancel completely. In the laboratory, such **radicals** react quickly to form more stable molecules, but the rare collisions in even “dense” parts of the interstellar medium allow radicals and molecular ions to have substantial abundances. Consequently, molecules with residual electronic angular momentum play a much larger role in astronomy than in terrestrial conditions. In addition, about one in a thousand stable molecules have a ground electronic state other than  ${}^1\Sigma$ , notably O<sub>2</sub> and SO (both  ${}^3\Sigma$ ). Thus, we need to consider the effects of including the residual angular momentum in rotational spectra. In addition, because excited states of electronic transitions are usually not  ${}^1\Sigma$ , we need this information before we pull together everything to discuss electronic transitions including vibrational and rotational structure. Finally, molecules with residual electronic angular momenta have relatively large Zeeman effects, so they are important for measuring magnetic fields.

We will restrict considerations to linear molecules for simplicity. As discussed in Chapter 6.3, it is the *projection* of the electronic angular momentum on the molecular symmetry axis that is (to first order) a good quantum number, not the angular momentum itself. The projection of the orbital angular momentum is labeled  $\Lambda$ , and the projection of the spin angular momentum is labeled  $\Sigma$  (not to be confused with the  $\Sigma$  state!). The sum of electronic orbital and spin angular momenta, projected on the symmetry axis, is labeled  $\Omega$ . Here conventions differ a bit, but generally,  $\Omega = |\Lambda + \Sigma|$ .

If we wish to reserve  $J$  for the total angular momentum (excluding nuclear spin, which we will deal with later), we need a new symbol for the rotation of the molecule. Here, conventions diverge, and I will use the notation of Townes and Schawlow and Gordy and Cook, in with  $O$  represents the rotation, and  $N = \Lambda + O$  is the total “orbital” angular momentum.

To understand how all these angular momenta couple, we turn to our old friend Hund. In

principle, there are five Hund's coupling cases, but only two, cases a and b, are very relevant. Note that, like Russell-Saunders versus  $jj$  coupling, these are idealized limits, and real molecules may be intermediate between them.

In Hund's case a, the electronic angular momenta couple to the symmetry axis ( $\vec{A}$ ) and form a projection,  $\Omega = |\Lambda + \Sigma|$ . This adds vectorially to the rotation  $\vec{O}$  to make  $\vec{J}$ .  $\vec{\Omega}$ ,  $\vec{A}$ , and  $\vec{O}$  precess around  $\vec{J}$ .

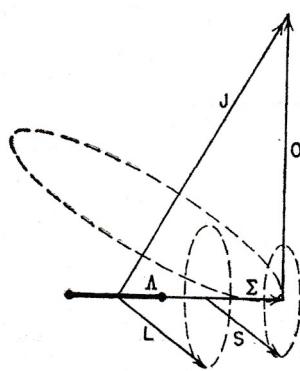


Figure 10.1: Hund's case a.  $\vec{L}$  and  $\vec{S}$  precess rapidly around the molecular axis, which precesses more slowly around  $\vec{J}$ , the total angular momentum.  $\vec{O}$  is the usual molecular rotation, which is perpendicular to the molecular axis. Figure from Townes and Schawlow.

Notice that this situation is similar to that of the symmetric top, with  $\Omega$  playing the role of  $K$ . There will be a series of ladders of different  $\Omega$ , with a lowest state having  $J = \Omega$ . A good example of Hund's case a is NO, which has a  $^2\Pi$  ground state. We use a subscript to indicate the value of  $\Omega$ . Here,  $\Lambda = 1$  and  $\Sigma = \frac{1}{2}$ , so  $\Omega$  can be either 0.5 or 1.5, giving us two ladders labeled  $^2\Pi_{\frac{1}{2}}$  and  $^2\Pi_{\frac{3}{2}}$ , where the subscript is the value of  $\Omega$ . The spin-orbit coupling in NO gives a “normal” doublet, with  $^2\Pi_{\frac{1}{2}}$  as the lowest state.

As the molecule rotates faster, the coupling of  $\vec{O}$  and the electronic angular momenta increases, and the electronic angular momenta may decouple from the molecular axis and instead couple to  $\vec{O}$ .  $\vec{S}$  decouples first, as it is really coupled indirectly through a spin-orbit interaction with  $\vec{L}$ . This leads to Hund's case b, in which  $\Lambda$  and  $\vec{O}$  couple to make at total angular momentum  $\vec{N}$  and  $\vec{S}$  then couples to  $\vec{N}$  to make  $\vec{J}$ . The spin-orbit coupling will be absent in molecules with  $L = 0$ , so the spin will couple to  $\vec{J}$  rather than  $\vec{A}$ . In this case, we have  $\vec{O} = \vec{N}$ .

Molecules of this type include CN ( $^2\Sigma$ ), O<sub>2</sub> ( $^3\Sigma$ ), and SO ( $^3\Sigma$ ). Let's consider first a  $^2\Sigma$  molecule. Starting from the rotational energy levels of a  $^1\Sigma$  molecule, we would observe a splitting of the levels depending on whether  $\vec{S}$  aligns parallel to anti-parallel to  $\vec{O} = \vec{N}$ . This is called spin doubling. For CN, for example, the  $O = 1 \rightarrow 0$  transition is split into two lines around 113.1 and 113.5 GHz.

For a  $^3\Sigma$  molecule, such as O<sub>2</sub>, we have  $S = 1$ , so each value of  $N = O$  is split into three, with  $J = N + S, N + S - 1, N - S$ , or  $N + 1, N, N - 1$  for  $S = 1$ . These will be split by spin-rotation interaction, so the energy level diagram will show a series of triplet states, called spin triplets or  $\rho$ -type triplets.

We can understand the energy levels and selection rules fairly simply. Their zeroth order energy

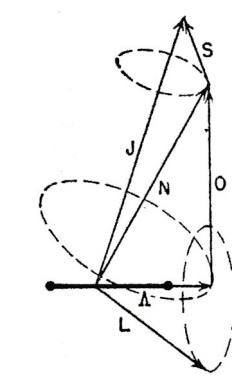


Figure 10.2: Hund's case b. Precession of the molecular axis is slower than the precession of  $\vec{L}$  about the axis, but faster than the precession of  $\vec{N}$  and  $\vec{S}$  about  $\vec{J}$ , the total angular momentum.  $\vec{O}$  is the usual molecular rotation, which is perpendicular to the molecular axis, and  $\vec{N} = \vec{\Lambda} + \vec{O}$ . Figure from Townes and Schawlow.

levels are the usual ones with  $N$  replacing  $J$ :  $E_N = B_v N(N + 1)$ . We will stick to the rigid rotor approximation for simplicity, but all the usual corrections can be applied. We now consider a perturbation caused by the interaction of the electronic spin magnetic moment with a weak magnetic field generated by the rotation. The magnetic field is generated because the rotation mixes a slight contribution from excited electronic states that are not  ${}^1\Sigma$  (a small breakdown of the Born-Oppenheimer approximation). Then the perturbation Hamiltonian is

$$H_M = -\gamma \vec{S} \cdot \vec{N} = -\gamma |\vec{S}| |\vec{N}| \cos(\vec{S}, \vec{N}) \quad (10.1)$$

where  $\gamma$  is the spin-rotation interaction constant. Analogously to the vector model of the spin-orbit interaction in the hydrogen atom (see equation 3.61), we have

$$\vec{J}^2 = \vec{N}^2 + \vec{S}^2 - 2|\vec{S}||\vec{N}| \cos(\vec{S}, \vec{N}) \quad (10.2)$$

So we can express  $H_M$  as

$$+\gamma/2 (\vec{J}^2 - \vec{N}^2 - \vec{S}^2) \quad (10.3)$$

yielding

$$E_M = +\gamma/2 (J(J + 1) - N(N + 1) - S(S + 1)) \quad (10.4)$$

For  ${}^2\Sigma$ ,  $S = \frac{1}{2}$ , so  $J = N + \frac{1}{2}$  and  $N - \frac{1}{2}$ , so we have  $E_M = \frac{1}{2}\gamma N$  or  $-\frac{1}{2}\gamma(N + 1)$ , so

$$E_{N+\frac{1}{2}} = B_v N(N + 1) + \frac{1}{2}\gamma N \quad (10.5)$$

and

$$E_{N-\frac{1}{2}} = B_v N(N + 1) - \frac{1}{2}\gamma(N + 1) \quad (10.6)$$

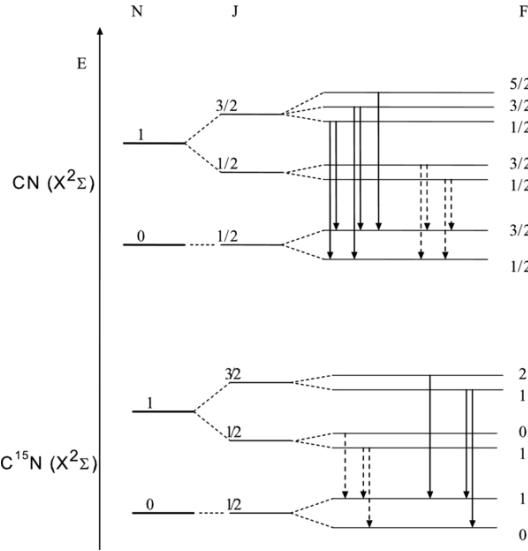


Figure 10.3: The energy levels for CN. For now, we have explained only the first splitting. The further (hyperfine) splitting to the right is due to interactions with nuclear spins.

The situation for  $^3\Sigma$  molecules would seem to be a straightforward extension, with  $J = N + 1, N, N - 1$ , but the two electron spins also have a spin-spin interaction, which is usually much larger than the spin-rotation interaction. If it is parameterized by  $\lambda$ . So we have three states with energies given by

$$\begin{aligned} E_{N+1} &= B_v N(N+1) + \gamma(N+1) - \frac{2\lambda(N+1)}{2N+3} \\ E_N &= B_v N(N+1) \\ E_{N-1} &= B_v N(N+1) - \gamma N - \frac{2\lambda N}{2N-1} \end{aligned} \quad (10.7)$$

In some cases, the spin-spin interaction can be larger than  $B_v$  so the “splitting” actually scrambles the usual rotational energy level structure. The energies in equation 10.11 are only approximate as the spin-spin interaction tends to align the spins along the molecular axis, making the molecule intermediate between case a and case b.

The selection rules for electric dipole transitions are

$$\Delta N = \pm 1; \quad \Delta J = 0, \pm 1 \quad (10.8)$$

One of the most important  $^3\Sigma$  molecules, both in space and in the Earth’s atmosphere, is  $O_2$ , which has no electric dipole moment owing to its symmetry. But it does have a magnetic dipole because of the unpaired spins, and magnetic dipole transitions of two distinct series are possible.

$$\begin{aligned} \text{Series1 : } \Delta N &= 0, \quad \Delta J = \pm 1 \\ \text{Series2 : } \Delta N &= \pm 2, \quad \Delta J = 0, \pm 1 \end{aligned} \quad (10.9)$$

The first series represents transitions within a spin triplet. Most of these for O<sub>2</sub> occur near 60 GHz or about 5 mm wavelength. The array of lines, pressure broadened by the Earth's atmosphere make the atmosphere essentially opaque around those wavelengths. One line ( $N = 1 \rightarrow 1, J = 1 \rightarrow 0$ ) occurs at about 118.8 GHz (2.3 mm) and the wings of this transition attenuate and add noise to observations of the CO  $J = 1 \rightarrow 0$  transition.

The second series starts at about 424 GHz with  $N = 3 \rightarrow 1; J = 2$ . After many years of searches, O<sub>2</sub> was finally detected unambiguously in space in several lines of this series by Goldsmith et al. 2011, Ap. J., 737, 96.

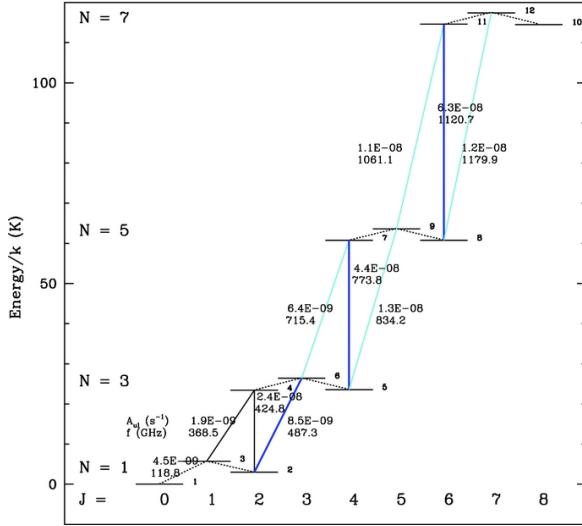


Figure 10.4: The energy level diagram of O<sub>2</sub>, showing transitions, frequencies, and Einstein A values. The transitions used to detect the molecule in space are indicated in blue. This figure is taken from Goldsmith et al. 2011, Ap. J., 737, 96.

Another important  $^3\Sigma$  molecule is SO, which has an electric dipole moment in addition to the magnetic one. In addition to the two series above, it has two more series of transitions, via electric dipole.

$$\begin{aligned} \text{Series3 : } & \Delta N = \pm 1, \quad \Delta J = \pm 1 \\ \text{Series4 : } & \Delta N = \pm 1, \quad \Delta J = 0 \end{aligned} \tag{10.10}$$

Series 3 is more or less pure rotational, while series 4 mixes rotation and spin transitions.

$$(10.11)$$

High rotational states, especially of light hydrides, can also tend toward case b. OH ( ${}^2\Pi$ ) is case b in high rotational states and intermediate for low rotational states. OH was the first interstellar molecule to be discovered with radio spectroscopy, and research is again active with this radical, so we consider it in more detail.

The spin-orbit coupling in OH causes the  $^2\Pi_3$  state to be the lowest (an “inverted” doublet). We would have two rotational ladders, as shown in figure 10.5. Rotational transitions ( $\Delta O = 1$ ) are in the far-infrared because OH has a very small moment of inertia (e.g., the longest wavelength in the  $^2\Pi_3$  ladder is  $119 \mu\text{m}$ ). Detection required airborne or spacecraft, and routine studies of these transitions became possible only with the *Herschel* space telescope.

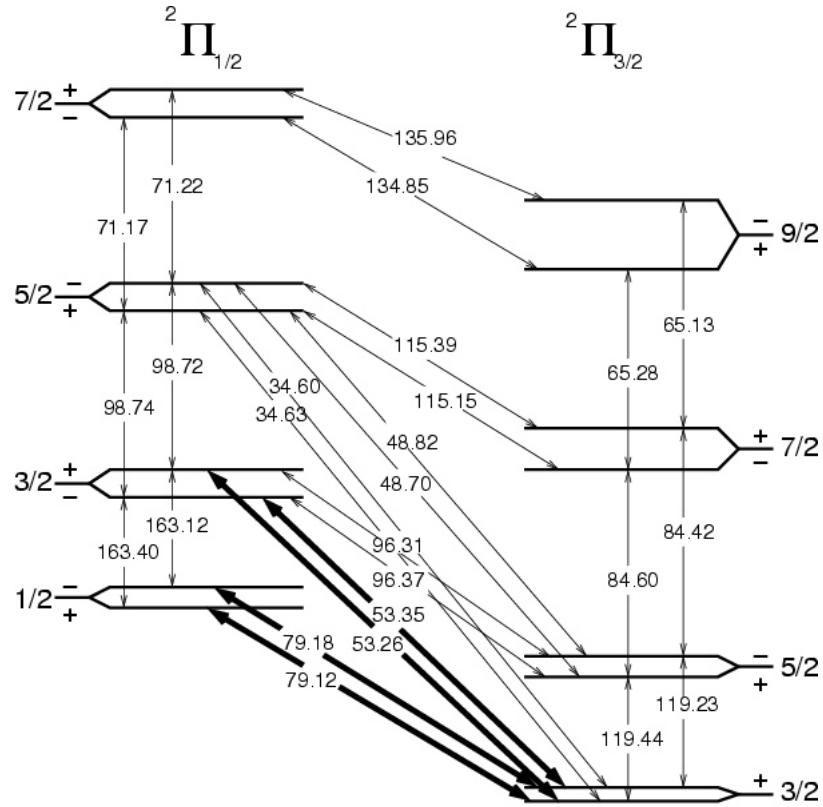


Figure 10.5: OH energy levels with  $\Lambda$  doubling.

Figure 10.5 reveals a splitting of each rotational level, in which the degeneracy between  $\pm\Lambda$  is split. By analogy to  $K$ -doublets in slightly asymmetric tops, these levels are called  $\Lambda$ -doublets. It is caused by a breakdown of the strong coupling of  $\vec{L}$  to  $\vec{A}$ , the molecular axis, owing to some slight interaction with the rotation of the molecule.

The simplest way to understand  $\Lambda$ -doubling is to pursue the analogy with  $K$ -type doubling, using H<sub>2</sub>CO, a slightly asymmetric top, which is a planar molecule, with two hydrogens off the line of the C-O axis. Using the notation of §8.4, the moment of inertia about the  $c$  axis is slightly greater than the moment of inertia about the  $b$  axis because the hydrogens are off the  $c$  axis. For OH, no nuclei are off the molecular axis, but we can think of the electrons as being off axis. Because OH is a radical, one valence electron of O is unpaired, so one  $\pi$  orbital contains two electrons and the other contains only one. Like H<sub>2</sub>CO, we have slightly different masses off axis and

thus slightly different moments of inertia. The rotational levels are thus split by a factor roughly equal to  $m_e/m_p \sim 1/1800$ . The actual splitting corresponds to a wavelength of about 18 cm for the  $^2\Pi_{\frac{3}{2}} J = \frac{3}{2}$  state. Transitions between these levels were detected from space in 1963, and they were found to be exhibiting maser emission soon after.

Continuing the analogy with asymmetric tops, the splitting and transition matrix element for  $\Lambda$ -doublets in  $\Lambda = 1$  states are approximately (cf. equation 8.34).

$$\Delta w = q_\Lambda J(J+1) \quad (10.12)$$

and

$$|\mu_{ij}|^2 = \mu^2 \frac{\Lambda^2}{J(J+1)} \quad (10.13)$$

The splitting is much less in states with  $\Lambda > 1$  because the orbitals are more symmetric about the axis.

Another important  $^2\Pi$  interstellar molecule with  $\Lambda$ -doubling is CH.

# Chapter 11

## Electronic Transitions

With our more complete picture of both rovibrational transitions and residual electronic angular momentum, we can turn to the topic of transitions between electronic states. In general, these electronic transitions also involve changes of vibrational and rotation states, so they are sometimes called **vibronic** transitions.

### 11.1 Generalities and Nomenclature

If we ignore the translational energy in equation 6.17, we can specify the total internal energy of the molecule as

$$E = E_\epsilon + E_v + E_r \quad (11.1)$$

where  $\epsilon$ ,  $v$ , and  $r$  represent whatever array of quantum numbers are needed to specify the electronic, vibrational, and rotational states, respectively. More generally,  $E_\epsilon$  is a function of internuclear distance, but electronic transitions are assumed to be fast compared to internuclear motions, so that transitions occur at nearly fixed separation.

With our usual convention that double primes indicate the lower state of a transition, we have the frequency of a vibronic transition,

$$\nu = \nu_{\epsilon' \epsilon''} + \nu_{v' v''} + \nu_{r' r''} \quad (11.2)$$

with

$$\nu_{\epsilon' \epsilon''} = \frac{E_{\epsilon'} - E_{\epsilon''}}{\hbar}, \quad \text{etc.} \quad (11.3)$$

An electronic band **system** consists of all the possible  $\nu_{v' v''}$  and  $\nu_{r' r''}$  for a given electronic transition (fixed  $\epsilon', \epsilon''$ ). If we also fix  $v', v''$ , we have a particular **band**. A set of bands with the same  $v''$  (or  $v'$ ) is called a  $v''$  (or  $v'$ ) **progression**. At modest temperatures, we can expect that only the lowest electronic and vibrational states will be occupied. Then we could only see an absorption spectrum with  $v'' = 0$  progressions. A group of bands with the same  $v' - v''$  is called a **band sequence**.

If we neglect the rotational structure for a moment, we can get the band center by adding  $\nu_{\epsilon'\epsilon''}$  to a suitable expression for the vibrational frequencies, such as equation 7.21. However, the vibrational frequencies and anharmonicities may be different for the different electronic states. Consequently, we use the Deslandres formula to calculate the band center:

$$\nu_D = \nu_{\epsilon'\epsilon''} + \omega_e'(v' + \frac{1}{2}) - \omega_e'x'(v' + \frac{1}{2})^2 - \omega_e''(v'' + \frac{1}{2}) + \omega_e''x_e''(v'' + \frac{1}{2})^2 \quad (11.4)$$

with all in the same units, such as  $\text{cm}^{-1}$ .

Now, let's consider rotational structure. We will discuss selection rules later, but for now assume that P, Q, and R branches are possible and consider the frequencies. An equation similar to 7.25 for the P and R branches is

$$\nu = \nu_D + (B' + B'')m + (B' - B'')m^2 \quad (11.5)$$

where  $\nu_D$  is given by the Deslandres formula, and  $m = -J$  for the P branch and  $m = J + 1$  for the R branch. For the Q branch, we have

$$\nu = \nu_D + (B' + B'')J + (B' - B'')J^2 \quad (11.6)$$

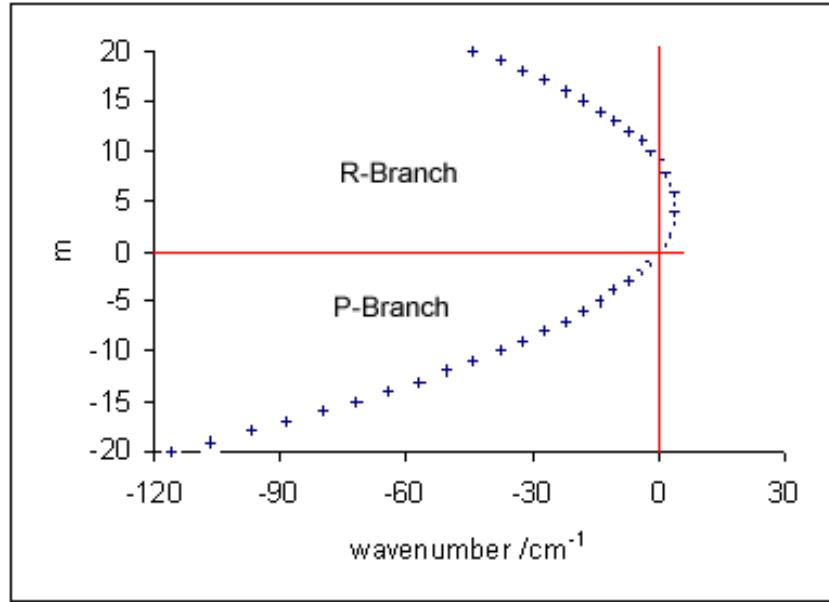


Figure 11.1: A Fortrat parabola. This illustration assumes  $B' = 0.8$  and  $B'' = 1.0 \text{ cm}^{-1}$ . In this case, the bandhead appears at relatively low  $J$  in the R branch. Figure from Wikipedia page on vibronic spectroscopy.

In both these equations, we have set  $J'' = J$ . Besides the use of the Deslandres formula, a very significant difference from the rovibrational formulas is that the  $B$  values in the upper and lower states are now in different electronic states, which may have minima at different  $R_0$ . Thus the

$B$  values may differ due to the leading term  $B_e \sim R_0^{-2}$ , not just in the second-order vibration-rotation interaction term. Consequently, band heads may form in either P or R branches and in fairly low  $J$  states. A plot of  $m$  versus frequency is called a **Forrat parabola**, illustrated in figure 11.1. Finally, note that there is no selection rule for the vibrational quantum numbers, even in the harmonic oscillator potential because the final and initial states are in different electronic states.

## 11.2 Selection Rules and Strengths

To calculate the strength of an individual transition, we would start with the following equation.

$$\langle \epsilon' v' r' | -e \sum_i \vec{r}_i | \epsilon'' v'' r'' \rangle = \langle v' r' | \vec{\mu}_{el}(R) | v'' r'' \rangle \quad (11.7)$$

where

$$\vec{\mu}_{el}(R) = \langle \epsilon' | -e \sum_i \vec{r}_i | \epsilon'' \rangle \quad (11.8)$$

is the electric dipole moment which depends on the sum of all electronic locations, but, for a diatomic molecule, it will generally be a parametric function of the internuclear separation ( $R$ ). If we assume that the transition is so fast compared to vibrational motions that  $R$  is essentially fixed, the **Franck-Condon Principle**, we can separate equation 11.7 into an electronic matrix element between rotational states and a **Frank-Condon** factor  $\langle v' | v'' \rangle$ :

$$\langle \epsilon' v' r' | \vec{\mu}_{el}(R) | \epsilon'' v'' r'' \rangle = \langle r' | \vec{\mu}_{el} | r'' \rangle \langle v' | v'' \rangle \quad (11.9)$$

Note that  $\langle v' | v'' \rangle \neq 0$  for  $v' = v''$  because the vibrational states are not in the same electronic state, hence are not necessarily orthogonal. Consequently, there is no selection rule for  $\Delta v$ . In fact, the overlap integral depends on the equilibrium radii in the two electronic states as well as the vibrational wave functions (see Fig. 11.2).

Given the behavior of the vibrational wave functions, a transition from  $v''$  to a high  $v'$  state will be favored if a vertical line can be drawn between the center of the  $v''$  state to the end (classical turning point) of the  $v'$  state. The favored transitions are illustrated in Figure 11.2.

The rotational part of the matrix element is  $\langle r' | \vec{\mu}_{el} | r'' \rangle$ . As with rovibrational transitions, this will factor into a magnitude and a part that depends only on the direction cosines. The part with the direction cosines will involve the usual rotational matrix elements while the magnitude will now be the matrix element of the dipole moment operator between different electronic states rather than the expectation value in a single state, as it was for the rovibrational transitions (see equations 7.29 to 7.36).

The selection rules arise as usual from symmetries like rotational, parity, and other reflection symmetries. We begin with the most fundamental and work our way into the most contingent rules. We will consider only electric dipole transitions.

1. Conservation of angular momentum gives us the usual rule,

$$\Delta J = 0, \pm 1; \quad J = 0 \not\rightarrow 0 \quad (11.10)$$

2. Parity must change. This is complicated because the parity operation is defined in the laboratory frame, but the wave functions are known in the molecule frame. It turns out that the

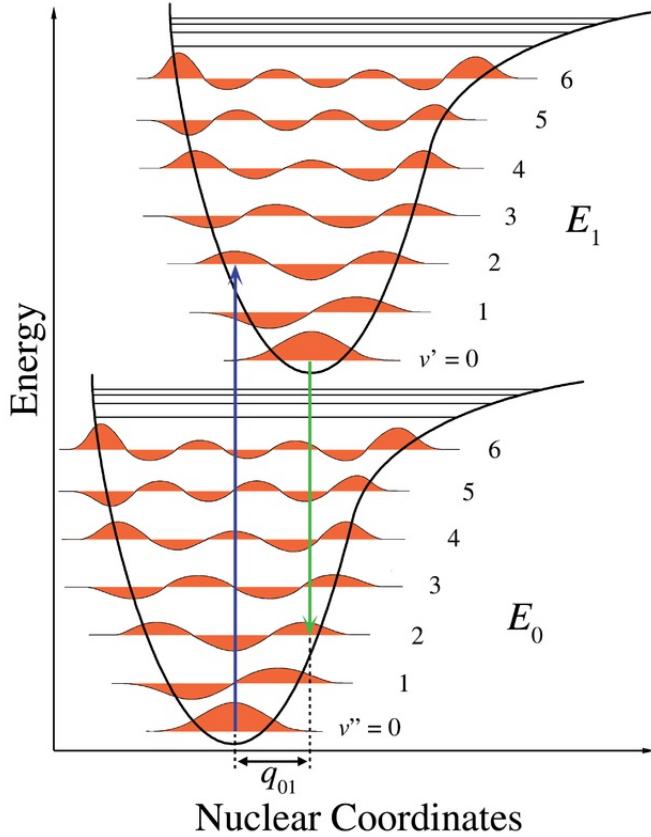


Figure 11.2: Illustration of the Franck-Condon principle. A transition between two electronic states tends to occur at constant internuclear separation. Absorption from the ground state to an excited electronic state with larger internuclear separation will likely occur to higher vibrational levels. Emission from the ground vibrational state of the excited electronic state is likely to result in a transition to the ground vibrational state of the ground electronic state. Figure from Wikipedia page by Mark Somoza.

inversion operator ( $i$ ) in the laboratory frame is equivalent to the  $\sigma_v$  operation in the molecular frame ( $x_i, y_i, z_i \rightarrow x_i, -y_i, z_i$ ). The nuclear wavefunctions are not involved in this parity. The parity is denoted simply as + or -. So the parity rule is

$$+ \leftrightarrow -, \quad + \leftrightarrow +, \quad - \leftrightarrow - \quad (11.11)$$

3. If the molecule consists of two equal charge nuclei, there is an inversion symmetry in the molecular frame, again applied only to the electronic orbital part of the wave function. This is the gerade/ungerade symmetry. Because the dipole moment operator has ungerade symmetry, we have

$$g \leftrightarrow u, \quad g \leftrightarrow g, \quad u \leftrightarrow u \quad (11.12)$$

4. If in addition, we have identical nuclei, we must have constructed explicitly symmetric or anti-symmetric nuclear spin functions, labeled  $a$  or  $s$ . Because the electric dipole moment does not affect the nuclear spins, it will not change this symmetry, so we have

$$a \leftrightarrow a \quad s \leftrightarrow s, \quad s \leftrightarrow a \quad (11.13)$$

Note that the labels apply to the wavefunction excluding nuclear spin. For two nuclei with  $I = \frac{1}{2}$ , there is a symmetric (triplet,  $S = 1$ ) and an antisymmetric (singlet,  $S = 0$ ) state. For fermions, the **total** wave function must be anti-symmetric, so the triplet is associated with the anti-symmetric wave function, excluding nuclear spin. For example, ortho-H<sub>2</sub> would be labeled  $a$ .

The rest of the selection rules depend on the coupling of angular momenta. If the two states involved in the transition belong to different Hund's coupling cases, then only the selection rules valid for **both** cases apply.

### 11.2.1 Selection Rules Applying to Both Hund's Case a and Hund's Case b

For both these cases,  $\Lambda$  is a good quantum number, so we have

$$\Delta\Lambda = 0 \pm 1 \quad (11.14)$$

corresponding to the  $\Delta M_L = 0, \pm 1$  in atoms. If the molecule has  $\Lambda = 0$ , it has  $\sigma_v$  symmetry, so we have  ${}^+\Sigma$  and  ${}^-\Sigma$  states. This symmetry cannot change in an electric dipole transition:

$${}^+\Sigma \leftrightarrow {}^+\Sigma, \quad {}^-\Sigma \leftrightarrow {}^-\Sigma, \quad {}^+\Sigma \leftrightarrow {}^-\Sigma, \quad (11.15)$$

Since states with  $\Lambda = 0$  have two degenerate states with opposite symmetry, either  ${}^+\Sigma$  or  ${}^-\Sigma$  states can transition to  $\Pi$  states. For both pure cases a and b, the total electronic spin is a good quantum number, so intercombination transitions are forbidden

$$\Delta S = 0 \quad (11.16)$$

as for atoms. Also, as with atoms, this is approximate; as  $\Lambda$  and  $S$  interact more strongly, intercombination bands will get stronger.

### 11.2.2 Selection Rules Applying only to Case a

In case a, the projection of  $\vec{S}$  on the molecular axis is a good quantum number, so we have a selection rule on  $\Sigma$ :

$$\Delta\Sigma = 0 \quad (11.17)$$

equivalent to  $\Delta M_S = 0$ , which applies to an atom in a strong electric or magnetic field. In addition, we have a rule for  $\Omega = |\Lambda + \Sigma|$ :

$$\Delta\Omega = 0, \pm 1 \quad (11.18)$$

which holds even for Hund's case c, when  $\Lambda$  and  $\Sigma$  lose their meaning. When  $\Omega = 0$ , there is no electronic angular momentum, so we revert to the simple rule

$$\Delta J = \pm 1 \quad (11.19)$$

so only P and R branches.

### 11.2.3 Selection Rules Applying only to Case b

In this case, the total **orbital** angular momentum,  $N = O + \Lambda$  is defined and we have

$$\Delta N = 0, \pm 1 \quad (11.20)$$

If the transition is  $\Sigma - \Sigma$ , then  $\Delta N = 0$  is not allowed.

### 11.2.4 Summary

To assemble all this information **Herzberg** considers each kind of transition case by case (pp 240-280 of *Diatom Molecules*). Alternatively, **Bernath** (pp. 309-324 of *Spectra of Atoms and Molecules*) considers first Singlet-Singlet transitions, then transitions from non-Singlet states.

For the simplest case,  ${}^1\Sigma - {}^1\Sigma$  transitions of diatomic molecules, we have only P and R branches, just as for rovibrational transitions. If one of the states has electronic angular momentum, we can think of the diatomic molecule as a symmetric top with  $\Lambda$  playing the role of  $K$ , the projection of the angular momentum on the molecular symmetry axis. So, we can have either parallel ( $\Delta\Lambda = 0$ ) or perpendicular ( $\Delta\Lambda = \pm 1$ ) bands, depending on whether the changing electric dipole moment is parallel or perpendicular to the molecular axis.

1. Parallel Bands ( $\Delta\Lambda = 0$ ). If both  $\Lambda'$  and  $\Lambda''$  are 0 ( $\Sigma$  states), we have only P and R branches, If  $\Lambda'' = \Lambda' \neq 0$  (e.g.,  ${}^1\Pi - {}^1\Pi$ ) there can also be Q branches, but they are weaker than the P and R branches.
2. Perpendicular Bands ( $\Delta\Lambda = \pm 1$ ). These transitions (e.g.,  ${}^1\Pi - {}^1\Sigma$ ) can have strong Q branches in addition to P and R branches.

The equations below give the Hönl-London factors for R, Q, and P branches for each case of the change of  $\Lambda$ . We have substituted  $J$  for  $J''$  for simplicity. Remember that these refer to the  $J$  value in the **lower** state. These can be used for vibrational transitions as well.

$$\begin{aligned} \Delta\Lambda = 0 \quad S_J^R &= \frac{(J+1+\Lambda)(J+1-\Lambda)}{J+1} \\ S_J^Q &= \frac{(2J+1)\Lambda^2}{J(J+1)} \\ S_J^P &= \frac{(J+\Lambda)(J-\Lambda)}{J} \end{aligned} \quad (11.21)$$

$$\begin{aligned} \Delta\Lambda = +1 \quad S_J^R &= \frac{(J+2+\Lambda)(J+1+\Lambda)}{4(J+1)} \\ S_J^Q &= \frac{(J+1+\Lambda)(J-\Lambda)(2J+1)}{4J(J+1)} \\ S_J^P &= \frac{(J-1-\Lambda)(J-\Lambda)}{4J} \end{aligned} \quad (11.22)$$

$$\begin{aligned}\Delta\Lambda = +1 \quad S_J^R &= \frac{(J+2-\Lambda)(J+1-\Lambda)}{4(J+1)} \\ S_J^Q &= \frac{(J+1-\Lambda)(J+\Lambda)(2J+1)}{4J(J+1)} \\ S_J^P &= \frac{(J-1+\Lambda)(J+\Lambda)}{4J}\end{aligned}\tag{11.23}$$

Putting all this together, the matrix element squared for an individual electronic, vibrational, and rotational transition from a lower state ( $l$ ) to an upper state ( $u$ ) is given by

$$|\mu_{l \rightarrow u}|^2 = |\mu_{el}|^2 |\langle v' | v'' \rangle|^2 S_J^x / (2J+1) \tag{11.24}$$

where  $x$  is R, P, or Q,  $|\langle v' | v'' \rangle|^2$  is the Franck-Condon factor, and  $|\mu_{el}|^2$  is the electronic dipole moment matrix element between final and initial states.

## Chapter 12

# Hyperfine Structure in Molecules

### 12.1 Introduction

The extreme frequency resolution and accuracy of heterodyne spectroscopy, together with the low velocity dispersions, in molecular clouds, allows separation of hyperfine structure especially at lower radio frequencies. The hyperfine splitting tends to range from kHz to MHz in frequency; Doppler broadening of  $1 \text{ km s}^{-1}$  will exceed 1 kHz for  $\nu > 300 \text{ MHz}$ . The hyperfine structure can be quite useful in constraining optical depths of lines. Because different hyperfine transitions will have different transition strengths, they will become optically thick at different column densities. Thus, observed ratios of hyperfine components different from the optically thin (sometimes called LTE) ratios can be used to measure the optical depth of the stronger components.

The hyperfine structure arises, as for atoms, from interactions with the nuclear charge distributions or spins. The possible interactions depend on the nuclear spin,  $I$ . If  $I = 0$ , only an electric charge is possible. If  $I \geq \frac{1}{2}$ , electric charge and magnetic dipole are allowed. If  $I \geq 1$ , electric charge, magnetic dipole, and electric quadrupole are allowed. Recalling that nuclei with even numbers of protons and even numbers of neutrons must have  $I = 0$ , we see that many of the most common isotopes in astronomy (e.g.,  $^{12}\text{C}$ ,  $^{16}\text{O}$ , and  $^{32}\text{S}$ ) will have no hyperfine contribution. Much of the common hyperfine structure arises when nitrogen ( $^{14}\text{N}$ ) or rarer isotopes of C and O ( $^{13}\text{C}$ ,  $^{17}\text{O}$ ) are present in the molecules (see Fig. 12.1). In addition, hydrogen ( $^1\text{H}$ ), with  $I = \frac{1}{2}$ , has a magnetic moment that can produce weak hyperfine splitting. If the molecule contains deuterium ( $^2\text{H}$ ), there is an electric quadrupole moment. One of the important molecules for determining conditions in very cold, dense molecular clouds is  $\text{N}_2\text{D}^+$ , a molecular ion with “interesting” hyperfine structure!

### 12.2 Molecules with Residual Electronic Angular Momentum

Molecules with residual electronic angular momentum have an interaction with the nucleus similar to that in atoms, as discussed in chapter 3.4. The nuclear magnetic moment interacts with the magnetic field created by the electron’s orbital or spin angular momentum. Assuming that the orbital electronic angular momentum is coupled to the molecular axis, only the component of  $\vec{L}$

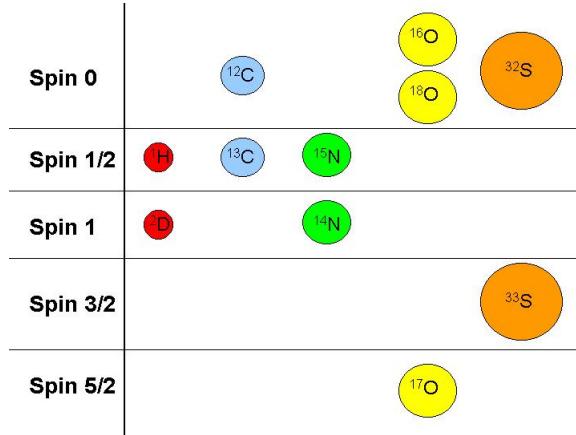


Figure 12.1: The spin of nuclei in common astrophysically observed molecules.

along the molecular axis will contribute. We label this as

$$\vec{L}_k = \Lambda \hat{k} \quad (12.1)$$

where  $\hat{k}$  is a unit vector along the molecular axis. The perturbation due to the interaction of the magnetic field from the electron's orbital motion with the nuclear magnetic moment is then

$$H_{IL} = \frac{2\mu_B\mu_I}{I} \sum_n \left\langle \frac{1}{r_n^3} \right\rangle \Lambda_n \vec{I} \cdot \hat{k} \quad (12.2)$$

where  $\vec{I}$  is the nuclear spin, the sum is over all electrons, and the  $n^{th}$  electron has a projection of its orbital angular momentum on the molecular axis of  $\Lambda_n$  and is a distance  $r_n$  from the nucleus. The contribution from paired electrons will cancel, so the only contribution is from unpaired electrons.

The interaction with the electronic spin is more complex. The classical interaction between two dipoles gives an Hamiltonian of

$$H_{IS}^1 = \frac{-2\mu_B\mu_I}{I} \left[ \frac{\vec{I} \cdot \vec{S}}{r^3} - \frac{3(\vec{I} \cdot \vec{r})(\vec{S} \cdot \vec{r})}{r^5} \right] \quad (12.3)$$

There is also a term proportional to  $\vec{I} \cdot \vec{S}$  and the probability of finding the electron at the origin:

$$H_{IS}^2 = \frac{16\pi}{3} \frac{\mu_B\mu_I}{I} \psi^2(0) \vec{I} \cdot \vec{S} \quad (12.4)$$

All three perturbations can be assembled into the following formula for the total perturbation:

$$H' = a \Lambda \vec{I} \cdot \hat{k} + b \vec{I} \cdot \vec{S} + c (\vec{I} \cdot \vec{k}) (\vec{S} \cdot \vec{k}) \quad (12.5)$$

with

$$a = \frac{2\mu_B\mu_I}{I} \left\langle \frac{1}{r^3} \right\rangle \quad (12.6)$$

$$b = \frac{2\mu_B\mu_I}{I} \left\langle \frac{8\pi\psi^2(0)}{3} - \frac{3\cos^2(\theta) - 1}{2r^3} \right\rangle \quad (12.7)$$

and

$$c = \frac{3\mu_B\mu_I}{I} \left\langle \frac{3\cos^2(\theta) - 1}{r^3} \right\rangle \quad (12.8)$$

where  $\theta$  is the angle between the molecular axis and the vector  $\vec{r}$  from the nucleus to the electron. The term with  $a$  enters only if there is electronic orbital angular momentum. If the electron's wave function is spherically symmetric, the terms with the cosine will have zero expectation value, so only the term with  $\psi^2(0)$  would enter. If the unpaired electron's wave function has much  $s$  type behavior, that term dominates the hyperfine interaction. These expressions are only approximate, but they give the general sense of the interaction.

To proceed further, we need to know how the nuclear spin couples to other angular momenta. There is a generalization of Hunds coupling cases that includes nuclear spin, illustrated in figure 12.2.

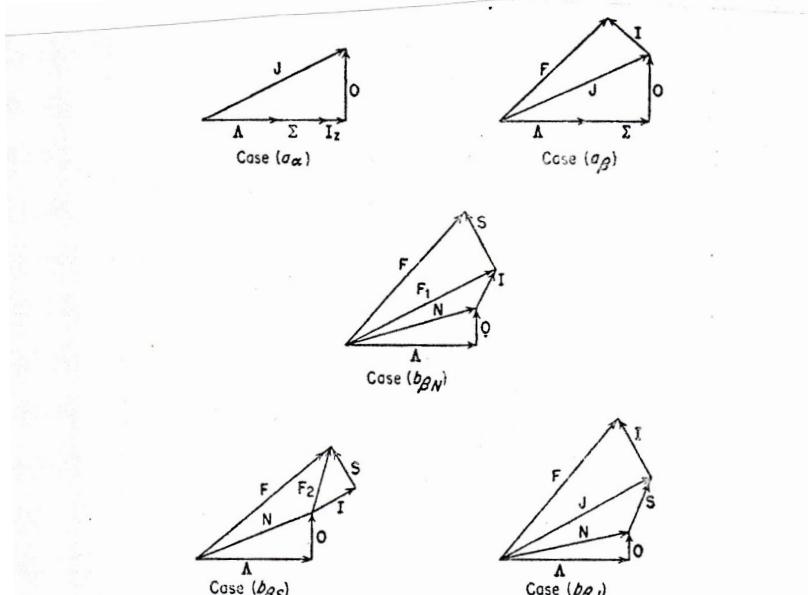


FIG. 8-1. Molecular coupling schemes including nuclear spin.

Figure 12.2: The possible Hunds coupling cases including nuclear spin,  $\vec{I}$  and total spin,  $\vec{F}$ . Figure from Townes and Schawlow, figure 8.1.

The subscript on case a or case b indicates the nuclear spin coupling:  $\alpha$  indicates strong coupling to the molecular axis;  $\beta$  indicates coupling to some other vector, indicated by a second subscript.

For case b, the nuclear spin is very unlikely to couple to the molecular axis because the electron spin, with its larger magnetic moment, has not done so; consequently, case  $b_\alpha$  does not exist.

For case a, equation 12.5 becomes

$$H' = [a\Lambda + (b + c)\Sigma] \vec{I} \cdot \hat{k} \quad (12.9)$$

because  $\vec{S} = \Sigma \hat{k}$  and  $\vec{S} \cdot \hat{k} = \Sigma$

NO, a  $^2\Pi$  molecule, is a good example of case  $a_\beta$ . The nitrogen magnetic moment couples to  $\vec{J}$ . The vector model gives

$$\vec{I} \cdot \hat{k} = \frac{(\vec{I} \cdot \vec{J})(\vec{J} \cdot \hat{k})}{J(J+1)} \quad (12.10)$$

and, since  $\vec{J} \cdot \hat{k} = \Lambda + \Sigma = \Omega$ , we have

$$E = [a\Lambda + (b + c)\Sigma] \frac{\Omega}{J(J+1)} \vec{I} \cdot \vec{J} \quad (12.11)$$

and

$$\vec{I} \cdot \vec{J} = \frac{F(F+1) - J(J+1) - I(I+1)}{2} \quad (12.12)$$

where  $F$  is the **total** angular momentum, including nuclear spin.

These expressions give a good description of the hyperfine effects in  $^2\Pi_{3/2}$  states. For the  $^2\Pi_{1/2}$  states of NO, and for OH, there is an additional effect called hyperfine doubling. Because the electron density (think of one orbital with two electrons, one with one) has different distributions for the two states of the  $\Lambda$  doublet (see Fig. 10.5), the spin-spin interaction is slightly different for the two states. This adds to the splitting represented by Equation 10.12, which is identical (53 MHz) for the two states. The net result is that four separate lines are seen in the  $\Lambda$  doublet transition of the ground state of OH, the  $^2\Pi_{3/2} J = 3/2$  state, as shown in figure 12.3.

The ratios of the line strengths, assuming LTE, is 1:5:9:1 for the transitions with frequencies of 1612:1665:1667:1720 MHz. Strong deviations from these ratios were an early indicator that maser action was occurring in interstellar OH. There is also hyperfine splitting in the higher levels of OH, as shown in figure 12.3.

Turning to case b molecules, CN is case  $b_{\beta_J}$ , with  $\vec{S}$  no longer parallel to  $\hat{k}$ , so the last term in equation 12.5 is more difficult to evaluate. The results are given by Townes and Schawlow, pg. 199.

From what we have said so far, it would appear that molecules without residual electronic angular momentum would have no hyperfine structure. However, as discussed in the next section, still finer effects are present.

## 12.3 Hyperfine Structure - Linear $^1\Sigma$ State

This section is a slightly modified version of a document written by Yancy Shirley.

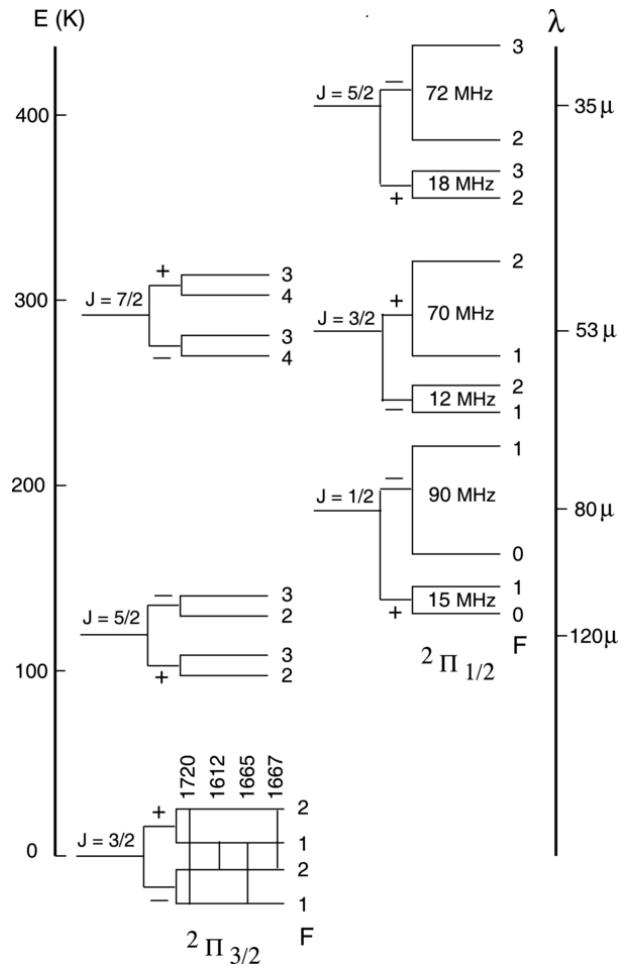


Figure 12.3: OH energy levels with hyperfine splitting.

### 12.3.1 Single Coupling Nucleus

Hyperfine structure is due to two separate interactions: electric quadrupole and magnetic dipole interactions. Typically, electric quadrupole coupling is much stronger for molecules in  ${}^1\Sigma$  states in contrast to the situation for states with residual electronic angular momentum, for which magnetic effects dominate (§12.2). As discussed in §12.1, for nuclei with nuclear spins,  $I$ , equal to 0 or  $1/2$ , the nuclear charge distribution is spherically symmetric and there is no electric quadrupole moment. Nuclei with nuclear spin greater than  $1/2$  have electric quadrupole moments. The nuclear electric quadrupole moment couples with the gradient of the electric field at the nucleus. Nuclei with a nuclear spin greater than 0 can have a dipole magnetic moment. Magnetic hyperfine coupling is the coupling between the nuclear magnetic moment and the magnetic field generated by molecular rotation. Table 12.1 and figure 12.1 list nuclear spins for nuclei commonly found in interstellar

molecules. Hyperfine structure is commonly observed in molecules that contain H ( $I = 1/2$ ), D ( $I = 1$ ),  $^{14}\text{N}$  ( $I = 1$ ), and  $^{17}\text{O}$  ( $I = 5/2$ ) nuclei (Fig. 12.1).

The hyperfine quantum number,  $F$ , is defined such that  $\vec{F} = \vec{I} + \vec{J}$  is the total molecular angular momentum, where  $\vec{I}$  and  $\vec{J}$  are the nuclear spin and rotational angular momentum of the molecule respectively ( $\vec{I}$  and  $\vec{J}$  precess around  $\vec{F}$ ; see Figure 12.4). Then,

$$\vec{F}^2 = (\vec{I} + \vec{J})^2 = \vec{I}^2 + \vec{J}^2 + 2\vec{I} \cdot \vec{J}, \quad (12.13)$$

$$F = J + I, J + I - 1, \dots, |J - I|. \quad (12.14)$$

Equation 12.14 is sometimes called a Clebsch-Gordon series.

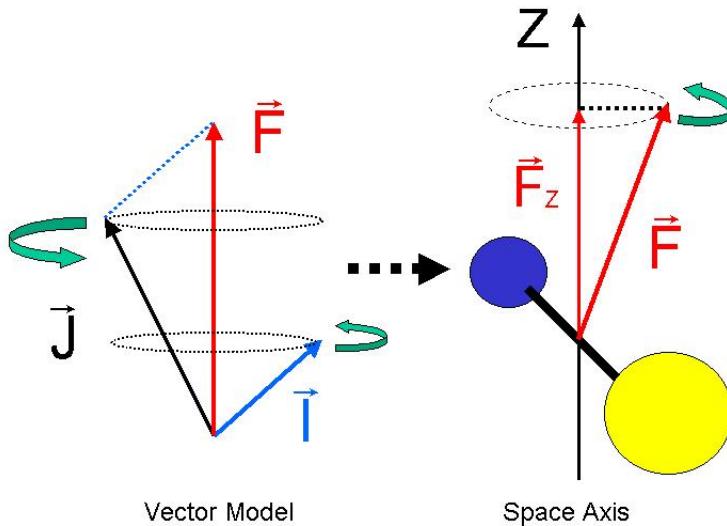


Figure 12.4: The vector model for the hyperfine structure of a linear molecule in a  ${}^1\Sigma$  state.

Selection rules stipulate that  $\Delta F = 0, \pm 1$  and  $\Delta J = \pm 1$ . Equation 12.4 determines which hyperfine levels exist and the selection rule determines which transitions exist. I shall discuss the hyperfine energy level structure for HCN ( $I(^{14}\text{N}) = 1$ ) and  $\text{C}^{17}\text{O}(I(^{17}\text{O}) = 5/2)$ . Almost all hyperfine structure in linear molecules in a  ${}^1\Sigma$  state can be understood from these two examples!

The  $^{14}\text{N}$  nucleus in HCN is responsible for strong hyperfine coupling (N.B. in Table 12.1,  $^{14}\text{N}$  has the strongest electric quadrupole moment by a factor of 4). The spin of  $^{14}\text{N}$  is  $I = 1$ . Equation 12.14 shows that the  $J = 0$  level remains unsplit ( $F = 1$ ), the  $J = 1$  level splits into a triplet ( $F = 2, 1, 0$ ), and the  $J = 2$  level also splits into a triplet ( $F = 3, 2, 1$ ). The selection rule allows 3 transitions from  $J = 1 \rightarrow 0$  and six hyperfine transitions from  $J = 2 \rightarrow 1$ . Since  $I = 1$  and the Clebsch-Gordon series (equation 12.14) terminates with  $F = |J - I|$ , all of the levels with  $J > 2$  will also be split into triplets. Figure 12.5 shows the energy levels for  $J = 0, 1$ , and  $2$ .

A good example of  $I = 5/2$  hyperfine structure is  $\text{C}^{17}\text{O}$ . Using equations 12.14, we see that the  $J = 0$  level remains unsplit ( $F = 5/2$ ), the  $J = 1$  level splits into a triplet ( $F = 7/2, 5/2, 3/2$ ), the  $J = 2$  level splits into 5 hyperfine levels ( $F = 9/2, 7/2, 5/2, 3/2, 1/2$ ), and the  $J = 3$  level into 6 hyperfine

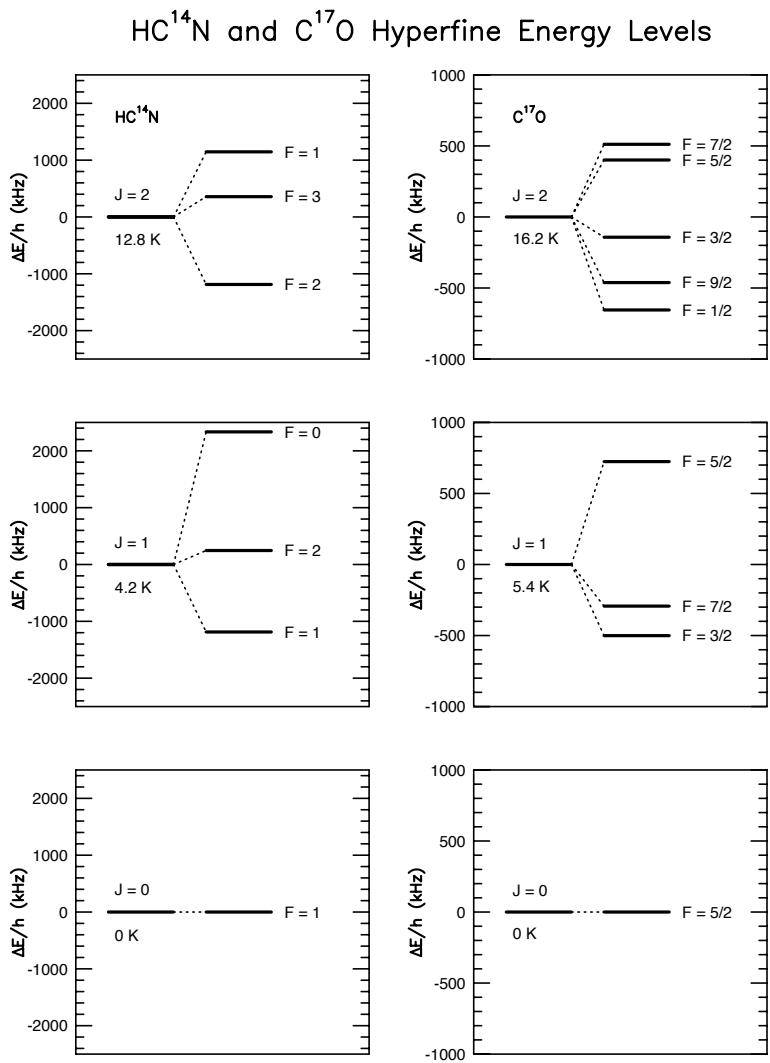


Figure 12.5: The hyperfine energy level structure of HCN and C<sup>17</sup>O for  $J = 0, 1$ , and  $2$ . The vertical axis is the shift in energy ( $\Delta E = E_Q + E_M$ ) in kHz. Note the vertical scale is different for HCN and C<sup>17</sup>O.

levels ( $F = 11/2, 9/2, 7/2, 5/2, 3/2, 1/2$ ). The selection rule allows three hyperfine transitions from  $J = 1 \rightarrow 0$ , nine hyperfine transitions from  $J = 2 \rightarrow 1$ , and 14 hyperfine transitions from  $J = 3 \rightarrow 2$ . All of the transitions for  $J = 2 \rightarrow 1$  are listed in Table 12.2 and the energy levels for  $J = 1$  and  $J = 2$  are shown in Figure 12.6. All levels with  $J \geq 3$  will be split into 6 hyperfine levels. Figure 12.6 shows the  $\text{C}^{17}\text{O}$  spectrum for  $J = 1 \rightarrow 0$  and  $J = 2 \rightarrow 1$ ,

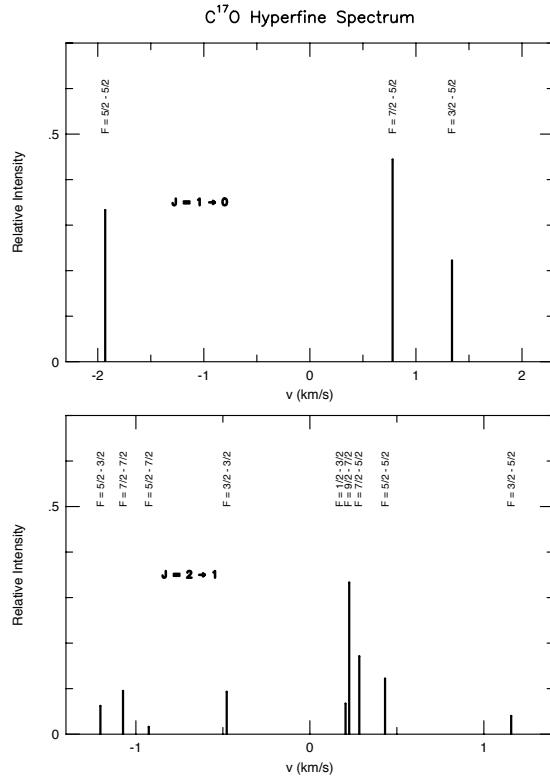


Figure 12.6: The hyperfine spectrum of  $\text{C}^{17}\text{O}$   $J = 2 \rightarrow 1$  and  $J = 1 \rightarrow 0$ . The vertical axis is relative intensity and the horizontal axis is velocity (km/s) in the rest frame of the molecule.

### 12.3.2 Frequency of Hyperfine Transitions and Relative Strength

The energy shifts due to electric quadrupole and magnetic hyperfine coupling are treated as perturbations of the rotational Hamiltonian. First I will treat the electric quadrupole coupling, then the magnetic hyperfine coupling. The following treatment is a distilled version of Ramsey (1953),

Townes & Schawlow 1975 (TS75), and Gordon & Cook 1984 (GC84).

To first order, we can write the perturbation due to an electric quadrupole moment as

$$\Delta E_Q = \langle I, J, F, M_F | \hat{H}_Q | I, J, F, M_F \rangle, \quad (12.15)$$

where we have assumed that  $J$  is a good quantum number (i.e. that  $\hat{H}_Q$  is diagonal). In reality,  $\hat{H}_Q$  is not exactly diagonal and  $J$  is perturbed by nuclear interactions. But, I will show later that second order effects are negligible). To solve Equation (11), we need to find an expression for the electric quadrupole Hamiltonian. Both TS75 and GC84 contain lucid derivations. The final solution for a linear diatomic molecule is given by

$$\hat{H}_Q = -eQq \frac{J \left[ 3(\vec{I} \cdot \vec{J})^2 + \frac{3}{2}\vec{I} \cdot \vec{J} - \vec{I}^2 \vec{J}^2 \right]}{2I(2I-1)(2J-1)J(2J+3)}, \quad (12.16)$$

where  $e$  is the charge,  $Q$  is the electric quadrupole moment, and  $q$  describes the magnitude of the component of the electric quadrupole tensor ( $q_{xx} = q_{yy} = -\frac{1}{2}q_{zz} = q$  because the field gradient at the nucleus is symmetric about the bond axis). By using Equation 12.16, we can solve for the energy perturbations by evaluating the eigenvalues for  $(\vec{I} \cdot \vec{J})^2$ ,  $\vec{I} \cdot \vec{J}$ , and  $\vec{I}^2 \vec{J}^2$ ,

$$\langle I, J, F | \vec{I} \cdot \vec{J} | I, J, F \rangle = \frac{1}{2} [F(F+1) - I(I+1) - J(J+1)] \equiv \frac{1}{2} C \quad (12.17)$$

$$\langle I, J, F | (\vec{I} \cdot \vec{J})^2 | I, J, F \rangle = \frac{1}{4} C^2 \quad (12.18)$$

$$\langle I, J, F | \vec{I}^2 \vec{J}^2 | I, J, F \rangle = I(I+1)J(J+1), \quad (12.19)$$

(N.B. even though the vector notation is used,  $\vec{I} \cdot \vec{J}$  etc. are operators). Plugging in the eigenvalues, we find

$$\Delta E_Q = -eQq \frac{\left[ \frac{3}{4}C(C+1) - I(I+1)J(J+1) \right]}{2I(2I-1)(2J-1)(2J+3)}, \quad (12.20)$$

where  $-eQq$  is the quadrupole coupling constant and the remaining expression on the right is Casimer's function (sometimes denoted  $Y(I, J, F)$ ).

As mentioned above, the Hamiltonian,  $\hat{H}_Q$ , is not exactly diagonal due to perturbations of the rotational angular momentum,  $J$ . As a result, second order perturbation theory could be used to obtain a more accurate energy shift. However, TS75 show that the magnitude of this effect is of the order  $\frac{eQq}{\nu}$ . Second order perturbations are typically 5 orders of magnitude smaller than electric quadrupole coupling and 3 orders of magnitude smaller than magnetic hyperfine interactions. First order perturbation theory is more than sufficient for the current resolution of submm spectrometers. Second order corrections can be found in TS75, Appendix 2.

As with nuclear electric quadrupole coupling, there is coupling between the nuclear magnetic moment and the magnetic field generated by molecular rotation. We can calculate the magnetic coupling energy with

$$\Delta E_M = \langle I, J, F, M_F | \hat{H}_M | I, J, F, M_F \rangle, \quad (12.21)$$

where  $\hat{H}_M = -\vec{\mu} \cdot \vec{H}_{eff}$  and  $\mu \sim \vec{I}$  is the nuclear magnetic moment and  $\vec{H}_{eff} \sim \vec{J}$  is the effective generated magnetic field. For a linear dipole molecule, the magnetic coupling Hamiltonian can be reduced to (again, see TS75 and GC84)

$$\hat{H}_M = C_I \vec{I} \cdot \vec{J}, \quad (12.22)$$

Table 12.1: Moments and Coupling Constants for Hyperfine Effects

Isotope	Spin I	Magnetic Moment <sup>a</sup>	Quadrupole Moment <sup>b</sup>	Molecule	eQq (MHz)	C <sub>I</sub> (kHz)	Ref
<sup>1</sup> H	$\frac{1}{2}$	+2.79284734(3)	...				
<sup>2</sup> D	1	+0.857438228(9)	+0.00286(2)	DCN	+0.1944	-0.6	1
<sup>12</sup> C	0	...	...				
<sup>13</sup> C	$\frac{1}{2}$	+0.7024118(14)	...				
<sup>14</sup> N	1	+40376100(6)	+0.0193(8)				
				NH <sub>3</sub>	-4.0842		1
				HCN	-4.7091	+10.4	1
				HC <sub>3</sub> N	-4.28		1
				CH <sub>3</sub> CN	-4.2244		1
<sup>15</sup> N	$\frac{1}{2}$	-0.28318884(5)	...				
<sup>16</sup> O	0	...	...				
<sup>17</sup> O	$\frac{5}{2}$	-1.89379(9)	-0.02578				
				C <sup>17</sup> O	+4.337	-30.4	2
				HC <sup>17</sup> O <sup>+</sup>	+4.595	-20	3
<sup>18</sup> O	0	...	...				
<sup>28</sup> Si	0	...	...				
<sup>32</sup> S	0	...	...				
<sup>33</sup> S	$\frac{3}{2}$	+0.6438212(14)	-0.064(10)				
				C <sup>33</sup> S	+12.83		1
<sup>34</sup> S	0	...	...				

a. Magnetic Moment in nuclear magnetons. 1  $\mu_B = 9.274 \times 10^{-21}$  erg·gauss<sup>-1</sup>.

b. Quadrupole Moment in barns. 1 barn =  $10^{24}$  cm<sup>2</sup>.

1. Gordy & Cook 1984; 2. Frerking & Langer 1981; 3. Dore et al. 2001

Table supplied by Yancy Shirley. Magnetic dipole and electric quadrupole data updated by NJE to values in Strong, Atomic Data and Nuclear Data Tables 90 (2005) 75.

Values in parentheses are uncertainties in the corresponding decimal places.

where  $C_I$  is the nuclear magnetic coupling constant (the only non-zero element of the nuclear magnetic coupling tensor,  $C_{xx} = C_{yy} = C_I$ ). Plugging in the eigenvalue of  $\vec{I} \cdot \vec{J}$  from Equation 12.17, we find that the energy perturbations due to magnetic coupling are

$$\Delta E_M = \frac{C_I}{2} [F(F+1) - I(I+1) - J(J+1)]. \quad (12.23)$$

The total frequency shifts relative to the unshifted frequency due to electric quadrupole and magnetic hyperfine coupling are given by

$$\nu_{if} - \nu_{unsplit} = \frac{\Delta E_i - \Delta E_f}{\hbar}, \quad (12.24)$$

where i is the initial F level, f is the final F level, and  $\Delta E = E_Q + E_M$ .

The relative strengths can be calculated by using irreducible tensor methods (see Gordy & Cook (1984) Chapter 15). We define the relative strength such that the sum of the relative strength, s, of all transitions from  $F' \rightarrow F$  for a given  $J' \rightarrow J$  are equal to one (see Rudolph 1968):

$$\sum_{F'F} s(IJ'F' \rightarrow IJF) = 1. \quad (12.25)$$

The relative line strengths are calculated in terms of a 6-j symbol,

$$s(IJ'F' \rightarrow IJF) = \frac{(2F+1)(2F'+1)}{(2I+1)} \left\{ \begin{array}{ccc} I & F' & J' \\ 1 & J & F \end{array} \right\}^2. \quad (12.26)$$

With the aid of  $6 - j$  Tables (Edmonds 1974 or Sobelman 1979), and the property that  $6 - j$  symbols are invariant with permutation of the columns, we find the appropriate  $6 - j$  symbol for each transition:

$$\left\{ \begin{array}{ccc} I & F & J \\ 1 & J-1 & F-1 \end{array} \right\}^2 = \frac{a(a+1)(a-2I-1)(a-2I)}{(2F-1)2F(2F+1)(2J-1)2J(2J+1)} \quad (12.27)$$

$$\left\{ \begin{array}{ccc} I & F & J \\ 1 & J-1 & F \end{array} \right\}^2 = \frac{2(a+1)(a-2I)(a-2F)(a-2J+1)}{2F(2F+1)(2F+2)(2J-1)2J(2J+1)} \quad (12.28)$$

$$\left\{ \begin{array}{ccc} I & F & J \\ 1 & J-1 & F+1 \end{array} \right\}^2 = \frac{(a-2F-1)(a-2F)(a-2J+1)(a-2J+2)}{(2F+1)(2F+2)(2F+3)(2J-1)2J(2J+1)}, \quad (12.29)$$

where  $a = F + J + I$ .

### 12.3.3 Multiple Coupling Nuclei

There are a few important linear molecules with multiple coupling nuclei. Important astrophysical examples include  $\text{N}_2\text{H}^+$ ,  $\text{N}_2\text{D}^+$ , and DCN. In general, there are two coupling schemes for multiple coupling nuclei. For hyperfine coupling when one nucleus has a much stronger coupling strength than the next nucleus (e.g.  $(eQq)_1 \gg (eQq)_2$ , etc.) we use the basis  $|JF_1F_2 \cdots F\rangle$  where

$$\begin{aligned} \vec{F}_1 &= \vec{I}_1 + \vec{J} \\ \vec{F}_2 &= \vec{I}_2 + \vec{F}_1 \\ &\vdots \\ \vec{F} &= \vec{I}_n + \vec{F}_{n-1} \end{aligned} \quad (12.30)$$

where  $I_i$  is the spin of the  $i^{th}$  nucleus and  $F_i = F_i + I_i, F_i + I_i - 1, \dots, |F_i - I_i|$  for each hyperfine quantum number. For hyperfine coupling when the nuclei have equal coupling strength we use the basis  $|J\mathcal{I}_1\mathcal{I}_2\mathcal{I}_3 \cdots \mathcal{I}F\rangle$  where

$$\begin{aligned} \vec{\mathcal{I}}_2 &= \vec{I}_1 + \vec{I}_2 \\ \vec{\mathcal{I}}_3 &= \vec{I}_2 + \vec{I}_3 \\ &\vdots \\ \vec{\mathcal{I}} &= \vec{I}_{n-1} + \vec{I}_n \\ \vec{F} &= \vec{\mathcal{I}} + \vec{J} \end{aligned} \quad (12.31)$$

Both  $\text{N}_2\text{H}^+$  and  $\text{N}_2\text{D}^+$  are examples of molecules with unequal coupling strengths and therefore use the first coupling scheme. The second coupling scheme is more important in molecules with higher degrees of symmetry (e.g.  $\text{D}_2\text{O}$  where the hyperfine coupling strengths of the two D nuclei are identical). I shall explain the observed hyperfine structure for  $\text{N}_2\text{H}^+$ .

The hyperfine energy levels for  $\text{N}_2\text{H}^+$  can be formed by using the Clebsch-Gordon series due to the outer nitrogen ( $F_1$ ) and due to the inner nitrogen ( $F$ ). The hyperfine coupling constants are:  $(eQq)_1 = -5.6902 \pm 0.0021$  MHz,  $(eQq)_2 = -1.3586 \pm 0.0038$  MHz,  $C_1 = 11.8 \pm 0.4$  kHz,  $C_2 = 8.7 \pm 0.6$  kHz (Caselli, Myers, & Thaddeus 1995). For the  $J=0$  level, the outer nitrogen results in only one energy level  $F_1 = 1$  (see Fig. 12.7). The inner nitrogen further splits the  $F_1 = 1$  level into

Table 12.2: Frequencies for Hyperfine Components of HCN and C<sup>17</sup>O

Molecule	$J_i$	$J_f$	$F_i$	$F_f$	$\Delta\nu_Q^a$ (kHz)	$\Delta\nu_M^a$ (kHz)	$\Delta v$ (km s <sup>-1</sup> )	$\nu$ (GHz)	Relative Intensity
HCN	1	0						88.631602	
			2	1	+235.4	+10.4	-0.83	88.631848	0.555
			1	1	-1177.3	-10.4	+4.02	88.630414	0.333
			0	1	+2354.5	-20.8	-7.90	88.633936	0.111
	2	1						177.261111	
			3	2	+100.9	+10.4	-0.19	177.261222	0.467
			2	2	-1412.7	-20.8	+2.43	177.259677	0.083
			2	1	0	0	0	177.261111	0.250
			1	2	+941.82	-41.6	-1.52	177.262011	0.005
			1	1	+2354.5	-20.8	-3.95	177.263445	0.083
			1	0	-1177.3	-10.4	+2.01	177.259923	0.111
C <sup>17</sup> O	1	0						112.359275	
			7/2	5/2	-216.8	-76.0	+0.78	112.358982	0.444
			5/2	5/2	+693.9	+30.4	-1.93	112.359999	0.333
			3/2	5/2	-607.2	+106.4	+1.34	112.358774	0.222
	2	1						224.714368	
			9/2	7/2	-92.9	-76.0	+0.22	224.714199	0.333
			7/2	7/2	+743.5	+60.8	-1.07	224.715172	0.095
			7/2	5/2	-167.3	-45.6	+0.28	224.714155	0.171
			5/2	7/2	+526.6	+167.2	-0.93	224.715062	0.016
			5/2	5/2	-384.1	+60.8	+0.43	224.714045	0.122
			5/2	3/2	+917.0	-15.2	-1.20	224.715270	0.062
			3/2	5/2	-1003.7	+136.8	+1.16	224.713501	0.040
			3/2	3/2	+297.4	+60.8	-0.48	224.714726	0.093
			1/2	3/2	-260.2	106.4	+0.20	224.714214	0.067

a. Quadrupole ( $\frac{E_Q}{\hbar}$ ) and magnetic ( $\frac{E_M}{\hbar}$ ) hyperfine frequency shifts with respect to the unshifted frequency.  
Table supplied by Yancy Shirley

three energy levels: F = 0, 1, 2. For the J = 1 level, the outer nitrogen coupling results in the triplet: F<sub>1</sub> = 0, 1, 2. The inner nitrogen splits the F<sub>1</sub> = 2 level into a triplet (F = 1, 2, 3), the F<sub>1</sub> = 1 level into a triplet (F = 0, 1, 2), and the F<sub>1</sub> = 0 level does not split and becomes F = 1. Thus, the J = 1 state is split into seven hyperfine levels and the J = 0 state is split into three energy levels when effects from both nuclei are included. If we continue to higher J levels, we will see that each level from J = 2 and higher is split into nine hyperfine levels.

Selection rules are similar to the single coupling case:  $\Delta F_1 = 0, \pm 1$ ,  $\Delta F = 0, \pm 1$  with  $0 \rightarrow 0$ , and  $\Delta J = \pm 1$ . Therefore, in the J = 1 → 0 transitions of N<sub>2</sub>H<sup>+</sup> and N<sub>2</sub>D<sup>+</sup>, there are fifteen allowed transitions. However, since the J = 0 splitting is very small, only seven transitions from the seven hyperfine J = 1 levels are observed (Figure 9). Recently the hyperfine structure of N<sub>2</sub>H<sup>+</sup> and N<sub>2</sub>D<sup>+</sup> up to the J = 3 → 2 transition has been calculated and observed (Gerin et al. 2001).

### 12.3.4 Multiple Coupling Nuclei - Energy Levels & Strength

Calculating the hyperfine energy levels for multiple coupling nuclei is more complicated. Irreducible tensor methods provide elegant and tractable techniques and are used in all modern papers (see The Journal of Molecular Spectroscopy, etc.). The hyperfine Hamiltonian for n coupling nuclei can be

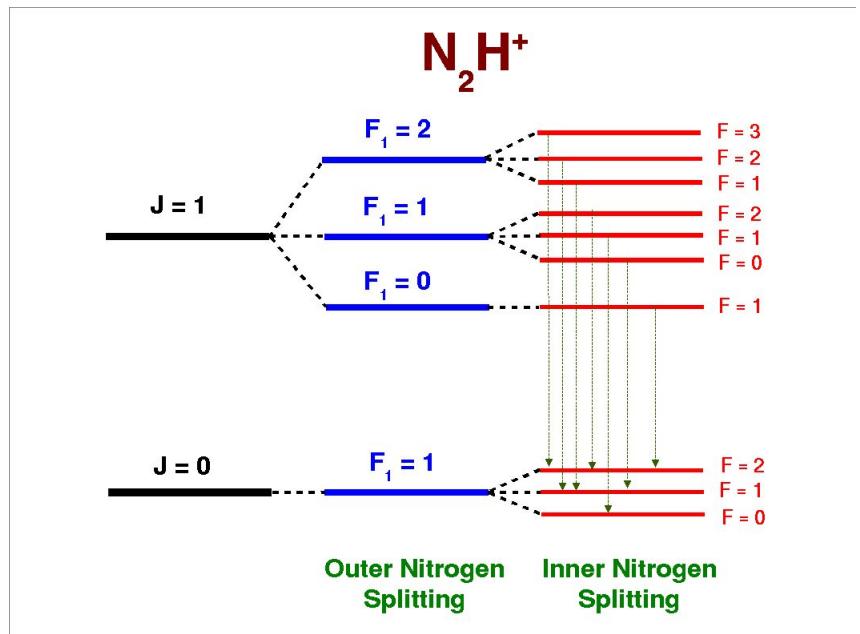


Figure 12.7: Diagrammatic representation of the hyperfine energy levels for  $\text{N}_2\text{H}^+$   $J=1 \rightarrow 0$ . The  $J=0$  level is not actually split into separate energy levels. The seven observed transitions from the seven upper hyperfine levels are shown.

expressed as

$$H_{hfs} = \sum_{i=1}^n \left[ \mathbf{V}^{(2)}(i) \cdot \mathbf{Q}^{(2)}(i) + \mathbf{m}^{(1)}(i) \cdot \mu^{(1)}(i) \right] + H_{spin-spin} \quad (12.32)$$

where  $\mathbf{V}^{(2)}(i)$  and  $\mathbf{Q}^{(2)}(i)$  are the second rank electric field gradient and electric quadrupole spherical tensors at the  $i^{th}$  nucleus,  $\mathbf{m}^{(1)}(i) \cdot \mu^{(1)}(i)$  is the magnetic dipole interaction in spherical tensor notation, and  $H_{spin-spin}$  is the hyperfine Hamiltonian due to spin-spin interactions between nuclei and is usually negligible for astrophysical spectroscopy (the interaction for 2 nuclei is given by  $H_{spin-spin} = -\sqrt{5/2}[\mathbf{D}_{21}^{(2)} \times \mathbf{I}_2^{(1)}]^{(1)} \cdot \mathbf{I}_1^{(1)}$ , where  $\mathbf{D}_{21}^{(2)}$  is the second rank spin coupling spherical tensor - see GC84 section 15.5). Expressions for the matrix elements of  $\mathbf{V}^{(2)}(i) \cdot \mathbf{Q}^{(2)}(i)$  and  $\mathbf{m}^{(1)}(i) \cdot \mu^{(1)}(i)$  for the two coupling schemes can be found in Chapter 15 of GC84 (e.g. equations 15.102, 15.103, 15.115, & 15.116). The matrix element for the electric quadrupole interaction and the magnetic dipole interaction for two unequal coupling nuclei in a linear molecule is given by (see GC84)

$$\begin{aligned} \Delta E_Q &= \frac{(-1)^t (eQq)_1 J}{2(2J+3) \left( \begin{array}{ccc} J & 2 & J \\ J & 0 & -J \end{array} \right) \left( \begin{array}{ccc} I_1 & 2 & I_1 \\ -I_1 & 0 & I_1 \end{array} \right)} \left\{ \begin{array}{ccc} F_1 & I_1 & J \\ 2 & J & I_1 \end{array} \right\} \\ &+ \frac{(-1)^r (eQq)_2 J (2F_1 + 1)}{2(2J+3) \left( \begin{array}{ccc} J & 2 & J \\ J & 0 & -J \end{array} \right) \left( \begin{array}{ccc} I_2 & 2 & I_2 \\ -I_2 & 0 & I_2 \end{array} \right)} \\ &\times \left\{ \begin{array}{ccc} I_2 & F_1 & J \\ 2 & J & F_1 \end{array} \right\} \left\{ \begin{array}{ccc} F & I_2 & F_1 \\ 2 & F_1 & I_2 \end{array} \right\} \end{aligned} \quad (12.33)$$

$$\begin{aligned} \Delta E_M &= (-1)^t C_1 [J(J+1)(2J+1)]^{1/2} [I_I(I_1+1)(2I_1+1)]^{1/2} \left\{ \begin{array}{ccc} F_1 & I_1 & J \\ 1 & J & I_1 \end{array} \right\} \\ &+ (-1)^{r+1} C_2 (2F_1 + 1) [J(J+1)(2J+1)]^{1/2} [I_2(I_2+1)(2I_2+1)]^{1/2} \\ &\times \left\{ \begin{array}{ccc} I_1 & F_1 & J \\ 1 & J & F_1 \end{array} \right\} \left\{ \begin{array}{ccc} F & I_2 & F_1 \\ 1 & F_1 & I_1 \end{array} \right\} \end{aligned} \quad (12.34)$$

where  $t = J + I_1 + F_1$  and  $r = J + I_1 + I_2 + 2F_1 + F$ . These equations represent the first order perturbation to the energy levels for unequal coupling nuclei. For equal or nearly equal coupling nuclei, a similar set of expressions based on the coupling scheme given in equation 12.32 and 12.33 is derived (see GC84). If we ignore the second term in  $\Delta E_Q$  and  $\Delta E_M$ , then we recover the same expressions derived §12.3.2 for a single coupling nucleus. We extend the method used to calculate the relative strengths to find the relative strength for two coupling nuclei (electric dipole transitions)

$$\begin{aligned} s(J'F'_1F' \rightarrow JF_1F) &= \frac{(2F'_1 + 1)(2F_1 + 1)(2F' + 1)(2F + 1)}{(2I_1 + 1)(2I_2 + 1)} \\ &\times \left\{ \begin{array}{ccc} I_1 & F'_1 & J' \\ 1 & J & F_1 \end{array} \right\}^2 \left\{ \begin{array}{ccc} I_2 & F' & F'_1 \\ 1 & F_1 & F \end{array} \right\}^2. \end{aligned} \quad (12.35)$$

The 3j and 6j symbols needed in equations (29 - 31) can be found in Edmonds (1974) or Sobelman (1979).

## 12.4 Hyperfine Structure in Non-linear Molecules

The basic techniques described so far can be extended to non-linear molecules. Two examples are worthy of note because of the importance of the hyperfine structure in constraining optical depths (recall §12.1). These are H<sub>2</sub>CO and NH<sub>3</sub>. Both have low frequency transitions caused by, respectively, K-type and inversion, doubling.

We will deal first with the simpler case, which is H<sub>2</sub>CO, with two spin- $\frac{1}{2}$  H nuclei. As discussed in chapter 8.5, these spins can combine into two states, with  $I = 0$  (para, B) or  $I = 1$  (ortho, A). Now, we will add the hyperfine effects of the two spins. The transition for which the hyperfine structure is important is the K-doublet transition between levels in the lowest state of ortho-H<sub>2</sub>CO,  $J_{K-1K+1} = 1_{10} \rightarrow 1_{11}$  (this transition is normally seen in absorption from the lower state ( $J_{K-1K+1} = 1_{11}$ ) to the higher state ( $J_{K-1K+1} = 1_{10}$ ) of the K-doublet, but we will use notation for emission to avoid confusion. This transition occurs at a frequency of 4829 MHz, or a wavelength around 6 cm.

With  $J = 1$  and  $I = 1$ , both states have allowed values of  $F = 0, 1, 2$  (see Eq. 12.14) and transitions are allowed to change  $F$  by zero or one with the usual prohibition that  $0 \not\rightarrow 0$ . There are thus six hyperfine transitions. Most of these are separated from the line center by less than 10 kHz and are thus blended in almost all sources, but the  $F = 1 \rightarrow 0$  component is split by -18.4 kHz (Tucker et al. 1971, Ap. J., 169, 429), corresponding to about 1 km s<sup>-1</sup>, enough to be clearly separate in very quiescent regions. Spectra showing the hyperfine splitting are shown in figure 12.8.

Ammonia (NH<sub>3</sub>) is considerably more complex, combining a quadrupole moment from the <sup>14</sup>N with three hydrogen spins. The inversion levels were discussed in §8.5, along with the division into A and E species, depending on the value of  $K$ . The relevant transition here is the lowest inversion transition, the  $(J, K) = (1, 1) \rightarrow (0, 0)$  transition at about 1.3 cm. In this state, the sum of the three hydrogen spins,  $I_H = 3/2$  to satisfy the Pauli exclusion principle. The nitrogen nuclear spin,  $I_N$  couples to  $J$  to form  $F_1$ , which then couples to  $I_H$  to form  $F$ . Since NH<sub>3</sub> is a <sup>1</sup> $\Sigma$  molecule, the nitrogen quadrupole moment produces the larger splitting, with the hydrogen magnetic moments providing smaller effects. The main source is the orientation of  $I_H$  with respect to  $J$  or  $F_1$ . There are also interactions with the nitrogen spin and with the rotation of the molecule. The rather complex combination of all of these effects is described by Townes and Schawlow, pp. 220-224. Here we simply indicate the net result with two figures from a classic paper, one (Fig. 12.9) showing the energy levels with the various sources of splitting and allowed transitions, and the second (Fig. 12.10) showing the astronomical spectrum toward a source with very low velocity dispersion. The splitting into a main group and inner and outer “satellite” lines with known, expected strengths provides an easy probe of optical depth.

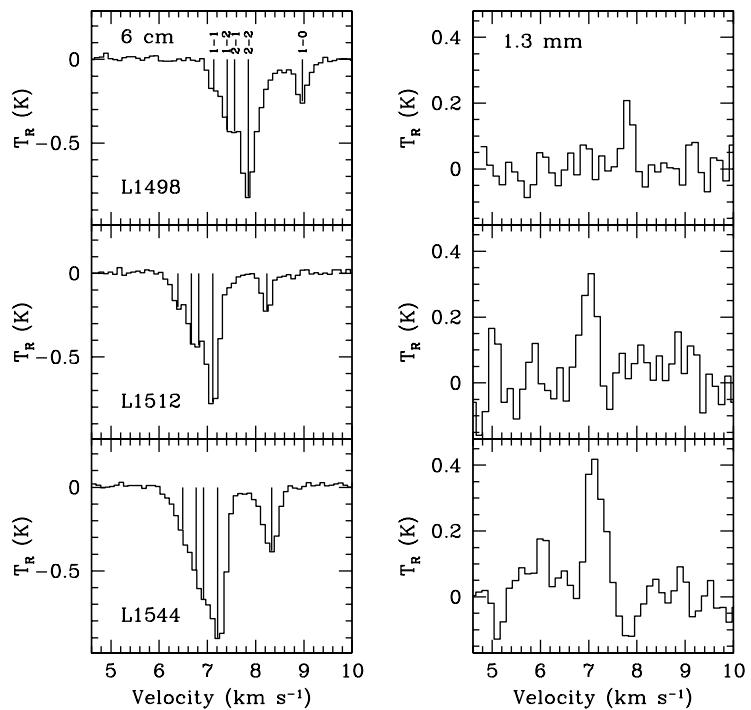


Figure 12.8: Observed spectra of the  $J_{K-1 K+1} = 1_{10} \rightarrow 1_{11}$  transition of  $\text{H}_2\text{CO}$  on the left, with hyperfine structure indicated, toward three starless, quiescent molecular cores, using the Arecibo telescope. The spectra on the right are the  $J_{K-1 K+1} = 3_{12} \rightarrow 2_{11}$  transition. No hyperfine structure is visible because of the much higher frequency (225 GHz). Figure from Young et al. (2004), Ap. J. 614, 252.

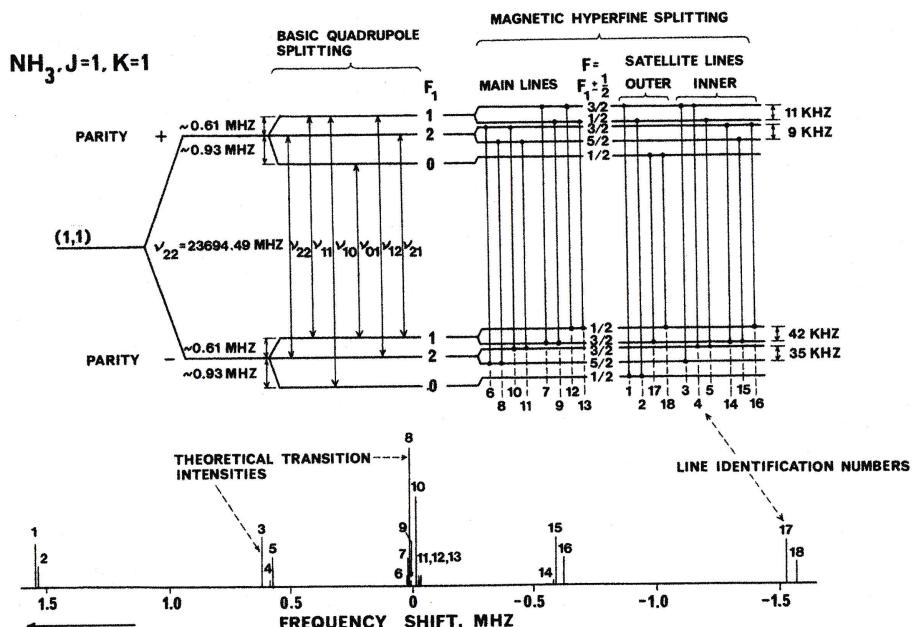


FIG. 1.—Sketch of the quadrupole and hyperfine energy levels for the  $(1,1)$  ammonia state with a spectral distribution profile. Note the many small components of the main line group and the conspicuously different splittings of the outer satellite groups.

Figure 12.9: Energy levels of the  $(J, K) = (1, 1) \rightarrow (0, 0)$  inversion levels of  $\text{NH}_3$ , with hyperfine structure due to the nitrogen and hydrogen nuclei. From Rydbeck et al. (1977) Ap. J. (Letters), 215, L35.

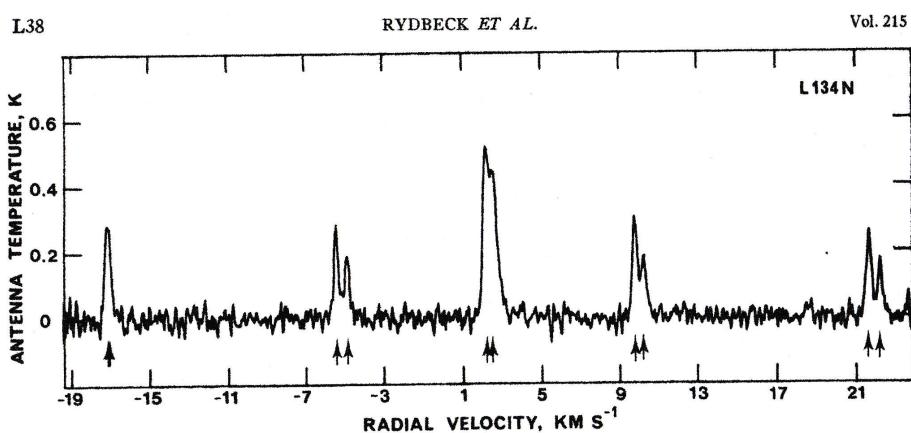


Figure 12.10: Spectrum of the  $(J, K) = (1, 1) \rightarrow (0, 0)$  inversion transition of  $\text{NH}_3$ , with hyperfine structure due to the nitrogen and hydrogen nuclei. From Rydbeck et al. (1977) Ap. J. (Letters), 215, L35.

# Appendix A

## Angular Momentum

### A.1 Clebsch-Gordon Coefficients

Suppose that we have two angular momenta that we can consider to be non-interacting, at least to zeroth order. Examples are the spin and orbital angular momenta of a single electron or the total angular momenta of two electrons. We represent these two angular momenta by the operators  $\vec{J}_1$  and  $\vec{J}_2$ , whose components commute: ( $[J_{1x}, J_{2y}] = 0$ , etc.). The eigenstates can be written as  $|j_1 m_1\rangle$  and  $|j_2 m_2\rangle$  where

$$\vec{J}_1^2 |j_1 m_1\rangle = j_1(j_1 + 1)\hbar^2 |j_1 m_1\rangle; \quad \vec{J}_2^2 |j_2 m_2\rangle = j_2(j_2 + 1)\hbar^2 |j_2 m_2\rangle \quad (\text{A.1})$$

Now we want to construct simultaneous eigenstates of  $\vec{J}_1$  and  $\vec{J}_2$ . These are given by the direct product:

$$|j_1 j_2 m_1 m_2\rangle = |j_1 m_1\rangle |j_2 m_2\rangle \quad (\text{A.2})$$

There will be  $(2j_1 + 1)(2j_2 + 1)$  of these.

Now consider the operators  $\vec{J} = \vec{J}_1 + \vec{J}_2$  and  $J_z$  (the  $z$  component of  $\vec{J}$ ). Because  $\vec{J}^2$ ,  $J_z$ ,  $\vec{J}_1^2$ , and  $\vec{J}_2^2$  all commute, there must be simultaneous eigenstates,  $|jm\rangle$ , such that

$$\vec{J}^2 |jm\rangle = j(j + 1)\hbar^2 |jm\rangle; \quad J_z |jm\rangle = m\hbar |jm\rangle \quad (\text{A.3})$$

How are these new  $|jm\rangle$  states related to the direct product states? The answer is the following unitary transformation:

$$|jm\rangle = \sum_{m_1 m_2} \langle j_1 j_2 m_1 m_2 | jm \rangle |j_1 m_1\rangle |j_2 m_2\rangle \quad (\text{A.4})$$

The coefficients,  $\langle j_1 j_2 m_1 m_2 | jm \rangle$ , are called Clebsch-Gordon coefficients. They tell us how to combine the vectors in the direct product representation to get a particular vector in the combined representation.

It is useful to work through the process for two spin  $\frac{1}{2}$  particles.

The Clebsch-Gordon coefficients are zero unless two conditions are met:

$$|j_1 - j_2| \leq j \leq |j_1 + j_2| \quad (\text{A.5})$$

and

$$m_1 + m_2 = m \quad \text{or} \quad m_1 + m_2 - m = 0 \quad (\text{A.6})$$

representing conservation of total and  $z$ -projected angular momenta. The second just says that the  $z$  components have to add up like scalars to the  $z$  component of the total. The first allows addition of angular momenta at different “angles” so that you get all the values between the limits that differ by integers. The expression A.5 is known as the *triangle inequality* and is often written as  $\Delta(j_1 j_2 j)$  as shorthand.

## A.2 3-J Symbols

It turns out to be easier to work with more symmetric coefficients called 3-J symbols. These are related to the Clebsch-Gordon coefficients by

$$\langle j_1 j_2 m_1 m_2 | jm \rangle = (-1)^{-j_1 + j_2 - m} \sqrt{2j+1} \begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & -m \end{pmatrix} \quad (\text{A.7})$$

The 3-J symbol,

$$\begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & m \end{pmatrix} = 0 \quad (\text{A.8})$$

unless  $\Delta(j_1 j_2 j)$  and  $m_1 + m_2 + m = 0$ . The condition on the  $m$  value explains why the minus sign appears in the definition A.7.

The 3-J symbols have the following nice symmetries:

1. An even transposition of columns has no effect on the value. This implies that  $\Delta(j_1 j_2 j) = \Delta(j j_1 j_2)$ , etc.
2. An odd transposition of columns introduced a phase factor,  $(-1)^{j_1 + j_2 + j}$ .

That is,

$$\begin{aligned} \begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & m \end{pmatrix} &= \begin{pmatrix} j_2 & j & j_1 \\ m_2 & m & m_1 \end{pmatrix} = \begin{pmatrix} j & j_1 & j_2 \\ m & m_1 & m_2 \end{pmatrix} \\ &= (-1)^{j_1 + j_2 + j} \begin{pmatrix} j_2 & j_1 & j \\ m_2 & m_1 & m \end{pmatrix} \end{aligned} \quad (\text{A.9})$$

We also have

$$\begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & m \end{pmatrix} = (-1)^{j_1 + j_2 + j} \begin{pmatrix} j_1 & j_2 & j \\ -m_1 & -m_2 & -m \end{pmatrix} \quad (\text{A.10})$$

The 3-J symbols also satisfy the following orthogonality relations:

$$\sum_{jm} (2j+1) \begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & m \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j \\ m'_1 & m'_2 & m \end{pmatrix} = \delta_{m_1 m'_1} \delta_{m_2 m'_2} \quad (\text{A.11})$$

and

$$\sum_{m_1 m_2} \begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & m \end{pmatrix} \begin{pmatrix} j_1 & j_2 & j' \\ m_1 & m_2 & m' \end{pmatrix} = \frac{1}{(2j+1)} \delta_{jj'} \delta_{mm'} \quad (\text{A.12})$$

The latter rule will be especially useful in evaluating matrix elements. A final, extremely useful result is that the integral of three spherical harmonics over all solid angles can be expressed in terms of 3-J symbols:

$$\int Y_{l_1 m_1}(\theta, \varphi) Y_{l_2 m_2}(\theta, \varphi) Y_{l_3 m_3}(\theta, \varphi) \sin(\theta) d\theta d\varphi = \sqrt{\frac{(2l_1+1)(2l_2+1)(2l_3+1)}{4\pi}} \begin{pmatrix} l_1 & l_2 & l_3 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l_1 & l_2 & l_3 \\ m_1 & m_2 & m_3 \end{pmatrix} \quad (\text{A.13})$$

Since  $Y_{lm}^* = (-1)^m Y_{l-m}$ , we can put A.13 into a form resembling the matrix element of one spherical harmonic between states that are represented by spherical harmonics.

### A.3 Behavior of Operators under Rotation: Spherical Tensors

Denote the rotation operator by  $U_R(\vec{\varphi})$ ; its action on a function of  $\vec{r}$ ,  $\psi(\vec{r})$ , is

$$U_R(\vec{\varphi})(\psi(\vec{r})) = \psi(\vec{r}'), \quad (\text{A.14})$$

where  $\vec{r}'$  is  $\vec{r}$  rotated by  $-\vec{\varphi}$ . For infinitesimal rotations,  $U_R = 1 - \varphi \frac{\partial}{\partial \varphi}$  but we can rewrite this in terms of the projection of the angular momentum operator on  $\vec{\varphi}$ ,  $J_{\hat{\varphi}} = i\hbar \frac{\partial}{\partial \varphi}$ :

$$U_R(\vec{\varphi}) = 1 + \frac{i}{\hbar} \vec{\varphi} \cdot \vec{J} \quad (\text{A.15})$$

Now suppose we have an operator  $A$ , and we want to know the matrix element of  $A$ , taken between rotated states,

$$\langle U_r \psi_1 | A | U_R \psi_2 \rangle = \langle \psi_1 | U_R^\dagger A U_R | \psi_2 \rangle = \quad (\text{A.16})$$

which defines the action of the rotation operator on  $A$ :

$$A' = U_R^\dagger A U_R \quad (\text{A.17})$$

For infinitesimal  $\vec{\varphi}$ ,

$$A' = A + \frac{i}{\hbar} [\vec{\varphi} \cdot \vec{J}, A] \quad (\text{A.18})$$

so the behavior of operators under rotation are defined by their commutation relations with  $\vec{J}$ . For example, a scalar operator commutes with  $\vec{J}$ :

$$[\vec{J}, A_s] = 0 \quad (\text{A.19})$$

which means that  $A'_s = A_s$ , while a vector operator does not:

$$[\vec{J}, \vec{A}_v] = \vec{J} \cdot \vec{A}_v \quad (\text{A.20})$$

which implies that  $\vec{A}'_v = \vec{A}_v + \vec{\varphi} \mathbf{x} \vec{A}_v$

More generally, we can classify operators by their behavior under rotation of the coordinate system. The object which generalizes scalars and vectors is a tensor, but the usual (Cartesian) definition of a tensor is unsuitable because the components of a given Cartesian tensor can be rearranged to form components that behave differently under rotation. For example, equation 2.57 was rearranged into a magnetic dipole and an electric quadrupole.

Therefore we define spherical (or irreducible) tensors. A spherical tensor of rank  $k$  is a set of  $2k+1$  entities that transform as do the spherical harmonics  $Y_{kq}$  under rotation of the coordinate system. So, we can represent an arbitrary spherical tensor  $T_{kq}$  by

$$T_{kq} = f(r) Y_{kq}(\theta, \varphi) \quad (\text{A.21})$$

where  $f(r)$  is an arbitrary function of  $r$  (only the magnitude, not the direction). The commutation relations for the spherical harmonics are

$$[J_z, Y_{kq}] = q\hbar Y_{kq} \quad (\text{A.22})$$

$$[J_{\pm}, Y_{kq}] = \hbar \sqrt{(k \mp q)(k \pm q + 1)} Y_{kq \pm 1} = \hbar \sqrt{(k(k+1) - q(q \pm 1)} Y_{kq \pm 1} \quad (\text{A.23})$$

where  $J_{\pm} = J_x \pm iJ_y$  are raising and lowering operators. So we know the commutation relations for all spherical tensors. To be specific,  $Y_{00}$  is a scalar,  $Y_{1m}$  is a vector,  $Y_{2m}$  is a spherical tensor of rank 2, etc.

## A.4 The Wigner-Eckart Theorem

Consider our tensor operator of rank  $k$  with components  $T_{kq}$ . We wish to know the matrix elements of this operator between states labeled by  $J$ , the eigenvalue of the total angular momentum, and by  $M$ , its projection on the  $z$  axis, and by some set of other quantum numbers, represented here by  $\gamma$ . Then we can label the states by  $|\gamma JM\rangle$ . Note that  $M$  and  $q$  depend on the orientation of the coordinate system, whereas the other quantum numbers do not. One may suspect that symmetry properties are sufficient to specify the dependence of the matrix element on  $M$  and  $q$ . This suspicion is confirmed by the Wigner-Eckart theorem:

$$\langle \gamma JM | T_{kq} | \gamma' J' M' \rangle = (-1)^{J-M} \langle \gamma J | |T_K| | \gamma' J' \rangle \begin{pmatrix} J & k & J' \\ -M & q & M' \end{pmatrix} \quad (\text{A.24})$$

The reduced matrix element  $\langle \gamma J | |T_k| | \gamma' J' \rangle$  does not depend on  $M, M'$ , or  $q$ . Together with the orthogonality relation for the 3-J symbols (A.12), this gives the following useful rule:

$$\sum_{MM'} |\langle \gamma JM | T_{kq} | \gamma' J' M' \rangle|^2 = \frac{1}{2k+1} |\langle \gamma J | |T_k| | \gamma' J' \rangle|^2 \quad (\text{A.25})$$

Furthermore, since the right hand side of equation A.25 does not depend on  $q$ ,

$$\sum_q \sum_{MM'} |\langle \gamma JM | T_{kq} | \gamma' J' M' \rangle|^2 = |\langle \gamma J | |T_k| | \gamma' J' \rangle|^2 \quad (\text{A.26})$$

These two last expressions depend on the 3-J symbol in A.24 not being zero. From the properties of 3-J symbols, expressed by the conditions following A.8, this implies that

$$|J - J'| \leq k \leq |J + J'| \quad (\text{A.27})$$

and

$$\Delta M = M - M' = q \quad (\text{A.28})$$

## A.5 The Dipole Moment Operator in Polarization Form

The dipole moment operator can be re-written in a particularly useful form for dealing with polarized light. This form is also more useful for deriving selection rules. We start with the standard form,

$$\vec{\mu} = -e\vec{r} = \mu_x \hat{x} + \mu_y \hat{y} + \mu_z \hat{z} \quad (\text{A.29})$$

which can be written in terms of spherical harmonics as

$$\vec{\mu} = \mu \sqrt{\frac{4\pi}{3}} \left[ \left( \frac{Y_{1-1} - Y_{11}}{\sqrt{2}} \right) \hat{x} + i \left( \frac{Y_{1-1} + Y_{11}}{\sqrt{2}} \right) \hat{y} + Y_{10} \hat{z} \right] \quad (\text{A.30})$$

where  $\mu$  is the magnitude of the electric dipole moment operator. We can now transform to axes given by  $\hat{e}_+$ ,  $\hat{e}_-$ , and  $\hat{z}$ , where  $\hat{e}_\pm$  represent opposite senses of circular polarization of light,

$$\hat{e}_\pm = \frac{\hat{x} \pm i\hat{y}}{\sqrt{2}} \quad (\text{A.31})$$

while  $\hat{z}$  is linear polarized light along the  $z$  axis. Then

$$\vec{\mu} = \mu [C_{1-1} \hat{e}_+ - C_{11} \hat{e}_- + C_{10} \hat{z}] \quad (\text{A.32})$$

where

$$C_{1m} = \sqrt{\frac{4\pi}{3}} Y_{1m} \quad (\text{A.33})$$

More generally,

$$C_{lm} = \sqrt{\frac{4\pi}{2l+1}} Y_{lm} \quad (\text{A.34})$$

are useful spherical functions and a pole of order  $k$  can be written in terms of  $C_{kq}$  (see equation 2.51).