

## Words to the reader about how to use this textbook

### I. What This Book Does and Does Not Contain

This text is intended for use by beginning graduate students and advanced upper division undergraduate students in all areas of chemistry.

It provides:

- (i) An introduction to the fundamentals of quantum mechanics as they apply to chemistry,
- (ii) Material that provides brief introductions to the subjects of molecular spectroscopy and chemical dynamics,
- (iii) An introduction to computational chemistry applied to the treatment of electronic structures of atoms, molecules, radicals, and ions,
- (iv) A large number of exercises, problems, and detailed solutions.

It does not provide much historical perspective on the development of quantum mechanics. Subjects such as the photoelectric effect, black-body radiation, the dual nature of electrons and photons, and the Davisson and Germer experiments are not even discussed.

To provide a text that students can use to gain introductory level knowledge of quantum mechanics as applied to chemistry problems, such a non-historical approach had to be followed. This text immediately exposes the reader to the machinery of quantum mechanics.

Sections 1 and 2 (i.e., Chapters 1-7), together with Appendices A, B, C and E, could constitute a one-semester course for most first-year Ph. D. programs in the U. S. A. Section 3 (Chapters 8-12) and selected material from other appendices or selections from Section 6 would be appropriate for a second-quarter or second-semester course. Chapters 13- 15 of Sections 4 and 5 would be of use for providing a link to a one-quarter or one-semester class covering molecular spectroscopy. Chapter 16 of Section 5 provides a brief introduction to chemical dynamics that could be used at the beginning of a class on this subject.

There are many quantum chemistry and quantum mechanics textbooks that cover material similar to that contained in Sections 1 and 2; in fact, our treatment of this material is generally briefer and less detailed than one finds in, for example, Quantum Chemistry, H. Eyring, J. Walter, and G. E. Kimball, J. Wiley and Sons, New York, N.Y. (1947), Quantum Chemistry, D. A. McQuarrie, University Science Books, Mill Valley, Ca. (1983), Molecular Quantum Mechanics, P. W. Atkins, Oxford Univ. Press, Oxford, England (1983), or Quantum Chemistry, I. N. Levine, Prentice Hall, Englewood Cliffs,

N. J. (1991), Depending on the backgrounds of the students, our coverage may have to be supplemented in these first two Sections.

By covering this introductory material in less detail, we are able, within the confines of a text that can be used for a one-year or a two-quarter course, to introduce the student to the more modern subjects treated in Sections 3, 5, and 6. Our coverage of modern quantum chemistry methodology is not as detailed as that found in Modern Quantum Chemistry, A. Szabo and N. S. Ostlund, Mc Graw-Hill, New York (1989), which contains little or none of the introductory material of our Sections 1 and 2.

By combining both introductory and modern up-to-date quantum chemistry material in a single book designed to serve as a text for one-quarter, one-semester, two-quarter, or one-year classes for first-year graduate students, we offer a unique product.

It is anticipated that a course dealing with atomic and molecular spectroscopy will follow the student's mastery of the material covered in Sections 1- 4. For this reason, beyond these introductory sections, this text's emphasis is placed on electronic structure applications rather than on vibrational and rotational energy levels, which are traditionally covered in considerable detail in spectroscopy courses.

In brief summary, this book includes the following material:

1. The Section entitled **The Basic Tools of Quantum Mechanics** treats the fundamental postulates of quantum mechanics and several applications to exactly soluble model problems. These problems include the conventional particle-in-a-box (in one and more dimensions), rigid-rotor, harmonic oscillator, and one-electron hydrogenic atomic orbitals. The concept of the Born-Oppenheimer separation of electronic and vibration-rotation motions is introduced here. Moreover, the vibrational and rotational energies, states, and wavefunctions of diatomic, linear polyatomic and non-linear polyatomic molecules are discussed here at an introductory level. This section also introduces the variational method and perturbation theory as tools that are used to deal with problems that can not be solved exactly.

2. The Section **Simple Molecular Orbital Theory** deals with atomic and molecular orbitals in a qualitative manner, including their symmetries, shapes, sizes, and energies. It introduces bonding, non-bonding, and antibonding orbitals, delocalized, hybrid, and Rydberg orbitals, and introduces Hückel-level models for the calculation of molecular orbitals as linear combinations of atomic orbitals (a more extensive treatment of

several semi-empirical methods is provided in Appendix F). This section also develops the Orbital Correlation Diagram concept that plays a central role in using Woodward-Hoffmann rules to predict whether chemical reactions encounter symmetry-imposed barriers.

### 3. The **Electronic Configurations, Term Symbols, and States**

Section treats the spatial, angular momentum, and spin symmetries of the many-electron wavefunctions that are formed as antisymmetrized products of atomic or molecular orbitals. Proper coupling of angular momenta (orbital and spin) is covered here, and atomic and molecular term symbols are treated. The need to include Configuration Interaction to achieve qualitatively correct descriptions of certain species' electronic structures is treated here. The role of the resultant Configuration Correlation Diagrams in the Woodward-Hoffmann theory of chemical reactivity is also developed.

4. The Section on **Molecular Rotation and Vibration** provides an introduction to how vibrational and rotational energy levels and wavefunctions are expressed for diatomic, linear polyatomic, and non-linear polyatomic molecules whose electronic energies are described by a single potential energy surface. Rotations of "rigid" molecules and harmonic vibrations of uncoupled normal modes constitute the starting point of such treatments.

5. The **Time Dependent Processes** Section uses time-dependent perturbation theory, combined with the classical electric and magnetic fields that arise due to the interaction of photons with the nuclei and electrons of a molecule, to derive expressions for the rates of transitions among atomic or molecular electronic, vibrational, and rotational states induced by photon absorption or emission. Sources of line broadening and time correlation function treatments of absorption lineshapes are briefly introduced. Finally, transitions induced by collisions rather than by electromagnetic fields are briefly treated to provide an introduction to the subject of theoretical chemical dynamics.

### 6. The Section on **More Quantitative Aspects of Electronic Structure Calculations**

introduces many of the computational chemistry methods that are used to quantitatively evaluate molecular orbital and configuration mixing amplitudes. The Hartree-Fock self-consistent field (SCF), configuration interaction (CI), multiconfigurational SCF (MCSCF), many-body and Møller-Plesset perturbation theories,

coupled-cluster (CC), and density functional or X<sup>n</sup>-like methods are included. The strengths and weaknesses of each of these techniques are discussed in some detail. Having mastered this section, the reader should be familiar with how potential energy hypersurfaces, molecular properties, forces on the individual atomic centers, and responses to externally applied fields or perturbations are evaluated on high speed computers.

## II. How to Use This Book: Other Sources of Information and Building Necessary Background

In most class room settings, the group of students learning quantum mechanics as it applies to chemistry have quite diverse backgrounds. In particular, the level of preparation in mathematics is likely to vary considerably from student to student, as will the exposure to symmetry and group theory. This text is organized in a manner that allows students to skip material that is already familiar while providing access to most if not all necessary background material. This is accomplished by dividing the material into sections, chapters and Appendices which fill in the background, provide methodological tools, and provide additional details.

The Appendices covering Point Group Symmetry and Mathematics Review are especially important to master. Neither of these two Appendices provides a first-principles treatment of their subject matter. The students are assumed to have fulfilled normal American Chemical Society mathematics requirements for a degree in chemistry, so only a review of the material especially relevant to quantum chemistry is given in the Mathematics Review Appendix. Likewise, the student is assumed to have learned or to be simultaneously learning about symmetry and group theory as applied to chemistry, so this subject is treated in a review and practical-application manner here. If group theory is to be included as an integral part of the class, then this text should be supplemented (e.g., by using the text Chemical Applications of Group Theory, F. A. Cotton, Interscience, New York, N. Y. (1963)).

The progression of sections leads the reader from the principles of quantum mechanics and several model problems which illustrate these principles and relate to chemical phenomena, through atomic and molecular orbitals, N-electron configurations, states, and term symbols, vibrational and rotational energy levels, photon-induced transitions among various levels, and eventually to computational techniques for treating chemical bonding and reactivity.

At the end of each Section, a set of **Review Exercises** and fully worked out answers are given. Attempting to work these exercises should allow the student to determine whether he or she needs to pursue additional background building via the **Appendices** .

In addition to the **Review Exercises** , sets of **Exercises and Problems**, and their solutions, are given at the end of each section.

The exercises are brief and highly focused on learning a particular skill. They allow the student to practice the mathematical steps and other material introduced in the section. The problems are more extensive and require that numerous steps be executed. They illustrate application of the material contained in the chapter to chemical phenomena and they help teach the relevance of this material to experimental chemistry. In many cases, new material is introduced in the problems, so all readers are encouraged to become actively involved in solving all problems.

To further assist the learning process, readers may find it useful to consult other textbooks or literature references. Several particular texts are recommended for additional reading, further details, or simply an alternative point of view. They include the following (in each case, the abbreviated name used in this text is given following the proper reference):

1. Quantum Chemistry, H. Eyring, J. Walter, and G. E. Kimball, J. Wiley and Sons, New York, N.Y. (1947)- EWK.
2. Quantum Chemistry, D. A. McQuarrie, University Science Books, Mill Valley, Ca. (1983)- McQuarrie.
3. Molecular Quantum Mechanics, P. W. Atkins, Oxford Univ. Press, Oxford, England (1983)- Atkins.
4. The Fundamental Principles of Quantum Mechanics, E. C. Kemble, McGraw-Hill, New York, N.Y. (1937)- Kemble.
5. The Theory of Atomic Spectra, E. U. Condon and G. H. Shortley, Cambridge Univ. Press, Cambridge, England (1963)- Condon and Shortley.
6. The Principles of Quantum Mechanics, P. A. M. Dirac, Oxford Univ. Press, Oxford, England (1947)- Dirac.
7. Molecular Vibrations, E. B. Wilson, J. C. Decius, and P. C. Cross, Dover Pub., New York, N. Y. (1955)- WDC.
8. Chemical Applications of Group Theory, F. A. Cotton, Interscience, New York, N. Y. (1963)- Cotton.
9. Angular Momentum, R. N. Zare, John Wiley and Sons, New York, N. Y. (1988)- Zare.

10. Introduction to Quantum Mechanics, L. Pauling and E. B. Wilson, Dover Publications, Inc., New York, N. Y. (1963)- Pauling and Wilson.
11. Modern Quantum Chemistry, A. Szabo and N. S. Ostlund, Mc Graw-Hill, New York (1989)- Szabo and Ostlund.
12. Quantum Chemistry, I. N. Levine, Prentice Hall, Englewood Cliffs, N. J. (1991)- Levine.
13. Energetic Principles of Chemical Reactions, J. Simons, Jones and Bartlett, Portola Valley, Calif. (1983),

# Section 1 The Basic Tools of Quantum Mechanics

## Chapter 1

*Quantum Mechanics Describes Matter in Terms of Wavefunctions and Energy Levels.  
Physical Measurements are Described in Terms of Operators Acting on Wavefunctions*

### I. Operators, Wavefunctions, and the Schrödinger Equation

The trends in chemical and physical properties of the elements described beautifully in the periodic table and the ability of early spectroscopists to fit atomic line spectra by simple mathematical formulas and to interpret atomic electronic states in terms of empirical quantum numbers provide compelling evidence that some relatively simple framework must exist for understanding the electronic structures of all atoms. The great predictive power of the concept of atomic valence further suggests that molecular electronic structure should be understandable in terms of those of the constituent atoms.

Much of quantum chemistry attempts to make more quantitative these aspects of chemists' view of the periodic table and of atomic valence and structure. By starting from 'first principles' and treating atomic and molecular states as solutions of a so-called Schrödinger equation, quantum chemistry seeks to determine what underlies the empirical quantum numbers, orbitals, the *aufbau* principle and the concept of valence used by spectroscopists and chemists, in some cases, even prior to the advent of quantum mechanics.

Quantum mechanics is cast in a language that is not familiar to most students of chemistry who are examining the subject for the first time. Its mathematical content and how it relates to experimental measurements both require a great deal of effort to master. With these thoughts in mind, the authors have organized this introductory section in a manner that first provides the student with a brief introduction to the two primary constructs of quantum mechanics, operators and wavefunctions that obey a Schrödinger equation, then demonstrates the application of these constructs to several chemically relevant model problems, and finally returns to examine in more detail the conceptual structure of quantum mechanics.

By learning the solutions of the Schrödinger equation for a few model systems, the student can better appreciate the treatment of the fundamental postulates of quantum mechanics as well as their relation to experimental measurement because the wavefunctions of the known model problems can be used to illustrate.

## A. Operators

*Each physically measurable quantity has a corresponding operator. The eigenvalues of the operator tell the values of the corresponding physical property that can be observed*

In quantum mechanics, any experimentally measurable physical quantity  $F$  (e.g., energy, dipole moment, orbital angular momentum, spin angular momentum, linear momentum, kinetic energy) whose classical mechanical expression can be written in terms of the cartesian positions  $\{q_i\}$  and momenta  $\{p_i\}$  of the particles that comprise the system of interest is assigned a corresponding quantum mechanical operator  $\mathbf{F}$ . Given  $F$  in terms of the  $\{q_i\}$  and  $\{p_i\}$ ,  $\mathbf{F}$  is formed by replacing  $p_j$  by  $-i\hbar / q_j$  and leaving  $q_j$  untouched.

For example, if

$$F = \sum_{i=1,N} (p_i^2/2m_i + 1/2 k(q_i - q_i^0)^2 + L(q_i - q_i^0)),$$

then

$$\mathbf{F} = \sum_{i=1,N} (-\hbar^2/2m_i \nabla_i^2 + 1/2 k(q_i - q_i^0)^2 + L(q_i - q_i^0))$$

is the corresponding quantum mechanical operator. Such an operator would occur when, for example, one describes the sum of the kinetic energies of a collection of particles (the  $\sum_{i=1,N} (p_i^2/2m_i)$  term, plus the sum of "Hookes' Law" parabolic potentials (the  $1/2 \sum_{i=1,N} k(q_i - q_i^0)^2$ ), and (the last term in  $F$ ) the interactions of the particles with an externally applied field whose potential energy varies linearly as the particles move away from their equilibrium positions  $\{q_i^0\}$ .

The sum of the z-components of angular momenta of a collection of  $N$  particles has

$$F = \sum_{j=1,N} (x_j p_{yj} - y_j p_{xj}),$$

and the corresponding operator is

$$\mathbf{F} = -i\hbar \sum_{j=1,N} (x_j \nabla_j - y_j \nabla_j).$$

The x-component of the dipole moment for a collection of  $N$  particles

has

$$F = \sum_{j=1, N} Z_j e x_j, \text{ and}$$

$$\mathbf{F} = \sum_{j=1, N} Z_j e \mathbf{x}_j ,$$

where  $Z_j e$  is the charge on the  $j^{\text{th}}$  particle.

The mapping from  $F$  to  $\mathbf{F}$  is straightforward only in terms of cartesian coordinates. To map a classical function  $F$ , given in terms of curvilinear coordinates (even if they are orthogonal), into its quantum operator is not at all straightforward. Interested readers are referred to Kemble's text on quantum mechanics which deals with this matter in detail. The mapping can always be done in terms of cartesian coordinates after which a transformation of the resulting coordinates and differential operators to a curvilinear system can be performed. The corresponding transformation of the kinetic energy operator to spherical coordinates is treated in detail in Appendix A. The text by EWK also covers this topic in considerable detail.

The relationship of these quantum mechanical operators to experimental measurement will be made clear later in this chapter. For now, suffice it to say that these operators define equations whose solutions determine the values of the corresponding physical property that can be observed when a measurement is carried out; only the values so determined can be observed. This should suggest the origins of quantum mechanics' prediction that some measurements will produce **discrete** or **quantized** values of certain variables (e.g., energy, angular momentum, etc.).

## B. Wavefunctions

*The eigenfunctions of a quantum mechanical operator depend on the coordinates upon which the operator acts; these functions are called wavefunctions*

In addition to operators corresponding to each physically measurable quantity, quantum mechanics describes the state of the system in terms of a wavefunction that is a function of the coordinates  $\{q_j\}$  and of time  $t$ . The function  $|\psi(q_j, t)|^2 = \psi^* \psi$  gives the probability density for observing the coordinates at the values  $q_j$  at time  $t$ . For a many-particle system such as the  $\text{H}_2\text{O}$  molecule, the wavefunction depends on many coordinates. For the  $\text{H}_2\text{O}$  example, it depends on the  $x$ ,  $y$ , and  $z$  (or  $r$ ,  $\theta$ , and  $\phi$ ) coordinates of the ten

electrons and the x, y, and z (or r,  $\theta$ , and  $\phi$ ) coordinates of the oxygen nucleus and of the two protons; a total of thirty-nine coordinates appear in  $\Psi$ .

In classical mechanics, the coordinates  $q_j$  and their corresponding momenta  $p_j$  are functions of time. The state of the system is then described by specifying  $q_j(t)$  and  $p_j(t)$ . In quantum mechanics, the concept that  $q_j$  is known as a function of time is replaced by the concept of the probability density for finding  $q_j$  at a particular value at a particular time  $t$ :  $|\Psi(q_j, t)|^2$ . Knowledge of the corresponding momenta as functions of time is also relinquished in quantum mechanics; again, only knowledge of the probability density for finding  $p_j$  with any particular value at a particular time  $t$  remains.

### C. The Schrödinger Equation

*This equation is an eigenvalue equation for the energy or Hamiltonian operator; its eigenvalues provide the energy levels of the system*

#### 1. The Time-Dependent Equation

*If the Hamiltonian operator contains the time variable explicitly, one must solve the time-dependent Schrödinger equation*

How to extract from  $\Psi(q_j, t)$  knowledge about momenta is treated below in Sec. III. A, where the structure of quantum mechanics, the use of operators and wavefunctions to make predictions and interpretations about experimental measurements, and the origin of 'uncertainty relations' such as the well known Heisenberg uncertainty condition dealing with measurements of coordinates and momenta are also treated.

Before moving deeper into understanding what quantum mechanics 'means', it is useful to learn how the wavefunctions  $\Psi$  are found by applying the basic equation of quantum mechanics, the Schrödinger equation, to a few exactly soluble model problems. Knowing the solutions to these 'easy' yet chemically very relevant models will then facilitate learning more of the details about the structure of quantum mechanics because these model cases can be used as 'concrete examples'.

The Schrödinger equation is a differential equation depending on time and on all of the spatial coordinates necessary to describe the system at hand (thirty-nine for the H<sub>2</sub>O example cited above). It is usually written

$$\mathbf{H} \Psi = i \hbar \frac{\partial \Psi}{\partial t}$$

where  $\psi(\mathbf{q},t)$  is the unknown wavefunction and  $\mathbf{H}$  is the operator corresponding to the total energy physical property of the system. This operator is called the Hamiltonian and is formed, as stated above, by first writing down the classical mechanical expression for the total energy (kinetic plus potential) in cartesian coordinates and momenta and then replacing all classical momenta  $\mathbf{p}_j$  by their quantum mechanical operators  $\mathbf{p}_j = -i\hbar \nabla / \mathbf{q}_j$ .

For the H<sub>2</sub>O example used above, the classical mechanical energy of all thirteen particles is

$$E = \sum_i \left\{ \frac{p_i^2}{2m_e} + \frac{1}{2} \sum_j e^2/r_{i,j} - \sum_a Z_a e^2/r_{i,a} \right\} + \sum_a \left\{ \frac{p_a^2}{2m_a} + \frac{1}{2} \sum_b Z_a Z_b e^2/r_{a,b} \right\},$$

where the indices  $i$  and  $j$  are used to label the ten electrons whose thirty cartesian coordinates are  $\{q_i\}$  and  $a$  and  $b$  label the three nuclei whose charges are denoted  $\{Z_a\}$ , and whose nine cartesian coordinates are  $\{q_a\}$ . The electron and nuclear masses are denoted  $m_e$  and  $\{m_a\}$ , respectively.

The corresponding Hamiltonian operator is

$$\mathbf{H} = \sum_i \left\{ -(\hbar^2/2m_e) \nabla_i^2 + \frac{1}{2} \sum_j e^2/r_{i,j} - \sum_a Z_a e^2/r_{i,a} \right\} + \sum_a \left\{ -(\hbar^2/2m_a) \nabla_a^2 + \frac{1}{2} \sum_b Z_a Z_b e^2/r_{a,b} \right\}.$$

Notice that  $\mathbf{H}$  is a second order differential operator in the space of the thirty-nine cartesian coordinates that describe the positions of the ten electrons and three nuclei. It is a second order operator because the momenta appear in the kinetic energy as  $p_j^2$  and  $p_a^2$ , and the quantum mechanical operator for each momentum  $\mathbf{p} = -i\hbar \nabla / \mathbf{q}$  is of first order.

The Schrödinger equation for the H<sub>2</sub>O example at hand then reads

$$\sum_i \left\{ -(\hbar^2/2m_e) \nabla_i^2 + \frac{1}{2} \sum_j e^2/r_{i,j} - \sum_a Z_a e^2/r_{i,a} \right\} + \sum_a \left\{ -(\hbar^2/2m_a) \nabla_a^2 + \frac{1}{2} \sum_b Z_a Z_b e^2/r_{a,b} \right\} \psi = i\hbar \frac{\partial}{\partial t} \psi.$$

## 2. The Time-Independent Equation

*If the Hamiltonian operator does not contain the time variable explicitly, one can solve the time-independent Schrödinger equation*

In cases where the classical energy, and hence the quantum Hamiltonian, do not contain terms that are explicitly time dependent (e.g., interactions with time varying external electric or magnetic fields would add to the above classical energy expression time dependent terms discussed later in this text), the separations of variables techniques can be used to reduce the Schrödinger equation to a time-independent equation.

In such cases,  $\mathbf{H}$  is not explicitly time dependent, so one can assume that  $\psi(\mathbf{q}_j, t)$  is of the form

$$\psi(\mathbf{q}_j, t) = \phi(\mathbf{q}_j) F(t).$$

Substituting this 'ansatz' into the time-dependent Schrödinger equation gives

$$\phi(\mathbf{q}_j) i \hbar \frac{dF}{dt} = \mathbf{H} \phi(\mathbf{q}_j) F(t).$$

Dividing by  $\phi(\mathbf{q}_j) F(t)$  then gives

$$F^{-1} i \hbar \frac{dF}{dt} = \mathbf{H} \phi(\mathbf{q}_j).$$

Since  $F(t)$  is only a function of time  $t$ , and  $\phi(\mathbf{q}_j)$  is only a function of the spatial coordinates  $\{\mathbf{q}_j\}$ , and because the left hand and right hand sides must be equal for all values of  $t$  and of  $\{\mathbf{q}_j\}$ , both the left and right hand sides must equal a constant. If this constant is called  $E$ , the two equations that are embodied in this separated Schrödinger equation read as follows:

$$\mathbf{H} \phi(\mathbf{q}_j) = E \phi(\mathbf{q}_j),$$

$$i \hbar \frac{dF(t)}{dt} = E F(t).$$

The first of these equations is called the time-independent Schrödinger equation; it is a so-called eigenvalue equation in which one is asked to find functions that yield a constant multiple of themselves when acted on by the Hamiltonian operator. Such functions are called eigenfunctions of  $\mathbf{H}$  and the corresponding constants are called eigenvalues of  $\mathbf{H}$ .

For example, if  $\mathbf{H}$  were of the form  $-\hbar^2/2M \nabla^2 = \mathbf{H}$ , then functions of the form  $\exp(i \mathbf{m} \cdot \mathbf{r})$  would be eigenfunctions because

$$\{-\hbar^2/2M \nabla^2\} \exp(i \mathbf{m} \cdot \mathbf{r}) = \{m^2 \hbar^2/2M\} \exp(i \mathbf{m} \cdot \mathbf{r}).$$

In this case,  $\{m^2 \hbar^2/2M\}$  is the eigenvalue.

When the Schrödinger equation can be separated to generate a time-independent equation describing the spatial coordinate dependence of the wavefunction, the eigenvalue  $E$  must be returned to the equation determining  $F(t)$  to find the time dependent part of the wavefunction. By solving

$$i\hbar dF(t)/dt = E F(t)$$

once  $E$  is known, one obtains

$$F(t) = \exp(-i Et/\hbar),$$

and the full wavefunction can be written as

$$\psi(\mathbf{q}_j, t) = \psi(\mathbf{q}_j) \exp(-i Et/\hbar).$$

For the above example, the time dependence is expressed by

$$F(t) = \exp(-i t \{m^2 \hbar^2/2M\}/\hbar).$$

Having been introduced to the concepts of operators, wavefunctions, the Hamiltonian and its Schrödinger equation, it is important to now consider several examples of the applications of these concepts. The examples treated below were chosen to provide the learner with valuable experience in solving the Schrödinger equation; they were also chosen because the models they embody form the most elementary chemical models of electronic motions in conjugated molecules and in atoms, rotations of linear molecules, and vibrations of chemical bonds.

## II. Examples of Solving the Schrödinger Equation

### A. Free-Particle Motion in Two Dimensions

*The number of dimensions depends on the number of particles and the number of spatial (and other) dimensions needed to characterize the position and motion of each particle*

## 1. The Schrödinger Equation

Consider an electron of mass  $m$  and charge  $e$  moving on a two-dimensional surface that defines the  $x,y$  plane (perhaps the electron is constrained to the surface of a solid by a potential that binds it tightly to a narrow region in the  $z$ -direction), and assume that the electron experiences a constant potential  $V_0$  at all points in this plane (on any real atomic or molecular surface, the electron would experience a potential that varies with position in a manner that reflects the periodic structure of the surface). The pertinent time independent Schrödinger equation is:

$$-\frac{\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \psi(x,y) + V_0 \psi(x,y) = E \psi(x,y).$$

Because there are no terms in this equation that couple motion in the  $x$  and  $y$  directions (e.g., no terms of the form  $x^a y^b$  or  $\frac{\partial}{\partial x} \frac{\partial}{\partial y}$  or  $\frac{\partial}{\partial x} \frac{\partial}{\partial y}$ ), separation of variables can be used to write  $\psi$  as a product  $\psi(x,y) = A(x)B(y)$ . Substitution of this form into the Schrödinger equation, followed by collecting together all  $x$ -dependent and all  $y$ -dependent terms, gives;

$$-\frac{\hbar^2}{2m} A^{-1} \frac{\partial^2 A}{\partial x^2} - \frac{\hbar^2}{2m} B^{-1} \frac{\partial^2 B}{\partial y^2} = E - V_0.$$

Since the first term contains no  $y$ -dependence and the second contains no  $x$ -dependence, both must actually be constant (these two constants are denoted  $E_x$  and  $E_y$ , respectively), which allows two separate Schrödinger equations to be written:

$$-\frac{\hbar^2}{2m} A^{-1} \frac{\partial^2 A}{\partial x^2} = E_x, \text{ and}$$

$$-\frac{\hbar^2}{2m} B^{-1} \frac{\partial^2 B}{\partial y^2} = E_y.$$

The total energy  $E$  can then be expressed in terms of these separate energies  $E_x$  and  $E_y$  as  $E_x + E_y = E - V_0$ . Solutions to the  $x$ - and  $y$ - Schrödinger equations are easily seen to be:

$$A(x) = \exp(ix(2mE_x/\hbar^2)^{1/2}) \text{ and } \exp(-ix(2mE_x/\hbar^2)^{1/2}),$$

$$B(y) = \exp(iy(2mE_y/\hbar^2)^{1/2}) \text{ and } \exp(-iy(2mE_y/\hbar^2)^{1/2}).$$

Two independent solutions are obtained for each equation because the x- and y-space Schrödinger equations are both second order differential equations.

## 2. Boundary Conditions

*The boundary conditions, not the Schrödinger equation, determine whether the eigenvalues will be discrete or continuous*

If the electron is entirely unconstrained within the x,y plane, the energies  $E_x$  and  $E_y$  can assume any value; this means that the experimenter can 'inject' the electron onto the x,y plane with any total energy  $E$  and any components  $E_x$  and  $E_y$  along the two axes as long as  $E_x + E_y = E$ . In such a situation, one speaks of the energies along both coordinates as being 'in the continuum' or 'not quantized'.

In contrast, if the electron is constrained to remain within a fixed area in the x,y plane (e.g., a rectangular or circular region), then the situation is qualitatively different. Constraining the electron to any such specified area gives rise to so-called boundary conditions that impose additional requirements on the above A and B functions. These constraints can arise, for example, if the potential  $V_0(x,y)$  becomes very large for x,y values outside the region, in which case, the probability of finding the electron outside the region is very small. Such a case might represent, for example, a situation in which the molecular structure of the solid surface changes outside the enclosed region in a way that is highly repulsive to the electron.

For example, if motion is constrained to take place within a rectangular region defined by  $0 \leq x \leq L_x$ ;  $0 \leq y \leq L_y$ , then the continuity property that all wavefunctions must obey (because of their interpretation as probability densities, which must be continuous) causes A(x) to vanish at 0 and at  $L_x$ . Likewise, B(y) must vanish at 0 and at  $L_y$ . To implement these constraints for A(x), one must linearly combine the above two solutions  $\exp(ix(2mE_x/\hbar^2)^{1/2})$  and  $\exp(-ix(2mE_x/\hbar^2)^{1/2})$  to achieve a function that vanishes at  $x=0$ :

$$A(x) = \exp(ix(2mE_x/\hbar^2)^{1/2}) - \exp(-ix(2mE_x/\hbar^2)^{1/2}).$$

One is allowed to linearly combine solutions of the Schrödinger equation that have the same energy (i.e., are degenerate) because Schrödinger equations are linear differential

equations. An analogous process must be applied to B(y) to achieve a function that vanishes at y=0:

$$B(y) = \exp(iy(2mE_y/\hbar^2)^{1/2}) - \exp(-iy(2mE_y/\hbar^2)^{1/2}).$$

Further requiring A(x) and B(y) to vanish, respectively, at x=L<sub>x</sub> and y=L<sub>y</sub>, gives equations that can be obeyed only if E<sub>x</sub> and E<sub>y</sub> assume particular values:

$$\exp(iL_x(2mE_x/\hbar^2)^{1/2}) - \exp(-iL_x(2mE_x/\hbar^2)^{1/2}) = 0, \text{ and}$$

$$\exp(iL_y(2mE_y/\hbar^2)^{1/2}) - \exp(-iL_y(2mE_y/\hbar^2)^{1/2}) = 0.$$

These equations are equivalent to

$$\sin(L_x(2mE_x/\hbar^2)^{1/2}) = \sin(L_y(2mE_y/\hbar^2)^{1/2}) = 0.$$

Knowing that sin( ) vanishes at =n , for n=1,2,3,..., (although the sin(n ) function vanishes for n=0, this function vanishes for all x or y, and is therefore unacceptable because it represents zero probability density at all points in space) one concludes that the energies E<sub>x</sub> and E<sub>y</sub> can assume only values that obey:

$$L_x(2mE_x/\hbar^2)^{1/2} = n_x ,$$

$$L_y(2mE_y/\hbar^2)^{1/2} = n_y , \text{ or}$$

$$E_x = n_x^2 \cdot \hbar^2/(2mL_x^2), \text{ and}$$

$$E_y = n_y^2 \cdot \hbar^2/(2mL_y^2), \text{ with } n_x \text{ and } n_y = 1, 2, 3, \dots$$

It is important to stress that it is the imposition of boundary conditions, expressing the fact that the electron is spatially constrained, that gives rise to quantized energies. In the absence of spatial confinement, or with confinement only at x =0 or L<sub>x</sub> or only at y =0 or L<sub>y</sub>, quantized energies would not be realized.

In this example, confinement of the electron to a finite interval along both the x and y coordinates yields energies that are quantized along both axes. If the electron were confined along one coordinate (e.g., between 0 < x < L<sub>x</sub>) but not along the other (i.e., B(y)

is either restricted to vanish at  $y=0$  or at  $y=L_y$  or at neither point), then the total energy  $E$  lies in the continuum; its  $E_x$  component is quantized but  $E_y$  is not. Such cases arise, for example, when a linear triatomic molecule has more than enough energy in one of its bonds to rupture it but not much energy in the other bond; the first bond's energy lies in the continuum, but the second bond's energy is quantized.

Perhaps more interesting is the case in which the bond with the higher dissociation energy is excited to a level that is not enough to break it but that is in excess of the dissociation energy of the weaker bond. In this case, one has two degenerate states- i. the strong bond having high internal energy and the weak bond having low energy ( $\psi_1$ ), and ii. the strong bond having little energy and the weak bond having more than enough energy to rupture it ( $\psi_2$ ). Although an experiment may prepare the molecule in a state that contains only the former component (i.e.,  $\psi = C_1 \psi_1 + C_2 \psi_2$  with  $C_1 \gg C_2$ ), coupling between the two degenerate functions (induced by terms in the Hamiltonian  $\mathbf{H}$  that have been ignored in defining  $\psi_1$  and  $\psi_2$ ) usually causes the true wavefunction  $\psi = \exp(-it\mathbf{H}/\hbar)$  to acquire a component of the second function as time evolves. In such a case, one speaks of internal vibrational energy flow giving rise to unimolecular decomposition of the molecule.

### 3. Energies and Wavefunctions for Bound States

*For discrete energy levels, the energies are specified functions that depend on quantum numbers, one for each degree of freedom that is quantized*

Returning to the situation in which motion is constrained along both axes, the resultant total energies and wavefunctions (obtained by inserting the quantum energy levels into the expressions for

$A(x)$   $B(y)$  are as follows:

$$E_x = n_x^2 \frac{\hbar^2}{2mL_x^2}, \text{ and}$$

$$E_y = n_y^2 \frac{\hbar^2}{2mL_y^2},$$

$$E = E_x + E_y ,$$

$$\psi(x,y) = (1/2L_x)^{1/2} (1/2L_y)^{1/2} [\exp(in_x x/L_x) - \exp(-in_x x/L_x)]$$

$$[\exp(in_y y/L_y) - \exp(-in_y y/L_y)], \text{ with } n_x \text{ and } n_y = 1, 2, 3, \dots$$

The two  $(1/2L)^{1/2}$  factors are included to guarantee that  $\psi$  is normalized:

$$\int |\psi(x,y)|^2 dx dy = 1.$$

Normalization allows  $|\psi(x,y)|^2$  to be properly identified as a probability density for finding the electron at a point  $x, y$ .

#### 4. Quantized Action Can Also be Used to Derive Energy Levels

There is another approach that can be used to find energy levels and is especially straightforward to use for systems whose Schrödinger equations are separable. The so-called classical **action** (denoted  $S$ ) of a particle moving with momentum  $\mathbf{p}$  along a path leading from initial coordinate  $\mathbf{q}_i$  at initial time  $t_i$  to a final coordinate  $\mathbf{q}_f$  at time  $t_f$  is defined by:

$$S = \int_{\mathbf{q}_i; t_i}^{\mathbf{q}_f; t_f} \mathbf{p} \cdot d\mathbf{q} .$$

Here, the momentum vector  $\mathbf{p}$  contains the momenta along all coordinates of the system, and the coordinate vector  $\mathbf{q}$  likewise contains the coordinates along all such degrees of freedom. For example, in the two-dimensional particle in a box problem considered above,  $\mathbf{q} = (x, y)$  has two components as does  $\mathbf{p} = (P_x, p_y)$ , and the action integral is:

$$S = \int_{x_i; y_i; t_i}^{x_f; y_f; t_f} (p_x dx + p_y dy) .$$

In computing such actions, it is essential to keep in mind the sign of the momentum as the particle moves from its initial to its final positions. An example will help clarify these matters.

For systems such as the above particle in a box example for which the Hamiltonian is separable, the action integral decomposed into a sum of such integrals, one for each degree of freedom. In this two-dimensional example, the additivity of  $H$ :

$$H = H_x + H_y = p_x^2/2m + p_y^2/2m + V(x) + V(y)$$

$$= -\hbar^2/2m \nabla_x^2 + V(x) - \hbar^2/2m \nabla_y^2 + V(y)$$

means that  $p_x$  and  $p_y$  can be independently solved for in terms of the potentials  $V(x)$  and  $V(y)$  as well as the energies  $E_x$  and  $E_y$  associated with each separate degree of freedom:

$$p_x = \pm \sqrt{2m(E_x - V(x))}$$

$$p_y = \pm \sqrt{2m(E_y - V(y))} ;$$

the signs on  $p_x$  and  $p_y$  must be chosen to properly reflect the motion that the particle is actually undergoing. Substituting these expressions into the action integral yields:

$$S = S_x + S_y$$

$$= \int_{x_i; t_i}^{x_f; t_f} \pm \sqrt{2m(E_x - V(x))} dx + \int_{y_i; t_i}^{y_f; t_f} \pm \sqrt{2m(E_y - V(y))} dy .$$

The relationship between these classical action integrals and existence of quantized energy levels has been shown to involve equating the classical action for motion on a closed path (i.e., a path that starts and ends at the same place after undergoing motion away from the starting point but eventually returning to the starting coordinate at a later time) to an integral multiple of Planck's constant:

$$S_{\text{closed}} = \int_{\mathbf{q}_i; t_i}^{\mathbf{q}_f; t_f} \mathbf{p} \cdot d\mathbf{q} = n h. \quad (n = 1, 2, 3, 4, \dots)$$

Applied to each of the independent coordinates of the two-dimensional particle in a box problem, this expression reads:

$$n_x h = \int_{x=0}^{x=L_x} \sqrt{2m(E_x - V(x))} dx + \int_{x=L_x}^{x=0} -\sqrt{2m(E_x - V(x))} dx$$

$$n_y h = \int_{y=0}^{y=L_y} \sqrt{2m(E_y - V(y))} dy + \int_{y=L_y}^{y=0} -\sqrt{2m(E_y - V(y))} dy .$$

Notice that the sign of the momenta are positive in each of the first integrals appearing above (because the particle is moving from  $x = 0$  to  $x = L_x$ , and analogously for  $y$ -motion, and thus has positive momentum) and negative in each of the second integrals (because the motion is from  $x = L_x$  to  $x = 0$  (and analogously for  $y$ -motion) and thus with negative momentum). Within the region bounded by  $0 \leq x \leq L_x$ ;  $0 \leq y \leq L_y$ , the potential vanishes, so  $V(x) = V(y) = 0$ . Using this fact, and reversing the upper and lower limits, and thus the sign, in the second integrals above, one obtains:

$$n_x h = 2 \int_{x=0}^{x=L_x} \sqrt{2mE_x} dx = 2 \sqrt{2mE_x} L_x$$

$$n_y h = 2 \int_{y=0}^{y=L_y} \sqrt{2mE_y} dy = 2 \sqrt{2mE_y} L_y.$$

Solving for  $E_x$  and  $E_y$ , one finds:

$$E_x = \frac{(n_x h)^2}{8mL_x^2}$$

$$E_y = \frac{(n_y h)^2}{8mL_y^2} .$$

These are the same quantized energy levels that arose when the wavefunction boundary conditions were matched at  $x = 0$ ,  $x = L_x$  and  $y = 0$ ,  $y = L_y$ . In this case, one says that the Bohr-Sommerfeld quantization condition:

$$n h = \int_{\mathbf{q}_i; \mathbf{t}_i}^{\mathbf{q}_f; \mathbf{t}_f} \mathbf{p} \cdot d\mathbf{q}$$

has been used to obtain the result.

## B. Other Model Problems

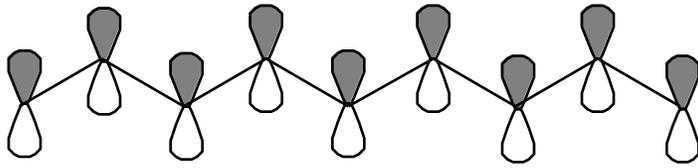
### 1. Particles in Boxes

*The particle-in-a-box problem provides an important model for several relevant chemical situations*

The above 'particle in a box' model for motion in two dimensions can obviously be extended to three dimensions or to one.

For two and three dimensions, it provides a crude but useful picture for electronic states on surfaces or in crystals, respectively. Free motion within a spherical volume gives rise to eigenfunctions that are used in nuclear physics to describe the motions of neutrons and protons in nuclei. In the so-called shell model of nuclei, the neutrons and protons fill separate s, p, d, etc orbitals with each type of nucleon forced to obey the Pauli principle. These orbitals are not the same in their radial 'shapes' as the s, p, d, etc orbitals of atoms because, in atoms, there is an additional radial potential  $V(r) = -Ze^2/r$  present. However, their angular shapes are the same as in atomic structure because, in both cases, the potential is independent of  $\theta$  and  $\phi$ . This same spherical box model has been used to describe the orbitals of valence electrons in clusters of mono-valent metal atoms such as  $Cs_n$ ,  $Cu_n$ ,  $Na_n$  and their positive and negative ions. Because of the metallic nature of these species, their valence electrons are sufficiently delocalized to render this simple model rather effective (see T. P. Martin, T. Bergmann, H. Göhlich, and T. Lange, J. Phys. Chem. 95, 6421 (1991)).

One-dimensional free particle motion provides a qualitatively correct picture for  $\pi$ -electron motion along the p orbitals of a delocalized polyene. The one cartesian dimension then corresponds to motion along the delocalized chain. In such a model, the box length L is related to the carbon-carbon bond length R and the number N of carbon centers involved in the delocalized network  $L=(N-1)R$ . Below, such a conjugated network involving nine centers is depicted. In this example, the box length would be eight times the C-C bond length.



## Conjugated Network with 9 Centers Involved

The eigenstates  $\psi_n(x)$  and their energies  $E_n$  represent orbitals into which electrons are placed. In the example case, if nine electrons are present (e.g., as in the 1,3,5,7-nonatetraene radical), the ground electronic state would be represented by a total wavefunction consisting of a product in which the lowest four  $\psi$ 's are doubly occupied and the fifth  $\psi$  is singly occupied:

$$\Psi = \psi_1 \psi_1 \psi_2 \psi_2 \psi_3 \psi_3 \psi_4 \psi_4 \psi_5 .$$

A product wavefunction is appropriate because the total Hamiltonian involves the kinetic plus potential energies of nine electrons. To the extent that this total energy can be represented as the sum of nine separate energies, one for each electron, the Hamiltonian allows a separation of variables

$$H = \sum_j H(j)$$

in which each  $H(j)$  describes the kinetic and potential energy of an individual electron. This (approximate) additivity of  $H$  implies that solutions of  $H \Psi = E \Psi$  are products of solutions to  $H(j) \psi_j(\mathbf{r}_j) = E_j \psi_j(\mathbf{r}_j)$ .

The two lowest  $\psi$ -excited states would correspond to states of the form

$$\psi_1^* \psi_1 \psi_2 \psi_2 \psi_3 \psi_3 \psi_4 \psi_5 \psi_5 , \text{ and}$$

$$\psi_1^* \psi_1 \psi_2 \psi_2 \psi_3 \psi_3 \psi_4 \psi_4 \psi_6 ,$$

where the spin-orbitals (orbitals multiplied by  $\alpha$  or  $\beta$ ) appearing in the above products depend on the coordinates of the various electrons. For example,

1 1 2 2 3 3 4 5 5

denotes

1 ( $\mathbf{r}_1$ ) 1 ( $\mathbf{r}_2$ ) 2 ( $\mathbf{r}_3$ ) 2 ( $\mathbf{r}_4$ ) 3 ( $\mathbf{r}_5$ ) 3 ( $\mathbf{r}_6$ ) 4 ( $\mathbf{r}_7$ ) 5  
 ( $\mathbf{r}_8$ ) 5 ( $\mathbf{r}_9$ ).

The electronic excitation energies within this model would be

$$E^* = \frac{\hbar^2}{2m} [5^2/L^2 - 4^2/L^2] \text{ and}$$

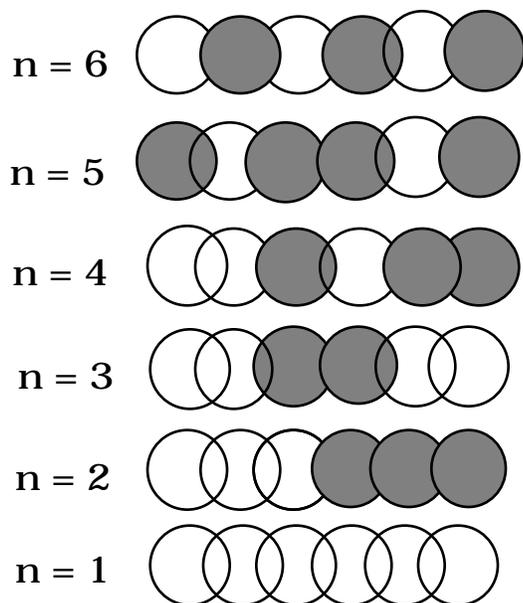
$E^* = \frac{\hbar^2}{2m} [6^2/L^2 - 5^2/L^2]$ , for the two excited-state functions described above. It turns out that this simple model of  $\pi$ -electron energies provides a qualitatively correct picture of such excitation energies.

This simple particle-in-a-box model does not yield orbital energies that relate to ionization energies unless the potential 'inside the box' is specified. Choosing the value of this potential  $V_0$  such that  $V_0 + \frac{\hbar^2}{2m} [5^2/L^2]$  is equal to minus the lowest ionization energy of the 1,3,5,7-nonatetraene radical, gives energy levels (as  $E = V_0 + \frac{\hbar^2}{2m} [n^2/L^2]$ ) which then are approximations to ionization energies.

The individual  $\pi$ -molecular orbitals

$$\psi_n = (2/L)^{1/2} \sin(n \pi x/L)$$

are depicted in the figure below for a model of the 1,3,5 hexatriene  $\pi$ -orbital system for which the 'box length'  $L$  is five times the distance  $R_{CC}$  between neighboring pairs of Carbon atoms.



$$(2/L)^{1/2} \sin(n \ x/L); L = 5 \times R_{CC}$$

In this figure, positive amplitude is denoted by the clear spheres and negative amplitude is shown by the darkened spheres; the magnitude of the  $k^{\text{th}}$  C-atom centered atomic orbital in the  $n^{\text{th}}$  -molecular orbital is given by  $(2/L)^{1/2} \sin(n \ kR_{CC}/L)$ .

This simple model allows one to estimate spin densities at each carbon center and provides insight into which centers should be most amenable to electrophilic or nucleophilic attack. For example, radical attack at the  $C_5$  carbon of the nine-atom system described earlier would be more facile for the ground state than for either  $\pi^*$  or  $\sigma^*$ . In the former, the unpaired spin density resides in  $\psi_5$ , which has non-zero amplitude at the  $C_5$  site  $x=L/2$ ; in  $\pi^*$  and  $\sigma^*$ , the unpaired density is in  $\psi_4$  and  $\psi_6$ , respectively, both of which have zero density at  $C_5$ . These densities reflect the values  $(2/L)^{1/2} \sin(n \ kR_{CC}/L)$  of the amplitudes for this case in which  $L = 8 \times R_{CC}$  for  $n = 5, 4,$  and  $6$ , respectively.

## 2. One Electron Moving About a Nucleus

The Hydrogenic atom problem forms the basis of much of our thinking about atomic structure. To solve the corresponding Schrödinger equation requires separation of the  $r$ ,  $\theta$ , and  $\phi$  variables

[Suggested Extra Reading- Appendix B: The Hydrogen Atom Orbitals]

The Schrödinger equation for a single particle of mass  $\mu$  moving in a central potential (one that depends only on the radial coordinate  $r$ ) can be written as

$$-\frac{\hbar^2}{2\mu} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \psi + V \sqrt{x^2+y^2+z^2} \psi = E \psi .$$

This equation is not separable in cartesian coordinates  $(x,y,z)$  because of the way  $x,y$ , and  $z$  appear together in the square root. However, it is separable in spherical coordinates

$$-\frac{\hbar^2}{2\mu r^2} \left( \frac{\partial^2}{\partial r^2} + \frac{1}{\sin^2 \theta} \left( \frac{\partial^2}{\partial \theta^2} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right) \right) \psi + V(r) \psi = E \psi .$$

Subtracting  $V(r)$  from both sides of the equation and multiplying by  $-\frac{2\mu r^2}{\hbar^2}$  then moving the derivatives with respect to  $r$  to the right-hand side, one obtains

$$\frac{1}{\sin^2 \theta} \left( \frac{\partial^2}{\partial \theta^2} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right) \psi = -\frac{2\mu r^2}{\hbar^2} (E - V(r)) \psi - \frac{\partial^2 \psi}{\partial r^2} .$$

Notice that the right-hand side of this equation is a function of  $r$  only; it contains no  $\theta$  or  $\phi$  dependence. Let's call the entire right hand side  $\mathbf{F}(r)$  to emphasize this fact.

To further separate the  $\theta$  and  $\phi$  dependence, we multiply by  $\sin^2 \theta$  and subtract the derivative terms from both sides to obtain

$$\frac{\partial^2 \psi}{\partial \theta^2} = \mathbf{F}(r) \sin^2 \theta \psi - \sin^2 \theta \frac{\partial^2 \psi}{\partial \phi^2} .$$

Now we have separated the  $\theta$  dependence from the  $\phi$  and  $r$  dependence. If we now substitute  $\psi = Q(\theta) Q(r)$  and divide by  $Q$ , we obtain

$$\frac{1}{2} \frac{d^2 Q}{dr^2} = \frac{1}{Q} F(r) \sin^2 \theta - \sin^2 \theta - \frac{Q}{r^2}.$$

Now all of the  $\theta$  dependence is isolated on the left hand side; the right hand side contains only  $r$  and  $Q$  dependence.

Whenever one has isolated the entire dependence on one variable as we have done above for the  $\theta$  dependence, one can easily see that the left and right hand sides of the equation must equal a constant. For the above example, the left hand side contains no  $r$  or  $Q$  dependence and the right hand side contains no  $\theta$  dependence. Because the two sides are equal, they both must actually contain no  $r$ ,  $\theta$ , or  $Q$  dependence; that is, they are constant.

For the above example, we therefore can set both sides equal to a so-called separation constant that we call  $-m^2$ . It will become clear shortly why we have chosen to express the constant in this form.

#### a. The $\theta$ Equation

The resulting  $\theta$  equation reads

$$\frac{d^2 Q}{d\theta^2} + m^2 Q = 0$$

which has as its most general solution

$$Q = A e^{im\theta} + B e^{-im\theta}.$$

We must require the function  $Q$  to be single-valued, which means that

$$Q(\theta) = Q(\theta + 2\pi) \text{ or,}$$

$$A e^{im\theta} (1 - e^{2im\pi}) + B e^{-im\theta} (1 - e^{-2im\pi}) = 0.$$

This is satisfied only when the separation constant is equal to an integer  $m = 0, \pm 1, \pm 2, \dots$  and provides another example of the rule that quantization comes from the boundary conditions on the wavefunction. Here  $m$  is restricted to certain discrete values because the wavefunction must be such that when you rotate through  $2\pi$  about the  $z$ -axis, you must get back what you started with.

#### b. The $r$ Equation

Now returning to the equation in which the  $\theta$  dependence was isolated from the  $r$  and  $Q$  dependence, and rearranging the  $\theta$  terms to the left-hand side, we have

$$\frac{1}{\sin^2 \theta} \frac{d^2 Q}{d\theta^2} - \frac{Q}{\sin^2 \theta} = F(r)Q.$$

In this equation we have separated  $\theta$  and  $r$  variations so we can further decompose the wavefunction by introducing  $Q = ( ) R(r)$ , which yields

$$\frac{1}{\sin^2 \theta} \frac{d}{d\theta} \left( \sin^2 \theta \frac{dQ}{d\theta} \right) - \frac{m^2}{\sin^2 \theta} Q = \frac{F(r)R}{R} = -\lambda,$$

where a second separation constant,  $-\lambda$ , has been introduced once the  $r$  and  $\theta$  dependent terms have been separated onto the right and left hand sides, respectively.

We now can write the  $\theta$  equation as

$$\frac{1}{\sin^2 \theta} \frac{d}{d\theta} \left( \sin^2 \theta \frac{dQ}{d\theta} \right) - \frac{m^2}{\sin^2 \theta} Q = -\lambda,$$

where  $m$  is the integer introduced earlier. To solve this equation for  $Q$ , we make the substitutions  $z = \cos \theta$  and  $P(z) = ( )$ , so  $\sqrt{1-z^2} = \sin \theta$ , and

$$\frac{dQ}{d\theta} = \frac{z}{\sin \theta} \frac{dP}{dz} = -\frac{z}{\sqrt{1-z^2}} \frac{dP}{dz}.$$

The range of values for  $\theta$  was  $0 < \theta < \pi$ , so the range for  $z$  is  $-1 < z < 1$ . The equation for  $P$ , when expressed in terms of  $P$  and  $z$ , becomes

$$\frac{d}{dz} \left( (1-z^2) \frac{dP}{dz} \right) - \frac{m^2 P}{1-z^2} + \lambda P = 0.$$

Now we can look for polynomial solutions for  $P$ , because  $z$  is restricted to be less than unity in magnitude. If  $m = 0$ , we first let

$$P = \sum_{k=0}^{\infty} a_k z^k,$$

and substitute into the differential equation to obtain

$$\sum_{k=0}^{\infty} (k+2)(k+1) a_{k+2} z^k - \sum_{k=0}^{\infty} (k+1)k a_k z^k + \sum_{k=0}^{\infty} a_k z^k = 0.$$

Equating like powers of  $z$  gives

$$a_{k+2} = \frac{a_k(k(k+1) - \lambda)}{(k+2)(k+1)}.$$

Note that for large values of  $k$

$$\frac{a_{k+2}}{a_k} = \frac{k^2 - 1 + \frac{1}{k}}{k^2 - 1 + \frac{2}{k} - 1 + \frac{1}{k}} = 1.$$

Since the coefficients do not decrease with  $k$  for large  $k$ , this series will diverge for  $z = \pm 1$  unless it truncates at finite order. This truncation only happens if the separation constant obeys  $l(l+1)$ , where  $l$  is an integer. So, once again, we see that a boundary condition (i.e., that the wavefunction be normalizable in this case) give rise to quantization. In this case, the values of  $l$  are restricted to  $l(l+1)$ ; before, we saw that  $m$  is restricted to  $0, \pm 1, \pm 2, \dots$ .

Since this recursion relation links every other coefficient, we can choose to solve for the even and odd functions separately. Choosing  $a_0$  and then determining all of the even  $a_k$  in terms of this  $a_0$ , followed by rescaling all of these  $a_k$  to make the function normalized generates an even solution. Choosing  $a_1$  and determining all of the odd  $a_k$  in like manner, generates an odd solution.

For  $l=0$ , the series truncates after one term and results in  $P_0(z) = 1$ . For  $l=1$  the same thing applies and  $P_1(z) = z$ . For  $l=2$ ,  $a_2 = -6 \frac{a_0}{2} = -3a_0$ , so one obtains  $P_2 = 3z^2 - 1$ , and so on. These polynomials are called Legendre polynomials.

For the more general case where  $m \neq 0$ , one can proceed as above to generate a polynomial solution for the  $\Theta$  function. Doing so, results in the following solutions:

$$P_l^m(z) = (1-z^2)^{\frac{|m|}{2}} \frac{d^{|m|} P_l(z)}{dz^{|m|}}.$$

These functions are called Associated Legendre polynomials, and they constitute the solutions to the  $\Theta$  problem for non-zero  $m$  values.

The above  $P$  and  $e^{im\phi}$  functions, when re-expressed in terms of  $\theta$  and  $\phi$ , yield the full angular part of the wavefunction for any centrosymmetric potential. These solutions are usually written as  $Y_{l,m}(\theta, \phi) = P_l^m(\cos \theta) (2\pi)^{\frac{1}{2}} \exp(im\phi)$ , and are called spherical harmonics. They provide the angular solution of the  $r, \theta, \phi$  Schrödinger equation for any problem in which the potential depends only on the radial coordinate. Such situations include all one-electron atoms and ions (e.g., H, He<sup>+</sup>, Li<sup>++</sup>, etc.), the rotational motion of a diatomic molecule (where the potential depends only on bond length  $r$ ), the motion of a nucleon in a spherically symmetrical "box" (as occurs in the shell model of nuclei), and the scattering of two atoms (where the potential depends only on interatomic distance).

c. The R Equation

Let us now turn our attention to the radial equation, which is the only place that the explicit form of the potential appears. Using our derived results and specifying  $V(r)$  to be the coulomb potential appropriate for an electron in the field of a nucleus of charge  $+Ze$ , yields:

$$\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) + \frac{2\mu}{\hbar^2} \left( E + \frac{Ze^2}{r} - \frac{l(l+1)}{r^2} \right) R = 0.$$

We can simplify things considerably if we choose rescaled length and energy units because doing so removes the factors that depend on  $\mu, \hbar$ , and  $e$ . We introduce a new radial coordinate  $\rho$  and a quantity  $\lambda$  as follows:

$$\rho = \frac{-8\mu E}{\hbar^2} \frac{1}{2} r, \text{ and } \lambda^2 = -\frac{\mu Z^2 e^4}{2E\hbar^2}.$$

Notice that if  $E$  is negative, as it will be for bound states (i.e., those states with energy below that of a free electron infinitely far from the nucleus and with zero kinetic energy),  $\lambda$  is real. On the other hand, if  $E$  is positive, as it will be for states that lie in the continuum,  $\lambda$  will be imaginary. These two cases will give rise to qualitatively different behavior in the solutions of the radial equation developed below.

We now define a function  $S$  such that  $S(\rho) = R(r)$  and substitute  $S$  for  $R$  to obtain:

$$\frac{1}{2} \frac{d}{d\rho} \left( \frac{dS}{d\rho} \right) + \left( -\frac{1}{4} - \frac{l(l+1)}{2\rho^2} + \lambda^2 \right) S = 0.$$

The differential operator terms can be recast in several ways using

$$\frac{1}{2} \frac{d}{d\rho} \left( \frac{dS}{d\rho} \right) = \frac{d^2 S}{d\rho^2} + \frac{1}{2} \frac{dS}{d\rho} = \frac{1}{2} \frac{d^2}{d\rho^2} (S).$$

It is useful to keep in mind these three embodiments of the derivatives that enter into the radial kinetic energy; in various contexts it will be useful to employ various of these.

The strategy that we now follow is characteristic of solving second order differential equations. We will examine the equation for  $S$  at large and small  $\rho$  values. Having found solutions at these limits, we will use a power series in  $\rho$  to "interpolate" between these two limits.

Let us begin by examining the solution of the above equation at small values of  $\rho$  to see how the radial functions behave at small  $r$ . As  $\rho \rightarrow 0$ , the second term in the brackets will dominate. Neglecting the other two terms in the brackets, we find that, for small values of  $\rho$  (or  $r$ ), the solution should behave like  $\rho^{-L}$  and because the function must be normalizable, we must have  $L > 0$ . Since  $L$  can be any non-negative integer, this suggests the following more general form for  $S(\rho)$ :

$$S(\rho) = \rho^{-L} e^{-a\rho}.$$

This form will insure that the function is normalizable since  $S(\infty) = 0$  as  $r \rightarrow \infty$  for all  $L$ , as long as  $a$  is a real quantity. If  $a$  is imaginary, such a form may not be normalized (see below for further consequences).

Turning now to the behavior of  $S$  for large  $L$ , we make the substitution of  $S(\infty)$  into the above equation and keep only the terms with the largest power of  $L$  (e.g., first term in brackets). Upon so doing, we obtain the equation

$$a^2 L e^{-a} = \frac{1}{4} L e^{-a},$$

which leads us to conclude that the exponent in the large- $L$  behavior of  $S$  is  $a = \frac{1}{2}$ .

Having found the small- and large- $L$  behaviors of  $S(\infty)$ , we can take  $S$  to have the following form to interpolate between large and small  $L$ -values:

$$S(\infty) = L e^{-\frac{1}{2}} P(\infty),$$

where the function  $L$  is expanded in an infinite power series in  $L$  as  $P(\infty) = \sum a_k L^k$ . Again substituting this expression for  $S$  into the above equation we obtain

$$P'' + P'(2L+2) + P(L-1) = 0,$$

and then substituting the power series expansion of  $P$  and solving for the  $a_k$ 's we arrive at:

$$a_{k+1} = \frac{(k+L+1) a_k}{(k+1)(k+2L+2)}.$$

For large  $k$ , the ratio of expansion coefficients reaches the limit  $\frac{a_{k+1}}{a_k} = \frac{1}{k}$ , which has the same behavior as the power series expansion of  $e^{-x}$ . Because the power series expansion of  $P$  describes a function that behaves like  $e^{-x}$  for large  $x$ , the resulting  $S(\infty)$

function would not be normalizable because the  $e^{-\frac{1}{2}}$  factor would be overwhelmed by this  $e^{-x}$  dependence. Hence, the series expansion of  $P$  must truncate in order to achieve a normalizable  $S$  function. Notice that if  $a$  is imaginary, as it will be if  $E$  is in the continuum, the argument that the series must truncate to avoid an exponentially diverging function no longer applies. Thus, we see a key difference between bound (with  $a$  real) and continuum (with  $a$  imaginary) states. In the former case, the boundary condition of non-divergence arises; in the latter, it does not.

To truncate at a polynomial of order  $n'$ , we must have  $n' - L + 1 = 0$ . This implies that the quantity  $L$  introduced previously is restricted to  $L = n' + 1$ , which is certainly an integer; let us call this integer  $n$ . If we label states in order of increasing  $n = 1, 2, 3, \dots$ , we see that doing so is consistent with specifying a maximum order ( $n'$ ) in the

$P(\rho)$  polynomial  $n' = 0, 1, 2, \dots$  after which the  $l$ -value can run from  $l = 0$ , in steps of unity up to  $l = n - 1$ .

Substituting the integer  $n$  for  $n'$ , we find that the energy levels are quantized because  $E$  is quantized (equal to  $n$ ):

$$E = - \frac{\mu Z^2 e^4}{2\hbar^2 n^2} \quad \text{and} \quad r = \frac{Zr_0}{n}.$$

Here, the length  $a_0$  is the so called Bohr radius  $a_0 = \frac{\hbar^2}{\mu e^2}$ ; it appears once the above  $E$ -

expression is substituted into the equation for  $P(\rho)$ . Using the recursion equation to solve for the polynomial's coefficients  $a_k$  for any choice of  $n$  and  $l$  quantum numbers generates a so-called Laguerre polynomial;  $P_{n-l-1}(\rho)$ . They contain powers of  $\rho$  from zero through  $n-l-1$ .

This energy quantization does not arise for states lying in the continuum because the condition that the expansion of  $P(\rho)$  terminate does not arise. The solutions of the radial equation appropriate to these scattering states (which relate to the scattering motion of an electron in the field of a nucleus of charge  $Z$ ) are treated on p. 90 of EWK.

In summary, separation of variables has been used to solve the full  $r, \theta, \phi$  Schrödinger equation for one electron moving about a nucleus of charge  $Z$ . The  $\theta$  and  $\phi$  solutions are the spherical harmonics  $Y_{L,m}(\theta, \phi)$ . The bound-state radial solutions

$$R_{n,L}(r) = S(\rho) = L e^{-\rho/2} P_{n-L-1}(\rho)$$

depend on the  $n$  and  $l$  quantum numbers and are given in terms of the Laguerre polynomials (see EWK for tabulations of these polynomials).

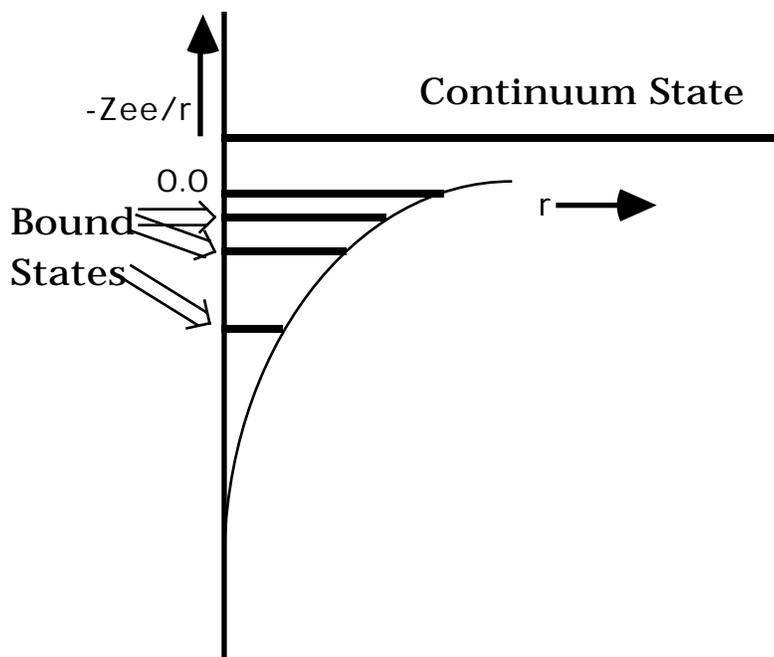
#### d. Summary

To summarize, the quantum numbers  $l$  and  $m$  arise through boundary conditions requiring that  $Y_{l,m}(\theta, \phi)$  be normalizable (i.e., not diverge) and  $Y_{l,m}(\theta, \phi) = Y_{l,m}(\theta + 2\pi, \phi)$ . In the texts by Atkins, EWK, and McQuarrie the differential equations obeyed by the  $\theta$  and  $\phi$  components of  $Y_{l,m}$  are solved in more detail and properties of the solutions are discussed. This differential equation involves the three-dimensional Schrödinger equation's angular kinetic energy operator. That is, the angular part of the above Hamiltonian is equal to  $-\hbar^2 L^2 / 2mr^2$ , where  $L^2$  is the square of the total angular momentum for the electron.

The radial equation, which is the only place the potential energy enters, is found to possess both bound-states (i.e., states whose energies lie below the asymptote at which the potential vanishes and the kinetic energy is zero) and continuum states lying energetically above this asymptote. The resulting hydrogenic wavefunctions (angular and radial) and

energies are summarized in Appendix B for principal quantum numbers  $n$  ranging from 1 to 3 and in Pauling and Wilson for  $n$  up to 5.

There are both bound and continuum solutions to the radial Schrödinger equation for the attractive coulomb potential because, at energies below the asymptote the potential confines the particle between  $r=0$  and an outer turning point, whereas at energies above the asymptote, the particle is no longer confined by an outer turning point (see the figure below).



The solutions of this one-electron problem form the qualitative basis for much of atomic and molecular orbital theory. For this reason, the reader is encouraged to use Appendix B to gain a firmer understanding of the nature of the radial and angular parts of these wavefunctions. The orbitals that result are labeled by  $n$ ,  $l$ , and  $m$  quantum numbers for the bound states and by  $l$  and  $m$  quantum numbers and the energy  $E$  for the continuum states. Much as the particle-in-a-box orbitals are used to qualitatively describe  $\pi$ -electrons in conjugated polyenes, these so-called hydrogen-like orbitals provide qualitative descriptions of orbitals of atoms with more than a single electron. By introducing the concept of screening as a way to represent the repulsive interactions among the electrons of an atom, an effective nuclear charge  $Z_{\text{eff}}$  can be used in place of  $Z$  in the  $\psi_{n,l,m}$  and  $E_{n,l}$  to generate approximate atomic orbitals to be filled by electrons in a many-electron atom. For

example, in the crudest approximation of a carbon atom, the two 1s electrons experience the full nuclear attraction so  $Z_{\text{eff}}=6$  for them, whereas the 2s and 2p electrons are screened by the two 1s electrons, so  $Z_{\text{eff}}= 4$  for them. Within this approximation, one then occupies two 1s orbitals with  $Z=6$ , two 2s orbitals with  $Z=4$  and two 2p orbitals with  $Z=4$  in forming the full six-electron wavefunction of the lowest-energy state of carbon.

### 3. Rotational Motion For a Rigid Diatomic Molecule

*This Schrödinger equation relates to the rotation of diatomic and linear polyatomic molecules. It also arises when treating the angular motions of electrons in any spherically symmetric potential*

A diatomic molecule with fixed bond length  $R$  rotating in the absence of any external potential is described by the following Schrödinger equation:

$$\frac{\hbar^2}{2\mu} \left\{ (R^2 \sin^2 \theta)^{-1} \left[ \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) + (R^2 \sin^2 \theta)^{-1} \frac{\partial^2}{\partial \phi^2} \right] \right\} \psi = E \psi$$

or

$$L^2 / 2\mu R^2 = E \quad .$$

The angles  $\theta$  and  $\phi$  describe the orientation of the diatomic molecule's axis relative to a laboratory-fixed coordinate system, and  $\mu$  is the reduced mass of the diatomic molecule  $\mu=m_1 m_2 / (m_1 + m_2)$ . The differential operators can be seen to be exactly the same as those that arose in the hydrogen-like-atom case, and, as discussed above, these  $\frac{\partial}{\partial \theta}$  and  $\frac{\partial^2}{\partial \phi^2}$  differential operators are identical to the  $L^2$  angular momentum operator whose general properties are analyzed in Appendix G. Therefore, the same spherical harmonics that served as the angular parts of the wavefunction in the earlier case now serve as the entire wavefunction for the so-called rigid rotor:  $\psi = Y_{J,M}(\theta, \phi)$ . As detailed later in this text, the eigenvalues corresponding to each such eigenfunction are given as:

$$E_J = \hbar^2 J(J+1) / (2\mu R^2) = B J(J+1)$$

and are independent of  $M$ . Thus each energy level is labeled by  $J$  and is  $2J+1$ -fold degenerate (because  $M$  ranges from  $-J$  to  $J$ ). The so-called rotational constant  $B$  (defined as  $\hbar^2 / 2\mu R^2$ ) depends on the molecule's bond length and reduced mass. Spacings between

successive rotational levels (which are of spectroscopic relevance because angular momentum selection rules often restrict  $J$  to 1,0, and -1) are given by

$$E = B (J+1)(J+2) - B J(J+1) = 2B(J+1).$$

These energy spacings are of relevance to microwave spectroscopy which probes the rotational energy levels of molecules.

The rigid rotor provides the most commonly employed approximation to the rotational energies and wavefunctions of linear molecules. As presented above, the model restricts the bond length to be fixed. Vibrational motion of the molecule gives rise to changes in  $R$  which are then reflected in changes in the rotational energy levels. The coupling between rotational and vibrational motion gives rise to rotational  $B$  constants that depend on vibrational state as well as dynamical couplings, called centrifugal distortions, that cause the total ro-vibrational energy of the molecule to depend on rotational and vibrational quantum numbers in a non-separable manner.

#### 4. Harmonic Vibrational Motion

*This Schrödinger equation forms the basis for our thinking about bond stretching and angle bending vibrations as well as collective phonon motions in solids*

The radial motion of a diatomic molecule in its lowest ( $J=0$ ) rotational level can be described by the following Schrödinger equation:

$$-\frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} \psi + V(r) \psi = E \psi,$$

where  $\mu$  is the reduced mass  $\mu = m_1 m_2 / (m_1 + m_2)$  of the two atoms.

By substituting  $\psi = F(r)/r$  into this equation, one obtains an equation for  $F(r)$  in which the differential operators appear to be less complicated:

$$-\frac{\hbar^2}{2\mu} \frac{d^2 F}{dr^2} + V(r) F = E F.$$

This equation is exactly the same as the equation seen above for the radial motion of the electron in the hydrogen-like atoms except that the reduced mass  $\mu$  replaces the electron mass  $m$  and the potential  $V(r)$  is not the coulomb potential.

If the potential is approximated as a quadratic function of the bond displacement  $x = r - r_e$  expanded about the point at which  $V$  is minimum:

$$V = 1/2 k(r-r_e)^2,$$

the resulting harmonic-oscillator equation can be solved exactly. Because the potential  $V$  grows without bound as  $x$  approaches

or  $-\infty$ , only bound-state solutions exist for this model problem; that is, the motion is confined by the nature of the potential, so no continuum states exist.

In solving the radial differential equation for this potential (see Chapter 5 of McQuarrie), the large- $r$  behavior is first examined. For large- $r$ , the equation reads:

$$d^2F/dx^2 = 1/2 k x^2 (2\mu/\hbar^2) F,$$

where  $x = r - r_e$  is the bond displacement away from equilibrium. Defining  $y = (\mu k/\hbar^2)^{1/4} x$  as a new scaled radial coordinate allows the solution of the large- $r$  equation to be written as:

$$F_{\text{large-}r} = \exp(-y^2/2).$$

The general solution to the radial equation is then taken to be of the form:

$$F = \exp(-y^2/2) \sum_{n=0}^{\infty} C_n,$$

where the  $C_n$  are coefficients to be determined. Substituting this expression into the full radial equation generates a set of recursion equations for the  $C_n$  amplitudes. As in the solution of the hydrogen-like radial equation, the series described by these coefficients is divergent unless the energy  $E$  happens to equal specific values. It is this requirement that the wavefunction not diverge so it can be normalized that yields energy quantization. The energies of the states that arise are given by:

$$E_n = \hbar (k/\mu)^{1/2} (n+1/2),$$

and the eigenfunctions are given in terms of the so-called Hermite polynomials  $H_n(y)$  as follows:

$$\psi_n(x) = (n! 2^n)^{-1/2} (\mu/\hbar)^{1/4} \exp(-\mu x^2/2) H_n(\mu^{1/2} x),$$

where  $\alpha = (k\mu/\hbar^2)^{1/2}$ . Within this harmonic approximation to the potential, the vibrational energy levels are evenly spaced:

$$E = E_{n+1} - E_n = \hbar (k/\mu)^{1/2}.$$

In experimental data such evenly spaced energy level patterns are seldom seen; most commonly, one finds spacings  $E_{n+1} - E_n$  that decrease as the quantum number  $n$  increases. In such cases, one says that the progression of vibrational levels displays anharmonicity.

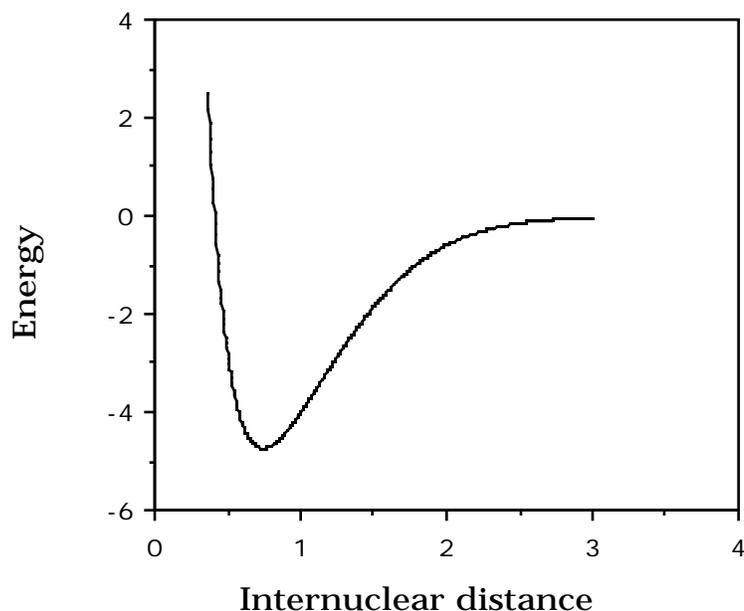
Because the  $H_n$  are odd or even functions of  $x$  (depending on whether  $n$  is odd or even), the wavefunctions  $\psi_n(x)$  are odd or even. This splitting of the solutions into two distinct classes is an example of the effect of symmetry; in this case, the symmetry is caused by the symmetry of the harmonic potential with respect to reflection through the origin along the  $x$ -axis. Throughout this text, many symmetries will arise; in each case, symmetry properties of the potential will cause the solutions of the Schrödinger equation to be decomposed into various symmetry groupings. Such symmetry decompositions are of great use because they provide additional quantum numbers (i.e., symmetry labels) by which the wavefunctions and energies can be labeled.

The harmonic oscillator energies and wavefunctions comprise the simplest reasonable model for vibrational motion. Vibrations of a polyatomic molecule are often characterized in terms of individual bond-stretching and angle-bending motions each of which is, in turn, approximated harmonically. This results in a total vibrational wavefunction that is written as a product of functions one for each of the vibrational coordinates.

Two of the most severe limitations of the harmonic oscillator model, the lack of anharmonicity (i.e., non-uniform energy level spacings) and lack of bond dissociation, result from the quadratic nature of its potential. By introducing model potentials that allow for proper bond dissociation (i.e., that do not increase without bound as  $x \rightarrow \infty$ ), the major shortcomings of the harmonic oscillator picture can be overcome. The so-called Morse potential (see the figure below)

$$V(r) = D_e (1 - \exp(-a(r-r_e)))^2,$$

is often used in this regard.



Here,  $D_e$  is the bond dissociation energy,  $r_e$  is the equilibrium bond length, and  $a$  is a constant that characterizes the 'steepness' of the potential and determines the vibrational frequencies. The advantage of using the Morse potential to improve upon harmonic-oscillator-level predictions is that its energy levels and wavefunctions are also known exactly. The energies are given in terms of the parameters of the potential as follows:

$$E_n = \hbar(k/\mu)^{1/2} \left\{ (n+1/2) - (n+1/2)^2 \hbar(k/\mu)^{1/2}/4D_e \right\},$$

where the force constant  $k$  is  $k=2D_e a^2$ . The Morse potential supports both bound states (those lying below the dissociation threshold for which vibration is confined by an outer turning point) and continuum states lying above the dissociation threshold. Its degree of anharmonicity is governed by the ratio of the harmonic energy  $\hbar(k/\mu)^{1/2}$  to the dissociation energy  $D_e$ .

### III. The Physical Relevance of Wavefunctions, Operators and Eigenvalues

*Having gained experience on the application of the Schrödinger equation to several of the more important model problems of chemistry, it is time to return to the issue of how the wavefunctions, operators, and energies relate to experimental reality.*

*In mastering the sections that follow the reader should keep in mind that :*

- i. It is the molecular system that possesses a set of characteristic wavefunctions and energy levels, but*
- ii. It is the experimental measurement that determines the nature by which these energy levels and wavefunctions are probed.*

This separation between the 'system' with its intrinsic set of energy levels and 'observation' or 'experiment' with its characteristic interaction with the system forms an important point of view used by quantum mechanics. It gives rise to a point of view in which the measurement itself can 'prepare' the system in a wavefunction that need not be any single eigenstate but can still be represented as a combination of the complete set of eigenstates. For the beginning student of quantum mechanics, these aspects of quantum mechanics are among the more confusing. If it helps, one should rest assured that all of the mathematical and 'rule' structure of this subject was created to permit the predictions of quantum mechanics to replicate what has been observed in laboratory experiments.

### ***Note to the Reader :***

*Before moving on to the next section, it would be very useful to work some of the Exercises and Problems. In particular, Exercises 3, 5, and 12 as well as problems 6, 8, and 11 provide insight that would help when the material of the next section is studied. The solution to Problem 11 is used throughout this section to help illustrate the concepts introduced here.*

#### **A. The Basic Rules and Relation to Experimental Measurement**

*Quantum mechanics has a set of 'rules' that link operators, wavefunctions, and eigenvalues to physically measurable properties. These rules have been formulated not in some arbitrary manner nor by derivation from some higher subject. Rather, the rules were designed to allow quantum mechanics to mimic the experimentally observed facts as revealed in mother nature's data. The extent to which these rules seem difficult to*

*understand usually reflects the presence of experimental observations that do not fit in with our common experience base.*

[Suggested Extra Reading- Appendix C: Quantum Mechanical Operators and Commutation]

The structure of quantum mechanics (QM) relates the wavefunction and operators  $\mathbf{F}$  to the 'real world' in which experimental measurements are performed through a set of rules (Dirac's text is an excellent source of reading concerning the historical development of these fundamentals). Some of these rules have already been introduced above. Here, they are presented in total as follows:

1. The time evolution of the wavefunction is determined by solving the time-dependent Schrödinger equation (see pp 23-25 of EWK for a rationalization of how the Schrödinger equation arises from the classical equation governing waves, Einstein's  $E=h\nu$ , and deBroglie's postulate that  $\lambda=h/p$ )

$$i\hbar \frac{\partial}{\partial t} \Psi = \mathbf{H} \Psi,$$

where  $\mathbf{H}$  is the Hamiltonian operator corresponding to the total (kinetic plus potential) energy of the system. For an isolated system (e.g., an atom or molecule not in contact with any external fields),  $\mathbf{H}$  consists of the kinetic and potential energies of the particles comprising the system. To describe interactions with an external field (e.g., an electromagnetic field, a static electric field, or the 'crystal field' caused by surrounding ligands), additional terms are added to  $\mathbf{H}$  to properly account for the system-field interactions.

If  $\mathbf{H}$  contains no explicit time dependence, then separation of space and time variables can be performed on the above Schrödinger equation  $\Psi = \psi \exp(-itE/\hbar)$  to give

$$\mathbf{H} \psi = E \psi.$$

In such a case, the time dependence of the state is carried in the phase factor  $\exp(-itE/\hbar)$ ; the spatial dependence appears in  $\psi(q_j)$ .

The so called time independent Schrödinger equation  $\mathbf{H} \psi = E \psi$  must be solved to determine the physically measurable energies  $E_k$  and wavefunctions  $\psi_k$  of the system. The most general solution to the full Schrödinger equation  $i\hbar \frac{\partial}{\partial t} \Psi = \mathbf{H} \Psi$  is then given by applying  $\exp(-i\mathbf{H}t/\hbar)$  to the wavefunction at some initial time ( $t=0$ )  $\Psi = \sum_k C_k \psi_k$  to obtain

$\psi(\mathbf{q}, t) = \sum_k C_k \psi_k \exp(-itE_k/\hbar)$ . The relative amplitudes  $C_k$  are determined by knowledge of the state at the initial time; this depends on how the system has been prepared in an earlier experiment. Just as Newton's laws of motion do not fully determine the time evolution of a classical system (i.e., the coordinates and momenta must be known at some initial time), the Schrödinger equation must be accompanied by initial conditions to fully determine  $\psi(\mathbf{q}, t)$ .

**Example :**

*Using the results of Problem 11 of this chapter to illustrate, the sudden ionization of  $N_2$  in its  $v=0$  vibrational state to generate  $N_2^+$  produces a vibrational wavefunction*

$$\psi_0 = \frac{1}{\sqrt{\pi}} e^{-x^2/2} = 3.53333 \text{Å}^{-1/2} e^{-(244.83 \text{Å}^{-2})(r-1.09769 \text{Å})^2}$$

*that was created by the fast ionization of  $N_2$ . Subsequent to ionization, this  $N_2$  function is not an eigenfunction of the new vibrational Schrödinger equation appropriate to  $N_2^+$ . As a result, this function will time evolve under the influence of the  $N_2^+$  Hamiltonian.*

*The time evolved wavefunction, according to this first rule, can be expressed in terms of the vibrational functions  $\{\psi_v\}$  and energies  $\{E_v\}$  of the  $N_2^+$  ion as*

$$\psi(\mathbf{q}, t) = \sum_v C_v \psi_v \exp(-i E_v t/\hbar).$$

*The amplitudes  $C_v$ , which reflect the manner in which the wavefunction is prepared (at  $t=0$ ), are determined by determining the component of each  $\psi_v$  in the function  $\psi$  at  $t=0$ . To do this, one uses*

$$C_v = \int \psi_v^* (\mathbf{q}, t=0) \psi(\mathbf{q}, t=0) d\mathbf{q} = C_v,$$

*which is easily obtained by multiplying the above summation by  $\psi_v^*$ , integrating, and using the orthonormality of the  $\{\psi_v\}$  functions.*

*For the case at hand, this results shows that by forming integrals involving products of the  $N_2$   $v=0$  function  $\psi_0$  ( $t=0$ )*

$$\psi_0 = -\frac{1}{4} e^{-x^2/2} = 3.53333 \text{Å}^{-\frac{1}{2}} e^{-(244.83 \text{Å}^{-2})(r-1.09769 \text{Å})^2}$$

and various  $N_2^+$  vibrational functions  $\psi_v$ , one can determine how  $\psi_0$  will evolve in time and the amplitudes of all  $\{\psi_v\}$  that it will contain. For example, the  $N_2$   $v=0$  function, upon ionization, contains the following amount of the  $N_2^+$   $v=0$  function:

$$C_0 = \int \psi_0^*(N_2^+) \psi_0(N_2) dr$$

$$= \int 3.47522 e^{-229.113(r-1.11642)^2} 3.53333 e^{-244.83(r-1.09769)^2} dr$$

As demonstrated in Problem 11, this integral reduces to 0.959. This means that the  $N_2$   $v=0$  state, subsequent to sudden ionization, can be represented as containing  $|0.959|^2 = 0.92$  fraction of the  $v=0$  state of the  $N_2^+$  ion.

This example relates to the well known Franck-Condon principle of spectroscopy in which squares of 'overlaps' between the initial electronic state's vibrational wavefunction and the final electronic state's vibrational wavefunctions allow one to estimate the probabilities of populating various final-state vibrational levels.

In addition to initial conditions, solutions to the Schrödinger equation must obey certain other constraints in form. They must be continuous functions of all of their spatial coordinates and must be single valued; these properties allow  $\psi^*$  to be interpreted as a probability density (i.e., the probability of finding a particle at some position can not be multivalued nor can it be 'jerky' or discontinuous). The derivative of the wavefunction must also be continuous except at points where the potential function undergoes an infinite jump (e.g., at the wall of an infinitely high and steep potential barrier). This condition relates to the fact that the momentum must be continuous except at infinitely 'steep' potential barriers where the momentum undergoes a 'sudden' reversal.

2. An experimental measurement of any quantity (whose corresponding operator is  $\mathbf{F}$ ) must result in one of the eigenvalues  $f_j$  of the operator  $\mathbf{F}$ . These eigenvalues are obtained by solving

$$\mathbf{F} \psi_j = f_j \psi_j,$$

where the  $\psi_j$  are the eigenfunctions of  $\mathbf{F}$ . Once the measurement of  $\mathbf{F}$  is made, for that sub-population of the experimental sample found to have the particular eigenvalue  $f_j$ , the wavefunction becomes  $\psi_j$ .

The equation  $\mathbf{H} \psi_k = E_k \psi_k$  is but a special case; it is an especially important case because much of the machinery of modern experimental chemistry is directed at placing the system in a particular energy quantum state by detecting its energy (e.g., by spectroscopic means).

The reader is strongly urged to also study Appendix C to gain a more detailed and illustrated treatment of this and subsequent rules of quantum mechanics.

3. The operators  $\mathbf{F}$  corresponding to all physically measurable quantities are Hermitian; this means that their matrix representations obey (see Appendix C for a description of the 'bra'  $\langle \psi |$  and 'ket'  $|\psi\rangle$  notation used below):

$$\langle \psi_j | \mathbf{F} | \psi_k \rangle = \langle \psi_k | \mathbf{F} | \psi_j \rangle^* = \langle \mathbf{F} \psi_j | \psi_k \rangle$$

in any basis  $\{\psi_j\}$  of functions appropriate for the action of  $\mathbf{F}$  (i.e., functions of the variables on which  $\mathbf{F}$  operates). As expressed through equality of the first and third elements above, Hermitian operators are often said to 'obey the turn-over rule'. This means that  $\mathbf{F}$  can be allowed to operate on the function to its right or on the function to its left if  $\mathbf{F}$  is Hermitian.

Hermiticity assures that the eigenvalues  $\{f_j\}$  are all real, that eigenfunctions  $\{\psi_j\}$  having different eigenvalues are orthogonal and can be normalized  $\langle \psi_j | \psi_k \rangle = \delta_{j,k}$ , and that eigenfunctions having the same eigenvalues can be made orthonormal (these statements are proven in Appendix C).

4. Once a particular value  $f_j$  is observed in a measurement of  $\mathbf{F}$ , this same value will be observed in all subsequent measurements of  $\mathbf{F}$  as long as the system remains undisturbed by measurements of other properties or by interactions with external fields. In fact, once  $f_j$  has been observed, the state of the system becomes an eigenstate of  $\mathbf{F}$  (if it already was, it remains unchanged):

$$\mathbf{F} \psi_j = f_j \psi_j.$$

This means that the measurement process itself may interfere with the state of the system and even determines what that state will be once the measurement has been made.

**Example:**

Again consider the  $v=0$   $N_2$  ionization treated in Problem 11 of this chapter. If, subsequent to ionization, the  $N_2^+$  ions produced were probed to determine their internal vibrational state, a fraction of the sample equal to  $|\langle (N_2; v=0) | (N_2^+; v=0) \rangle|^2 = 0.92$  would be detected in the  $v=0$  state of the  $N_2^+$  ion. For this sub-sample, the vibrational wavefunction becomes, and remains from then on,

$$\psi(t) = (N_2^+; v=0) \exp(-i t E_{v=0}^+ / \hbar),$$

where  $E_{v=0}^+$  is the energy of the  $N_2^+$  ion in its  $v=0$  state. If, at some later time, this sub-sample is again probed, all species will be found to be in the  $v=0$  state.

5. The probability  $P_k$  of observing a particular value  $f_k$  when  $F$  is measured, given that the system wavefunction is  $\psi$  prior to the measurement, is given by expanding  $\psi$  in terms of the complete set of normalized eigenstates of  $F$

$$\psi = \sum_j |j\rangle \langle j | \psi \rangle$$

and then computing  $P_k = |\langle k | \psi \rangle|^2$ . For the special case in which  $\psi$  is already one of the eigenstates of  $F$  (i.e.,  $\psi = |k\rangle$ ), the probability of observing  $f_j$  reduces to  $P_j = \delta_{j,k}$ . The set of numbers  $C_j = \langle j | \psi \rangle$  are called the expansion coefficients of  $\psi$  in the basis of the  $\{|j\rangle\}$ . These coefficients, when collected together in all possible products as  $D_{j,i} = C_i^* C_j$  form the so-called density matrix  $D_{j,i}$  of the wavefunction  $\psi$  within the  $\{|j\rangle\}$  basis.

**Example:**

If  $F$  is the operator for momentum in the  $x$ -direction and  $\psi(x,t)$  is the wave function for  $x$  as a function of time  $t$ , then the above expansion corresponds to a Fourier transform of

$$\psi(x,t) = \frac{1}{2} \int_{-\infty}^{\infty} \exp(ikx) \exp(-ikx') \psi(x',t) dx' dk.$$

Here  $(1/2)^{1/2} \exp(ikx)$  is the normalized eigenfunction of  $\hat{F} = -i\hbar / x$  corresponding to momentum eigenvalue  $\hbar k$ . These momentum eigenfunctions are orthonormal:

$$\int_{-\infty}^{\infty} \exp(-ikx) \exp(ik'x) dx = 2\pi \delta(k-k'),$$

and they form a complete set of functions in  $x$ -space

$$\int_{-\infty}^{\infty} \exp(-ikx) \exp(ikx') dk = 2\pi \delta(x-x')$$

because  $\hat{F}$  is a Hermitian operator. The function  $\int_{-\infty}^{\infty} \exp(-ikx') \psi(x',t) dx'$  is called the momentum-space transform of  $\psi(x,t)$  and is denoted  $\phi(k,t)$ ; it gives, when used as  $|\phi(k,t)|^2$ , the probability density for observing momentum values  $\hbar k$  at time  $t$ .

### Another Example:

Take the initial  $\psi$  to be a superposition state of the form

$$\psi = a(2p_0 + 2p_{-1} - 2p_1) + b(3p_0 - 3p_{-1}),$$

where the  $a$  and  $b$  are amplitudes that describe the admixture of  $2p$  and  $3p$  functions in this wavefunction. Then:

a. If  $L^2$  were measured, the value  $2\hbar^2$  would be observed with probability  $3|a|^2 + 2|b|^2 = 1$ , since all of the functions in  $\psi$  are  $p$ -type orbitals. After said measurement, the wavefunction would still be this same  $\psi$  because this entire  $\psi$  is an eigenfunction of  $L^2$ .

b. If  $L_z$  were measured for this

$$\psi = a(2p_0 + 2p_{-1} - 2p_1) + b(3p_0 - 3p_{-1}),$$

the values  $0\hbar$ ,  $1\hbar$ , and  $-1\hbar$  would be observed (because these are the only functions with non-zero  $C_m$  coefficients for the  $L_z$  operator) with respective probabilities  $|a|^2 + |b|^2$ ,  $|a|^2$ , and  $|a|^2 + |-b|^2$ .

c. After  $L_z$  were measured, if the sub-population for which  $-1\hbar$  had been detected were subjected to measurement of  $L^2$  the value  $2\hbar^2$  would certainly be found because the new wavefunction

$$\psi = \{-a 2p_{-1} - b 3p_{-1}\} (|a|^2 + |b|^2)^{-1/2}$$

is still an eigenfunction of  $L^2$  with this eigenvalue.

d. Again after  $L_z$  were measured, if the sub-population for which  $-1\hbar$  had been observed and for which the wavefunction is now

$$\psi = \{-a 2p_{-1} - b 3p_{-1}\} (|a|^2 + |b|^2)^{-1/2}$$

were subjected to measurement of the energy (through the Hamiltonian operator), two values would be found. With probability

$|a|^2 (|a|^2 + |b|^2)^{-1}$  the energy of the  $2p_{-1}$  orbital would be observed; with probability  $|b|^2 (|a|^2 + |b|^2)^{-1}$ , the energy of the  $3p_{-1}$  orbital would be observed.

If  $\psi$  is a function of several variables (e.g., when  $\psi$  describes more than one particle in a composite system), and if  $F$  is a property that depends on a subset of these variables (e.g., when  $F$  is a property of one of the particles in the composite system), then the expansion  $\psi = \sum_j |j\rangle \langle j| \psi$  is viewed as relating only to  $\psi$ 's dependence on the subset of variables related to  $F$ . In this case, the integrals  $\langle k| \psi \rangle$  are carried out over only these variables; thus the probabilities  $P_k = |\langle k| \psi \rangle|^2$  depend parametrically on the remaining variables.

### **Example:**

Suppose that  $\psi(r, \phi)$  describes the radial ( $r$ ) and angular ( $\phi$ ) motion of a diatomic molecule constrained to move on a planar surface. If an experiment were performed to measure the component of the rotational angular momentum of the diatomic molecule perpendicular to the surface ( $L_z = -i\hbar \frac{\partial}{\partial \phi}$ ), only values equal to  $m\hbar$  ( $m=0, 1, -1, 2, -2, 3, -3, \dots$ ) could be observed, because these are the eigenvalues of  $L_z$ :

$$L_z \psi_m = -i\hbar \frac{\partial}{\partial \phi} \psi_m = m\hbar \psi_m, \text{ where}$$

$$\psi_m = (1/2\pi)^{1/2} \exp(im\phi).$$

The quantization of  $L_z$  arises because the eigenfunctions  $\psi_m(\phi)$  must be periodic in  $\phi$ :

$$\psi(\phi + 2\pi) = \psi(\phi).$$

Such quantization (i.e., constraints on the values that physical properties can realize) will be seen to occur whenever the pertinent wavefunction is constrained to obey a so-called boundary condition (in this case, the boundary condition is  $\psi(\phi + 2\pi) = \psi(\phi)$ ).

Expanding the  $\phi$ -dependence of  $\psi$  in terms of the  $\psi_m$

$$\psi = \sum_{m < m/ >} \psi_m(\phi)$$

allows one to write the probability that  $m\hbar$  is observed if the angular momentum  $L_z$  is measured as follows:

$$P_m = |\langle m | \psi \rangle|^2 = \left| \int \psi_m^*(\phi) \psi(r, \phi) d\phi \right|^2.$$

If one is interested in the probability that  $m\hbar$  be observed when  $L_z$  is measured regardless of what bond length  $r$  is involved, then it is appropriate to integrate this expression over the  $r$ -variable about which one does not care. This, in effect, sums contributions from all  $r$ -values to obtain a result that is independent of the  $r$  variable. As a result, the probability reduces to:

$$P_m = \int \psi_m^*(\phi) \left\{ \int \psi_m^*(r, \phi) \psi(r, \phi) r dr \right\} d\phi d\phi',$$

which is simply the above result integrated over  $r$  with a volume element  $r dr$  for the two-dimensional motion treated here.

If, on the other hand, one were able to measure  $L_z$  values when  $r$  is equal to some specified bond length (this is only a hypothetical example; there is no known way to perform such a measurement), then the probability would equal:

$$P_m r dr = r dr \int \psi_m^*(\phi) \psi_m^*(r, \phi) \psi(r, \phi) \psi(\phi) d\phi d\phi' = |\langle m | \psi \rangle|^2 r dr.$$

6. Two or more properties F,G, J whose corresponding Hermitian operators **F**, **G**, **J** commute

$$\mathbf{FG}-\mathbf{GF}=\mathbf{FJ}-\mathbf{JF}=\mathbf{GJ}-\mathbf{JG}=0$$

have complete sets of simultaneous eigenfunctions (the proof of this is treated in Appendix C). This means that the set of functions that are eigenfunctions of one of the operators can be formed into a set of functions that are also eigenfunctions of the others:

$$\mathbf{F} \psi = f_j \psi \implies \mathbf{G} \psi = g_j \psi \implies \mathbf{J} \psi = j \psi.$$

**Example:**

The  $p_x$ ,  $p_y$  and  $p_z$  orbitals are eigenfunctions of the  $\mathbf{L}^2$  angular momentum operator with eigenvalues equal to  $L(L+1)\hbar^2 = 2\hbar^2$ . Since  $\mathbf{L}^2$  and  $\mathbf{L}_z$  commute and act on the same (angle) coordinates, they possess a complete set of simultaneous eigenfunctions.

Although the  $p_x$ ,  $p_y$  and  $p_z$  orbitals are not eigenfunctions of  $\mathbf{L}_z$ , they can be combined to form three new orbitals:  $p_0 = p_z$ ,  $p_1 = 2^{-1/2} [p_x + i p_y]$ , and  $p_{-1} = 2^{-1/2} [p_x - i p_y]$  that are still eigenfunctions of  $\mathbf{L}^2$  but are now eigenfunctions of  $\mathbf{L}_z$  also (with eigenvalues  $0\hbar$ ,  $1\hbar$ , and  $-1\hbar$ , respectively).

It should be mentioned that if two operators do not commute, they may still have some eigenfunctions in common, but they will not have a complete set of simultaneous eigenfunctions. For example, the  $\mathbf{L}_z$  and  $\mathbf{L}_x$  components of the angular momentum operator do not commute; however, a wavefunction with  $L=0$  (i.e., an S-state) is an eigenfunction of both operators.

The fact that two operators commute is of great importance. It means that once a measurement of one of the properties is carried out, subsequent measurement of that property or of any of the other properties corresponding to mutually commuting operators can be made without altering the system's value of the properties measured earlier. Only subsequent measurement of another property whose operator does not commute with  $\mathbf{F}$ ,  $\mathbf{G}$ , or  $\mathbf{J}$  will destroy precise knowledge of the values of the properties measured earlier.

**Example:**

Assume that an experiment has been carried out on an atom to measure its total angular momentum  $L^2$ . According to quantum mechanics, only values equal to  $L(L+1)\hbar^2$  will be observed. Further assume, for the particular experimental sample subjected to observation, that values of  $L^2$  equal to  $2\hbar^2$  and  $0\hbar^2$  were detected in relative amounts of 64 % and 36 %, respectively. This means that the atom's original wavefunction could be represented as:

$$\psi = 0.8 P + 0.6 S,$$

where  $P$  and  $S$  represent the  $P$ -state and  $S$ -state components of  $\psi$ . The squares of the amplitudes 0.8 and 0.6 give the 64 % and 36 % probabilities mentioned above.

Now assume that a subsequent measurement of the component of angular momentum along the lab-fixed  $z$ -axis is to be measured for that sub-population of the original sample found to be in the  $P$ -state. For that population, the wavefunction is now a pure  $P$ -function:

$$\psi' = P.$$

However, at this stage we have no information about how much of this  $\psi'$  is of  $m = 1, 0$ , or  $-1$ , nor do we know how much  $2p, 3p, 4p, \dots np$  components this state contains.

Because the property corresponding to the operator  $L_z$  is about to be measured, we express the above  $\psi'$  in terms of the eigenfunctions of  $L_z$ :

$$\psi' = P = \sum_{m=1,0,-1} C'_m P_m.$$

When the measurement of  $L_z$  is made, the values  $1\hbar, 0\hbar$ , and  $-1\hbar$  will be observed with probabilities given by  $|C'_1|^2, |C'_0|^2$ , and  $|C'_{-1}|^2$ , respectively. For that sub-population found to have, for example,  $L_z$  equal to  $-1\hbar$ , the wavefunction then becomes

$$\psi'' = P_{-1}.$$

At this stage, we do not know how much of  $2p_{-1}, 3p_{-1}, 4p_{-1}, \dots np_{-1}$  this wavefunction contains. To probe this question another subsequent measurement of the energy (corresponding to the  $H$  operator) could be made. Doing so would allow the amplitudes in the expansion of the above  $\psi'' = P_{-1}$

$$\psi = \sum_n C_n \psi_n$$

to be found.

The kind of experiment outlined above allows one to find the content of each particular component of an initial sample's wavefunction. For example, the original wavefunction has

$0.64 |C_n|^2 / |C_m|^2$  fractional content of the various  $\psi_n$  functions. It is analogous to the other examples considered above because all of the operators whose properties are measured commute.

### Another Example:

Let us consider an experiment in which we begin with a sample (with wavefunction  $\psi$ ) that is first subjected to measurement of  $L_z$  and then subjected to measurement of  $L^2$  and then of the energy. In this order, one would first find specific values (integer multiples of  $\hbar$ ) of  $L_z$  and one would express  $\psi$  as

$$\psi = \sum_m D_m \psi_m$$

At this stage, the nature of each  $\psi_m$  is unknown (e.g., the  $\psi_1$  function can contain  $\psi_{1,1}$ ,  $\psi_{1,0}$ ,  $\psi_{1,-1}$ , etc. components); all that is known is that  $\psi_m$  has  $m\hbar$  as its  $L_z$  value.

Taking that sub-population ( $|D_m|^2$  fraction) with a particular  $m\hbar$  value for  $L_z$  and subjecting it to subsequent measurement of  $L^2$  requires the current wavefunction  $\psi_m$  to be expressed as

$$\psi_m = \sum_L D_{L,m} \psi_{L,m}$$

When  $L^2$  is measured the value  $L(L+1)\hbar^2$  will be observed with probability  $|D_{L,m}|^2$ , and the wavefunction for that particular sub-population will become

$$\psi = \sum_{L,m} D_{L,m} \psi_{L,m}$$

At this stage, we know the value of  $L$  and of  $m$ , but we do not know the energy of the state. For example, we may know that the present sub-population has  $L=1$ ,  $m=-1$ , but we have no knowledge (yet) of how much  $2p_{-1}$ ,  $3p_{-1}$ , ...  $np_{-1}$  the system contains.

To further probe the sample, the above sub-population with  $L=1$  and  $m=-1$  can be subjected to measurement of the energy. In this case, the function  $\psi_{1,-1}$  must be expressed as

$$\psi_{1,-1} = \sum_n D_n'' \psi_{n,-1}.$$

When the energy measurement is made, the state  $\psi_{n,-1}$  will be found  $|D_n''|^2$  fraction of the time.

The fact that  $L_z$ ,  $L^2$ , and  $H$  all commute with one another (i.e., are mutually commutative) makes the series of measurements described in the above examples more straightforward than if these operators did not commute.

In the first experiment, the fact that they are mutually commutative allowed us to expand the 64 % probable  $L^2$  eigenstate with  $L=1$  in terms of functions that were eigenfunctions of the operator for which measurement was about to be made without destroying our knowledge of the value of  $L^2$ . That is, because  $L^2$  and  $L_z$  can have simultaneous eigenfunctions, the  $L = 1$  function can be expanded in terms of functions that are eigenfunctions of both  $L^2$  and  $L_z$ . This in turn, allowed us to find experimentally the sub-population that had, for example  $-1 \hbar$  as its value of  $L_z$  while retaining knowledge that the state remains an eigenstate of  $L^2$  (the state at this time had  $L = 1$  and  $m = -1$  and was denoted  $\psi_{1,-1}$ ). Then, when this  $\psi_{1,-1}$  state was subjected to energy measurement, knowledge of the energy of the sub-population could be gained without giving up knowledge of the  $L^2$  and  $L_z$  information; upon carrying out said measurement, the state became  $\psi_{1,-1}$ .

We therefore conclude that the act of carrying out an experimental measurement disturbs the system in that it causes the system's wavefunction to become an eigenfunction of the operator whose property is measured. If two properties whose corresponding operators commute are measured, the measurement of the second property does not destroy knowledge of the first property's value gained in the first measurement.

On the other hand, as detailed further in Appendix C, if the two properties ( $F$  and  $G$ ) do not commute, the second measurement destroys knowledge of the first property's value. After the first measurement,  $\psi$  is an eigenfunction of  $F$ ; after the second measurement, it becomes an eigenfunction of  $G$ . If the two non-commuting operators' properties are measured in the opposite order, the wavefunction first is an eigenfunction of  $G$ , and subsequently becomes an eigenfunction of  $F$ .

It is thus often said that 'measurements for operators that do not commute interfere with one another'. The simultaneous measurement of the position and momentum along the

same axis provides an example of two measurements that are incompatible. The fact that  $x$  and  $p_x = -i\hbar \frac{d}{dx}$  do not commute is straightforward to demonstrate:

$$\{x(-i\hbar \frac{d}{dx}) - (-i\hbar \frac{d}{dx})x\} = i\hbar \neq 0.$$

Operators that commute with the Hamiltonian and with one another form a particularly important class because each such operator permits each of the energy eigenstates of the system to be labelled with a corresponding quantum number. These operators are called **symmetry operators**. As will be seen later, they include angular momenta (e.g.,  $L^2, L_z, S^2, S_z$ , for atoms) and point group symmetries (e.g., planes and rotations about axes). Every operator that qualifies as a symmetry operator provides a quantum number with which the energy levels of the system can be labeled.

7. If a property  $F$  is measured for a large number of systems all described by the same  $\psi$ , the average value  $\langle F \rangle$  of  $F$  for such a set of measurements can be computed as

$$\langle F \rangle = \langle \psi | \mathbf{F} | \psi \rangle.$$

Expanding  $\psi$  in terms of the complete set of eigenstates of  $\mathbf{F}$  allows  $\langle F \rangle$  to be rewritten as follows:

$$\langle F \rangle = \sum_j f_j |\langle \psi | j \rangle|^2,$$

which clearly expresses  $\langle F \rangle$  as the product of the probability  $P_j$  of obtaining the particular value  $f_j$  when the property  $F$  is measured and the value  $f_j$  of the property in such a measurement. This same result can be expressed in terms of the density matrix  $D_{i,j}$  of the state  $\psi$  defined above as:

$$\begin{aligned} \langle F \rangle &= \sum_{i,j} \langle \psi | i \rangle \langle i | \mathbf{F} | j \rangle \langle j | \psi \rangle = \sum_{i,j} C_i^* \langle i | \mathbf{F} | j \rangle C_j \\ &= \sum_{i,j} D_{j,i} \langle i | \mathbf{F} | j \rangle = \text{Tr}(\mathbf{D}\mathbf{F}). \end{aligned}$$

Here,  $\mathbf{D}\mathbf{F}$  represents the matrix product of the density matrix  $D_{j,i}$  and the matrix representation  $F_{i,j} = \langle i | \mathbf{F} | j \rangle$  of the  $\mathbf{F}$  operator, both taken in the  $\{|j\rangle\}$  basis, and  $\text{Tr}$  represents the matrix trace operation.

As mentioned at the beginning of this Section, this set of rules and their relationships to experimental measurements can be quite perplexing. The structure of quantum mechanics embodied in the above rules was developed in light of new scientific observations (e.g., the photoelectric effect, diffraction of electrons) that could not be interpreted within the conventional pictures of classical mechanics. Throughout its development, these and other experimental observations placed severe constraints on the structure of the equations of the new quantum mechanics as well as on their interpretations. For example, the observation of discrete lines in the emission spectra of atoms gave rise to the idea that the atom's electrons could exist with only certain discrete energies and that light of specific frequencies would be given off as transitions among these quantized energy states took place.

Even with the assurance that quantum mechanics has firm underpinnings in experimental observations, students learning this subject for the first time often encounter difficulty. Therefore, it is useful to again examine some of the model problems for which the Schrödinger equation can be exactly solved and to learn how the above rules apply to such concrete examples.

The examples examined earlier in this Chapter and those given in the Exercises and Problems serve as useful models for chemically important phenomena: electronic motion in polyenes, in solids, and in atoms as well as vibrational and rotational motions. Their study thus far has served two purposes; it allowed the reader to gain some familiarity with applications of quantum mechanics and it introduced models that play central roles in much of chemistry. Their study now is designed to illustrate how the above seven rules of quantum mechanics relate to experimental reality.

## B. An Example Illustrating Several of the Fundamental Rules

The physical significance of the time independent wavefunctions and energies treated in Section II as well as the meaning of the seven fundamental points given above can be further illustrated by again considering the simple two-dimensional electronic motion model.

If the electron were prepared in the eigenstate corresponding to  $n_x = 1$ ,  $n_y = 2$ , its total energy would be

$$E = \frac{\hbar^2}{2m} \left[ \frac{1^2}{L_x^2} + \frac{2^2}{L_y^2} \right].$$

If the energy were experimentally measured, this and only this value would be observed, and this same result would hold for all time as long as the electron is undisturbed.

If an experiment were carried out to measure the momentum of the electron along the y-axis, according to the second postulate above, only values equal to the eigenvalues of  $-\hbar^2 \nabla^2 / y$  could be observed. The  $p_y$  eigenfunctions (i.e., functions that obey  $p_y \psi = -\hbar^2 \nabla^2 / y \psi = c \psi$ ) are of the form

$$(1/L_y)^{1/2} \exp(ik_y y),$$

where the momentum  $\hbar k_y$  can achieve any value; the  $(1/L_y)^{1/2}$  factor is used to normalize the eigenfunctions over the range  $0 \leq y \leq L_y$ . It is useful to note that the y-dependence of  $\psi$  as expressed above [ $\exp(i2\pi y/L_y) - \exp(-i2\pi y/L_y)$ ] is already written in terms of two such eigenstates of  $-\hbar^2 \nabla^2 / y$ :

$$-\hbar^2 \nabla^2 / y \exp(i2\pi y/L_y) = 2\hbar^2 \pi^2 / L_y \exp(i2\pi y/L_y), \text{ and}$$

$$-\hbar^2 \nabla^2 / y \exp(-i2\pi y/L_y) = -2\hbar^2 \pi^2 / L_y \exp(-i2\pi y/L_y).$$

Thus, the expansion of  $\psi$  in terms of eigenstates of the property being measured dictated by the fifth postulate above is already accomplished. The only two terms in this expansion correspond to momenta along the y-axis of  $2\hbar^2 \pi^2 / L_y$  and  $-2\hbar^2 \pi^2 / L_y$ ; the probabilities of observing these two momenta are given by the squares of the expansion coefficients of  $\psi$  in terms of the normalized eigenfunctions of  $-\hbar^2 \nabla^2 / y$ . The functions  $(1/L_y)^{1/2} \exp(i2\pi y/L_y)$  and

$(1/L_y)^{1/2} \exp(-i2\pi y/L_y)$  are such normalized eigenfunctions; the expansion coefficients of these functions in  $\psi$  are  $2^{-1/2}$  and  $-2^{-1/2}$ , respectively. Thus the momentum  $2\hbar^2 \pi^2 / L_y$  will be observed with probability  $(2^{-1/2})^2 = 1/2$  and  $-2\hbar^2 \pi^2 / L_y$  will be observed with probability  $(-2^{-1/2})^2 = 1/2$ . If the momentum along the x-axis were experimentally measured, again only two values  $\hbar^2 \pi^2 / L_x$  and  $-\hbar^2 \pi^2 / L_x$  would be found, each with a probability of 1/2.

The average value of the momentum along the x-axis can be computed either as the sum of the probabilities multiplied by the momentum values:

$$\langle p_x \rangle = 1/2 [\hbar^2 \pi^2 / L_x - \hbar^2 \pi^2 / L_x] = 0,$$

or as the so-called expectation value integral shown in the seventh postulate:

$$\langle p_x \rangle = \int \psi^* (-i\hbar \frac{\partial}{\partial x}) \psi dx dy.$$

Inserting the full expression for  $\psi(x,y)$  and integrating over  $x$  and  $y$  from 0 to  $L_x$  and  $L_y$ , respectively, this integral is seen to vanish. This means that the result of a large number of measurements of  $p_x$  on electrons each described by the same  $\psi$  will yield zero net momentum along the  $x$ -axis.; half of the measurements will yield positive momenta and half will yield negative momenta of the same magnitude.

The time evolution of the full wavefunction given above for the  $n_x=1, n_y=2$  state is easy to express because this  $\psi$  is an energy eigenstate:

$$\psi(x,y,t) = \psi(x,y) \exp(-iEt/\hbar).$$

If, on the other hand, the electron had been prepared in a state  $\psi(x,y)$  that is not a pure eigenstate (i.e., cannot be expressed as a single energy eigenfunction), then the time evolution is more complicated. For example, if at  $t=0$   $\psi$  were of the form

$$\begin{aligned} \psi &= (2/L_x)^{1/2} (2/L_y)^{1/2} [a \sin(2\pi x/L_x) \sin(\pi y/L_y) \\ &+ b \sin(\pi x/L_x) \sin(2\pi y/L_y) ], \end{aligned}$$

with  $a$  and  $b$  both real numbers whose squares give the probabilities of finding the system in the respective states, then the time evolution operator  $\exp(-i\mathbf{H}t/\hbar)$  applied to  $\psi$  would yield the following time dependent function:

$$\begin{aligned} \psi &= (2/L_x)^{1/2} (2/L_y)^{1/2} [a \exp(-iE_{2,1} t/\hbar) \sin(2\pi x/L_x) \\ &\sin(\pi y/L_y) + b \exp(-iE_{1,2} t/\hbar) \sin(\pi x/L_x) \sin(2\pi y/L_y) ], \end{aligned}$$

where

$$E_{2,1} = \frac{\hbar^2}{2m} [ \frac{4\pi^2}{L_x^2} + \frac{\pi^2}{L_y^2} ], \text{ and}$$

$$E_{1,2} = \frac{\hbar^2}{2m} [ \frac{\pi^2}{L_x^2} + \frac{4\pi^2}{L_y^2} ].$$

The probability of finding  $E_{2,1}$  if an experiment were carried out to measure energy would be  $|a \exp(-iE_{2,1} t/\hbar)|^2 = |a|^2$ ; the probability for finding  $E_{1,2}$  would be  $|b|^2$ . The spatial probability distribution for finding the electron at points  $x,y$  will, in this case, be given by:

$$|\psi\rangle^2 = |a|^2 |\psi_{2,1}\rangle^2 + |b|^2 |\psi_{1,2}\rangle^2 + 2ab \psi_{2,1} \psi_{1,2} \cos(Et/\hbar),$$

where  $E$  is  $E_{2,1} - E_{1,2}$ ,

$$\psi_{2,1} = (2/L_x)^{1/2} (2/L_y)^{1/2} \sin(2x/L_x) \sin(y/L_y),$$

and

$$\psi_{1,2} = (2/L_x)^{1/2} (2/L_y)^{1/2} \sin(x/L_x) \sin(2y/L_y).$$

This spatial distribution is not stationary but evolves in time. So in this case, one has a wavefunction that is not a pure eigenstate of the Hamiltonian (one says that  $\psi$  is a superposition state or a non-stationary state) whose average energy remains constant ( $E = E_{2,1} |a|^2 + E_{1,2} |b|^2$ ) but whose spatial distribution changes with time.

Although it might seem that most spectroscopic measurements would be designed to prepare the system in an eigenstate (e.g., by focusing on the sample light whose frequency matches that of a particular transition), such need not be the case. For example, if very short laser pulses are employed, the Heisenberg uncertainty broadening ( $\Delta E \Delta t \sim \hbar$ ) causes the light impinging on the sample to be very non-monochromatic (e.g., a pulse time of  $1 \times 10^{-12}$  sec corresponds to a frequency spread of approximately  $5 \text{ cm}^{-1}$ ). This, in turn, removes any possibility of preparing the system in a particular quantum state with a resolution of better than  $30 \text{ cm}^{-1}$  because the system experiences time oscillating electromagnetic fields whose frequencies range over at least  $5 \text{ cm}^{-1}$ ).

*Essentially all of the model problems that have been introduced in this Chapter to illustrate the application of quantum mechanics constitute widely used, highly successful 'starting-point' models for important chemical phenomena. As such, it is important that students retain working knowledge of the energy levels, wavefunctions, and symmetries that pertain to these models.*

Thus far, exactly soluble model problems that represent one or more aspects of an atom or molecule's quantum-state structure have been introduced and solved. For example, electronic motion in polyenes was modeled by a particle-in-a-box. The harmonic oscillator and rigid rotor were introduced to model vibrational and rotational motion of a diatomic molecule.

As chemists, we are used to thinking of electronic, vibrational, rotational, and translational energy levels as being (at least approximately) separable. On the other hand, we are aware that situations exist in which energy can flow from one such degree of freedom to another (e.g., electronic-to-vibrational energy flow occurs in radiationless relaxation and vibration-rotation couplings are important in molecular spectroscopy). It is important to understand how the simplifications that allow us to focus on electronic or vibrational or rotational motion arise, how they can be obtained from a first-principles derivation, and what their limitations and range of accuracy are.

## Chapter 2

### *Approximation Methods Can be Used When Exact Solutions to the Schrödinger Equation Can Not be Found.*

In applying quantum mechanics to 'real' chemical problems, one is usually faced with a Schrödinger differential equation for which, to date, no one has found an analytical solution. This is equally true for electronic and nuclear-motion problems. It has therefore proven essential to develop and efficiently implement mathematical methods which can provide approximate solutions to such eigenvalue equations. Two methods are widely used in this context- the variational method and perturbation theory. These tools, whose use permeates virtually all areas of theoretical chemistry, are briefly outlined here, and the details of perturbation theory are amplified in Appendix D.

#### I. The Variational Method

For the kind of potentials that arise in atomic and molecular structure, the Hamiltonian  $H$  is a Hermitian operator that is bounded from below (i.e., it has a lowest eigenvalue). Because it is Hermitian, it possesses a complete set of orthonormal eigenfunctions  $\{ \psi_j \}$ . Any function  $\psi$  that depends on the same spatial and spin variables on which  $H$  operates and obeys the same boundary conditions that the  $\{ \psi_j \}$  obey can be expanded in this complete set

$$\psi = \sum_j C_j \psi_j.$$

The expectation value of the Hamiltonian for any such function can be expressed in terms of its  $C_j$  coefficients and the exact energy levels  $E_j$  of  $H$  as follows:

$$\langle \psi | H | \psi \rangle = \sum_{ij} C_i C_j \langle \psi_i | H | \psi_j \rangle = \sum_j |C_j|^2 E_j.$$

If the function  $\psi$  is normalized, the sum  $\sum_j |C_j|^2$  is equal to unity. Because  $H$  is bounded from below, all of the  $E_j$  must be greater than or equal to the lowest energy  $E_0$ . Combining the latter two observations allows the energy expectation value of  $\psi$  to be used to produce a very important inequality:

$$\langle \psi | H | \psi \rangle \geq E_0.$$

The equality can hold only if  $\psi$  is equal to  $\psi_0$ ; if  $\psi$  contains components along any of the other  $\psi_j$ , the energy of  $\psi$  will exceed  $E_0$ .

This upper-bound property forms the basis of the so-called variational method in which 'trial wavefunctions' are constructed:

i. To guarantee that  $\psi$  obeys all of the boundary conditions that the exact  $\psi_j$  do and that  $\psi$  is of the proper spin and space symmetry and is a function of the same spatial and spin coordinates as the  $\psi_j$ ;

ii. With parameters embedded in  $\psi$  whose 'optimal' values are to be determined by making  $\langle \psi | H | \psi \rangle$  a minimum.

It is perfectly acceptable to vary any parameters in  $\psi$  to attain the lowest possible value for  $\langle \psi | H | \psi \rangle$  because the proof outlined above constrains this expectation value to be above the true lowest eigenstate's energy  $E_0$  for any  $\psi$ . The philosophy then is that the  $\psi$  that gives the lowest  $\langle \psi | H | \psi \rangle$  is the best because its expectation value is closest to the exact energy.

Quite often a trial wavefunction is expanded as a linear combination of other functions

$$\psi = \sum_J C_J \psi_J.$$

In these cases, one says that a 'linear variational' calculation is being performed. The set of functions  $\{\psi_J\}$  are usually constructed to obey all of the boundary conditions that the exact state  $\psi_0$  obeys, to be functions of the the same coordinates as  $\psi_0$ , and to be of the same spatial and spin symmetry as  $\psi_0$ . Beyond these conditions, the  $\{\psi_J\}$  are nothing more than members of a set of functions that are convenient to deal with (e.g., convenient to evaluate Hamiltonian matrix elements  $\langle \psi_I | H | \psi_J \rangle$ ) and that can, in principle, be made complete if more and more such functions are included.

For such a trial wavefunction, the energy depends quadratically on the 'linear variational'  $C_J$  coefficients:

$$\langle \psi | H | \psi \rangle = \sum_{IJ} C_I C_J \langle \psi_I | H | \psi_J \rangle.$$

Minimization of this energy with the constraint that  $\psi$  remain normalized ( $\langle \psi | \psi \rangle = 1 = \sum_{IJ} C_I C_J \langle \psi_I | \psi_J \rangle$ ) gives rise to a so-called secular or eigenvalue-eigenvector problem:

$$\sum_J [\langle \psi_I | H | \psi_J \rangle - E \langle \psi_I | \psi_J \rangle] C_J = \sum_J [H_{IJ} - E S_{IJ}] C_J = 0.$$

If the functions  $\{\psi_J\}$  are orthonormal, then the overlap matrix  $S$  reduces to the unit matrix and the above generalized eigenvalue problem reduces to the more familiar form:

$$\sum_J H_{IJ} C_J = E C_I.$$

The secular problem, in either form, has as many eigenvalues  $E_i$  and eigenvectors  $\{C_{iJ}\}$  as the dimension of the  $H_{IJ}$  matrix as  $N$ . It can also be shown that between successive pairs of the eigenvalues obtained by solving the secular problem at least one exact eigenvalue must occur (i.e.,  $E_{i+1} > E_{\text{exact}} > E_i$ , for all  $i$ ). This observation is referred to as 'the bracketing theorem'.

Variational methods, in particular the linear variational method, are the most widely used approximation techniques in quantum chemistry. To implement such a method one needs to know the Hamiltonian  $H$  whose energy levels are sought and one needs to construct a trial wavefunction in which some 'flexibility' exists (e.g., as in the linear variational method where the  $C_J$  coefficients can be varied). In **Section 6** this tool will be used to develop several of the most commonly used and powerful molecular orbital methods in chemistry.

## II. Perturbation Theory

[Suggested Extra Reading- Appendix D; Time Independent Perturbation Theory]

Perturbation theory is the second most widely used approximation method in quantum chemistry. It allows one to estimate the splittings and shifts in energy levels and changes in wavefunctions that occur when an external field (e.g., an electric or magnetic field or a field that is due to a surrounding set of 'ligands'- a crystal field) or a field arising when a previously-ignored term in the Hamiltonian is applied to a species whose 'unperturbed' states are known. These 'perturbations' in energies and wavefunctions are expressed in terms of the (complete) set of unperturbed eigenstates.

Assuming that all of the wavefunctions  $\psi_k$  and energies  $E_k^0$  belonging to the unperturbed Hamiltonian  $H^0$  are known

$$H^0 \psi_k = E_k^0 \psi_k,$$

and given that one wishes to find eigenstates ( $\psi_k$  and  $E_k$ ) of the perturbed Hamiltonian

$$H = H^0 + V,$$

perturbation theory expresses  $\psi_k$  and  $E_k$  as power series in the perturbation strength  $\lambda$ :

$$\psi_k = \sum_{n=0}^{\infty} \lambda^n \psi_k^{(n)}$$

$$E_k = \sum_{n=0}^{\infty} \lambda^n E_k^{(n)}.$$

The systematic development of the equations needed to determine the  $E_k^{(n)}$  and the  $\psi_k^{(n)}$  is presented in Appendix D. Here, we simply quote the few lowest-order results.

The zeroth-order wavefunctions and energies are given in terms of the solutions of the unperturbed problem as follows:

$$\psi_k^{(0)} = \psi_k \text{ and } E_k^{(0)} = E_k^0.$$

This simply means that one must be willing to identify one of the unperturbed states as the 'best' approximation to the state being sought. This, of course, implies that one must therefore strive to find an unperturbed model problem, characterized by  $H^0$  that represents the true system as accurately as possible, so that one of the  $\psi_k$  will be as close as possible to  $\psi_k$ .

The first-order energy correction is given in terms of the zeroth-order (i.e., unperturbed) wavefunction as:

$$E_k^{(1)} = \langle \psi_k | V | \psi_k \rangle,$$

which is identified as the average value of the perturbation taken with respect to the unperturbed function  $\psi_k$ . The so-called first-order wavefunction  $\psi_k^{(1)}$  expressed in terms of the complete set of unperturbed functions  $\{\psi_j\}$  is:

$$\psi_k^{(1)} = \sum_{j \neq k} \frac{\langle \psi_j | V | \psi_k \rangle}{E_k^0 - E_j^0} \psi_j.$$

The second-order energy correction is expressed as follows:

$$E_k^{(2)} = \sum_{j \neq k} \frac{|\langle j | V | k \rangle|^2}{E_k^{(0)} - E_j^{(0)}},$$

and the second-order correction to the wavefunction is expressed as

$$\psi_k^{(2)} = \sum_{j \neq k} \frac{\langle j | V | k \rangle}{E_k^{(0)} - E_j^{(0)}} \psi_j^{(1)} - \frac{\langle k | V | k \rangle}{E_k^{(0)} - E_k^{(0)}} \psi_k^{(1)}$$

An essential point about perturbation theory is that the energy corrections  $E_k^{(n)}$  and wavefunction corrections  $\psi_k^{(n)}$  are expressed in terms of integrals over the unperturbed wavefunctions  $\psi_k$  involving the perturbation (i.e.,  $\langle j | V | i \rangle$ ) and the unperturbed energies  $E_j^{(0)}$ . Perturbation theory is most useful when one has, in hand, the solutions to an unperturbed Schrödinger equation that is reasonably 'close' to the full Schrödinger equation whose solutions are being sought. In such a case, it is likely that low-order corrections will be adequate to describe the energies and wavefunctions of the full problem.

It is important to stress that although the solutions to the full 'perturbed' Schrödinger equation are expressed, as above, in terms of sums over all states of the unperturbed Schrödinger equation, it is improper to speak of the perturbation as creating excited-state species. For example, the polarization of the 1s orbital of the Hydrogen atom caused by the application of a static external electric field of strength  $E$  along the z-axis is described, in first-order perturbation theory, through the sum

$$\sum_{n=2} \frac{\langle np_0 | E e r \cos \theta | 1s \rangle}{E_{1s} - E_{np_0}}$$

over all  $p_z = p_0$  orbitals labeled by principal quantum number  $n$ . The coefficient multiplying each  $p_0$  orbital depends on the energy gap corresponding to the 1s-to- $np$  'excitation' as well as the electric dipole integral  $\langle np_0 | E e r \cos \theta | 1s \rangle$  between the 1s orbital and the  $np_0$  orbital.

This sum describes the polarization of the 1s orbital in terms of functions that have  $p_0$  symmetry; by combining an s orbital and  $p_0$  orbitals, one can form a 'hybrid-like' orbital that is nothing but a distorted 1s orbital. The appearance of the excited  $np_0$  orbitals has

nothing to do with forming excited states; these  $np_0$  orbitals simply provide a set of functions that can describe the response of the  $1s$  orbital to the applied electric field.

The relative strengths and weaknesses of perturbation theory and the variational method, as applied to studies of the electronic structure of atoms and molecules, are discussed in Section 6.

## Chapter 3

*The Application of the Schrödinger Equation to the Motions of Electrons and Nuclei in a Molecule Lead to the Chemists' Picture of Electronic Energy Surfaces on Which Vibration and Rotation Occurs and Among Which Transitions Take Place.*

### I. The Born-Oppenheimer Separation of Electronic and Nuclear Motions

Many elements of chemists' pictures of molecular structure hinge on the point of view that separates the electronic motions from the vibrational/rotational motions and treats couplings between these (approximately) separated motions as 'perturbations'. It is essential to understand the origins and limitations of this separated-motions picture.

To develop a framework in terms of which to understand when such separability is valid, one thinks of an atom or molecule as consisting of a collection of  $N$  electrons and  $M$  nuclei each of which possesses kinetic energy and among which coulombic potential energies of interaction arise. To properly describe the motions of all these particles, one needs to consider the full Schrödinger equation  $H = E$ , in which the Hamiltonian  $H$  contains the sum (denoted  $H_e$ ) of the kinetic energies of all  $N$  electrons and the coulomb potential energies among the  $N$  electrons and the  $M$  nuclei as well as the kinetic energy  $T$  of the  $M$  nuclei

$$T = \sum_{a=1,M} \left( -\frac{\hbar^2}{2m_a} \right) \nabla_a^2,$$

$$H = H_e + T$$

$$H_e = \sum_j \left\{ \left( -\frac{\hbar^2}{2m_e} \right) \nabla_j^2 - \sum_a Z_a e^2 / r_{j,a} \right\} + \sum_{j < k} e^2 / r_{j,k} \\ + \sum_{a < b} Z_a Z_b e^2 / R_{a,b}.$$

Here,  $m_a$  is the mass of the nucleus  $a$ ,  $Z_a e^2$  is its charge, and  $\nabla_a^2$  is the Laplacian with respect to the three cartesian coordinates of this nucleus (this operator  $\nabla_a^2$  is given in spherical polar coordinates in Appendix A);  $r_{j,a}$  is the distance between the  $j^{\text{th}}$  electron and the  $a^{\text{th}}$  nucleus,  $r_{j,k}$  is the distance between the  $j^{\text{th}}$  and  $k^{\text{th}}$  electrons,  $m_e$  is the electron's mass, and  $R_{a,b}$  is the distance from nucleus  $a$  to nucleus  $b$ .

The full Hamiltonian  $H$  thus contains differential operators over the  $3N$  electronic coordinates (denoted  $r$  as a shorthand) and the  $3M$  nuclear coordinates (denoted  $R$  as a shorthand). In contrast, the electronic Hamiltonian  $H_e$  is a Hermitian differential operator in

r-space but not in R-space. Although  $H_e$  is indeed a function of the R-variables, it is not a differential operator involving them.

Because  $H_e$  is a Hermitian operator in r-space, its eigenfunctions  $\psi_i(r|\mathbf{R})$  obey  $H_e \psi_i(r|\mathbf{R}) = E_i(\mathbf{R}) \psi_i(r|\mathbf{R})$

for any values of the R-variables, and form a complete set of functions of r for any values of R. These eigenfunctions and their eigenvalues  $E_i(\mathbf{R})$  depend on R only because the potentials appearing in  $H_e$  depend on R. The  $\psi_i$  and  $E_i$  are the electronic wavefunctions and electronic energies whose evaluations are treated in the next three Chapters.

The fact that the set of  $\{\psi_i\}$  is, in principle, complete in r-space allows the full (electronic and nuclear) wavefunction  $\Psi$  to have its r-dependence expanded in terms of the  $\psi_i$ :

$$\Psi(r, \mathbf{R}) = \sum_i \psi_i(r|\mathbf{R}) \chi_i(\mathbf{R}).$$

The  $\chi_i(\mathbf{R})$  functions, carry the remaining R-dependence of  $\Psi$  and are determined by insisting that  $\Psi$  as expressed here obey the full Schrödinger equation:

$$(H_e + T - E) \Psi = 0.$$

Projecting this equation against  $\langle \psi_j(r|\mathbf{R}) |$  (integrating only over the electronic coordinates because the  $\psi_j$  are orthonormal only when so integrated) gives:

$$[ (E_j(\mathbf{R}) - E) \chi_j(\mathbf{R}) + T \chi_j(\mathbf{R}) ] = - \sum_i \{ \langle \psi_j | T | \psi_i \rangle (\mathbf{R}) \chi_i(\mathbf{R}) + \sum_{a=1, M} ( - \hbar^2 / m_a ) \langle \psi_j | \nabla_a | \psi_i \rangle (\mathbf{R}) \cdot \nabla_a \chi_i(\mathbf{R}) \},$$

where the  $(\mathbf{R})$  notation in  $\langle \psi_j | T | \psi_i \rangle (\mathbf{R})$  and  $\langle \psi_j | \nabla_a | \psi_i \rangle (\mathbf{R})$  has been used to remind one that the integrals  $\langle \dots \rangle$  are carried out only over the r coordinates and, as a result, still depend on the R coordinates.

In the **Born-Oppenheimer** (BO) approximation, one neglects the so-called non-adiabatic or non-BO couplings on the right-hand side of the above equation. Doing so yields the following equations for the  $\chi_i(\mathbf{R})$  functions:

$$[ (E_j(\mathbf{R}) - E) \chi_j^0(\mathbf{R}) + T \chi_j^0(\mathbf{R}) ] = 0,$$

where the superscript in  $\psi_i^0(\mathbf{R})$  is used to indicate that these functions are solutions within the BO approximation only.

These BO equations can be recognized as the equations for the translational, rotational, and vibrational motion of the nuclei on the 'potential energy surface'  $E_j(\mathbf{R})$ . That is, within the BO picture, the electronic energies  $E_j(\mathbf{R})$ , considered as functions of the nuclear positions  $\mathbf{R}$ , provide the potentials on which the nuclei move. The electronic and nuclear-motion aspects of the Schrödinger equation are thereby separated.

#### A. Time Scale Separation

*The physical parameters that determine under what circumstances the BO approximation is accurate relate to the motional time scales of the electronic and vibrational/rotational coordinates.*

The range of accuracy of this separation can be understood by considering the differences in time scales that relate to electronic motions and nuclear motions under ordinary circumstances. In most atoms and molecules, the electrons orbit the nuclei at speeds much in excess of even the fastest nuclear motions (the vibrations). As a result, the electrons can adjust 'quickly' to the slow motions of the nuclei. This means it should be possible to develop a model in which the electrons 'follow' smoothly as the nuclei vibrate and rotate.

This picture is that described by the BO approximation. Of course, one should expect large corrections to such a model for electronic states in which 'loosely held' electrons exist. For example, in molecular Rydberg states and in anions, where the outer valence electrons are bound by a fraction of an electron volt, the natural orbit frequencies of these electrons are not much faster (if at all) than vibrational frequencies. In such cases, significant breakdown of the BO picture is to be expected.

#### B. Vibration/Rotation States for Each Electronic Surface

*The BO picture is what gives rise to the concept of a manifold of potential energy surfaces on which vibrational/rotational motions occur.*

Even within the BO approximation, motion of the nuclei on the various electronic energy surfaces is different because the nature of the chemical bonding differs from surface to surface. That is, the vibrational/rotational motion on the ground-state surface is certainly

not the same as on one of the excited-state surfaces. However, there are a complete set of wavefunctions  $\psi_{j,m}^0(\mathbf{R})$  and energy levels  $E_{j,m}^0$  for each surface  $E_j(\mathbf{R})$  because  $T + E_j(\mathbf{R})$  is a Hermitian operator in  $\mathbf{R}$ -space for each surface (labelled  $j$ ):

$$[T + E_j(\mathbf{R})] \psi_{j,m}^0(\mathbf{R}) = E_{j,m}^0 \psi_{j,m}^0.$$

The eigenvalues  $E_{j,m}^0$  must be labelled by the electronic surface ( $j$ ) on which the motion occurs as well as to denote the particular state ( $m$ ) on that surface.

## II. Rotation and Vibration of Diatomic Molecules

For a diatomic species, the vibration-rotation (V/R) kinetic energy operator can be expressed as follows in terms of the bond length  $R$  and the angles  $\theta$  and  $\phi$  that describe the orientation of the bond axis relative to a laboratory-fixed coordinate system:

$$T_{V/R} = -\frac{\hbar^2}{2\mu} \left\{ \frac{1}{R^2} \frac{\partial}{\partial R} \left( R^2 \frac{\partial}{\partial R} \right) - \frac{L^2}{R^2} \right\},$$

where the square of the rotational angular momentum of the diatomic species is

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

Because the potential  $E_j(\mathbf{R})$  depends on  $R$  but not on  $\theta$  or  $\phi$ , the V/R function  $\psi_{j,m}^0$  can be written as a product of an angular part and an  $R$ -dependent part; moreover, because  $L^2$  contains the full angle-dependence of  $T_{V/R}$ ,  $\psi_{j,n}^0$  can be written as

$$\psi_{j,n}^0 = Y_{J,M}(\theta, \phi) F_{j,J,v}(R).$$

The general subscript  $n$ , which had represented the state in the full set of  $3M-3$   $\mathbf{R}$ -space coordinates, is replaced by the three quantum numbers  $J, M$ , and  $v$  (i.e., once one focuses on the three specific coordinates  $R$ ,  $\theta$ , and  $\phi$ , a total of three quantum numbers arise in place of the symbol  $n$ ).

Substituting this product form for  $\psi_{j,n}^0$  into the V/R equation gives:

$$-\frac{\hbar^2}{2\mu} \left\{ \frac{1}{R^2} \frac{\partial}{\partial R} \left( R^2 \frac{\partial}{\partial R} \right) - \frac{\hbar^2 J(J+1)}{R^2} \right\} F_{j,J,v}(R)$$

$$+ E_j(R) F_{j,J,v}(R) = E_{j,J,v}^0 F_{j,J,v}$$

as the equation for the vibrational (i.e., R-dependent) wavefunction within electronic state j and with the species rotating with  $J(J+1) \hbar^2$  as the square of the total angular momentum and a projection along the laboratory-fixed Z-axis of  $M\hbar$ . The fact that the  $F_{j,J,v}$  functions do not depend on the M quantum number derives from the fact that the  $T_{V/R}$  kinetic energy operator does not explicitly contain  $J_Z$ ; only  $J^2$  appears in  $T_{V/R}$ .

The solutions for which  $J=0$  correspond to vibrational states in which the species has no rotational energy; they obey

$$- \hbar^2/2\mu \left\{ \frac{d^2}{dR^2} + \frac{J(J+1)\hbar^2}{2\mu R^2} \right\} F_{j,0,v}(R)$$

$$+ E_j(R) F_{j,0,v}(R) = E_{j,0,v}^0 F_{j,0,v}$$

The differential-operator parts of this equation can be simplified somewhat by substituting  $F = R^{-1} \psi$  and thus obtaining the following equation for the new function  $\psi$ :

$$- \hbar^2/2\mu \left\{ \frac{d^2}{dR^2} + \frac{J(J+1)\hbar^2}{2\mu R^2} \right\} \psi_{j,0,v}(R) + E_j(R) \psi_{j,0,v}(R) = E_{j,0,v}^0 \psi_{j,0,v}$$

Solutions for which  $J \neq 0$  require the vibrational wavefunction and energy to respond to the presence of the 'centrifugal potential' given by  $\hbar^2 J(J+1)/(2\mu R^2)$ ; these solutions obey the full coupled V/R equations given above.

#### A. Separation of Vibration and Rotation

It is common, in developing the working equations of diatomic-molecule rotational/vibrational spectroscopy, to treat the coupling between the two degrees of freedom using perturbation theory as developed later in this chapter. In particular, one can expand the centrifugal coupling  $\hbar^2 J(J+1)/(2\mu R^2)$  around the equilibrium geometry  $R_e$  (which depends, of course, on j):

$$\hbar^2 J(J+1)/(2\mu R^2) = \hbar^2 J(J+1)/(2\mu [R_e^2 (1 + \delta R/R_e)^2])$$

$$= \hbar^2 J(J+1)/(2\mu R_e^2) [1 - 2 \delta R/R_e + \dots],$$

and treat the terms containing powers of the bond length displacement  $\delta R^k$  as perturbations. The zeroth-order equations read:

$$-\frac{\hbar^2}{2\mu} \left\{ \frac{d^2}{dR^2} + \frac{J(J+1)}{R^2} \right\} F_{j,J,v}^0(R) + E_j(R) F_{j,J,v}^0(R) + \frac{\hbar^2 J(J+1)}{2\mu R_e^2} F_{j,J,v}^0 = E_{j,J,v}^0 F_{j,J,v}^0,$$

and have solutions whose energies separate

$$E_{j,J,v}^0 = \frac{\hbar^2 J(J+1)}{2\mu R_e^2} + E_{j,v}$$

and whose wavefunctions are independent of J (because the coupling is not R-dependent in zeroth order)

$$F_{j,J,v}^0(R) = F_{j,v}(R).$$

Perturbation theory is then used to express the corrections to these zeroth order solutions as indicated in Appendix D.

## B. The Rigid Rotor and Harmonic Oscillator

Treatment of the rotational motion at the zeroth-order level described above introduces the so-called 'rigid rotor' energy levels and wavefunctions:  $E_J = \frac{\hbar^2 J(J+1)}{2\mu R_e^2}$  and  $Y_{J,M}(\theta, \phi)$ ; these same quantities arise when the diatomic molecule is treated as a rigid rod of length  $R_e$ . The spacings between successive rotational levels within this approximation are

$$E_{J+1,J} = 2hcB(J+1),$$

where the so-called rotational constant B is given in  $\text{cm}^{-1}$  as

$$B = \frac{h}{8\pi^2 c\mu R_e^2}.$$

The rotational level J is  $(2J+1)$ -fold degenerate because the energy  $E_J$  is independent of the M quantum number of which there are  $(2J+1)$  values for each J:  $M = -J, -J+1, -J+2, \dots, J-2, J-1, J$ .

The explicit form of the zeroth-order vibrational wavefunctions and energy levels,  $F_{j,v}^0$  and  $E_{j,v}^0$ , depends on the description used for the electronic potential energy surface

$E_j(R)$ . In the crudest useful approximation,  $E_j(R)$  is taken to be a so-called harmonic potential

$$E_j(R) = 1/2 k_j (R-R_e)^2 ;$$

as a consequence, the wavefunctions and energy levels reduce to

$$E_{j,v}^0 = E_j(R_e) + \hbar \omega_j (v + 1/2), \text{ and}$$

$$F_{j,v}^0(R) = [2^v v! ]^{-1/2} ( \mu / \hbar )^{1/4} \exp(- (R-R_e)^2/2) H_v ( \omega_j^{1/2} (R-R_e)),$$

where  $\omega_j = (k_j \mu)^{1/2}/\hbar$  and  $H_v(y)$  denotes the Hermite polynomial defined by:

$$H_v(y) = (-1)^v \exp(y^2) d^v/dy^v \exp(-y^2).$$

The solution of the vibrational differential equation

$$-\hbar^2/2\mu \{ R^{-2} / R(R^2 / R) \} F_{j,v}(R) + E_j(R) F_{j,v}(R) = E_{j,v} F_{j,v}$$

is treated in EWK, Atkins, and McQuarrie.

These harmonic-oscillator solutions predict evenly spaced energy levels (i.e., no anharmonicity) that persist for all  $v$ . It is, of course, known that molecular vibrations display anharmonicity (i.e., the energy levels move closer together as one moves to higher  $v$ ) and that quantized vibrational motion ceases once the bond dissociation energy is reached.

### C. The Morse Oscillator

The Morse oscillator model is often used to go beyond the harmonic oscillator approximation. In this model, the potential  $E_j(R)$  is expressed in terms of the bond dissociation energy  $D_e$  and a parameter  $a$  related to the second derivative  $k$  of  $E_j(R)$  at  $R_e$   $k = (d^2E_j/dR^2) = 2a^2D_e$  as follows:

$$E_j(R) - E_j(R_e) = D_e \{ 1 - \exp(-a(R-R_e)) \}^2 .$$

The Morse oscillator energy levels are given by

$$E_{j,v}^0 = E_j(R_e) + \hbar \sqrt{k/\mu} (v+1/2) - \hbar^2/4 (k/\mu D_e) (v+1/2)^2;$$

the corresponding eigenfunctions are also known analytically in terms of hypergeometric functions (see, for example, Handbook of Mathematical Functions, M. Abramowitz and I. A. Stegun, Dover, Inc. New York, N. Y. (1964)). Clearly, the Morse solutions display anharmonicity as reflected in the negative term proportional to  $(v+1/2)^2$ .

#### D. Perturbative Treatment of Vibration-Rotation Coupling

#### III. Rotation of Polyatomic Molecules

*To describe the orientations of a diatomic or linear polyatomic molecule requires only two angles (usually termed  $\theta$  and  $\phi$ ). For any non-linear molecule, three angles (usually  $\theta$ ,  $\phi$ , and  $\chi$ ) are needed. Hence the rotational Schrödinger equation for a non-linear molecule is a differential equation in three-dimensions.*

There are  $3M-6$  vibrations of a non-linear molecule containing  $M$  atoms; a linear molecule has  $3M-5$  vibrations. The linear molecule requires two angular coordinates to describe its orientation with respect to a laboratory-fixed axis system; a non-linear molecule requires three angles.

#### A. Linear Molecules

The rotational motion of a linear polyatomic molecule can be treated as an extension of the diatomic molecule case. One obtains the  $Y_{J,M}(\theta, \phi)$  as rotational wavefunctions and, within the approximation in which the centrifugal potential is approximated at the equilibrium geometry of the molecule ( $R_e$ ), the energy levels are:

$$E_J^0 = J(J+1) \hbar^2/(2I).$$

Here the total moment of inertia  $I$  of the molecule takes the place of  $\mu R_e^2$  in the diatomic molecule case

$$I = \sum_a m_a (R_a - R_{\text{CofM}})^2;$$

$m_a$  is the mass of atom  $a$  whose distance from the center of mass of the molecule is  $(R_a - R_{\text{CofM}})$ . The rotational level with quantum number  $J$  is  $(2J+1)$ -fold degenerate again because there are  $(2J+1)$   $M$ - values.

## B. Non-Linear Molecules

For a non-linear polyatomic molecule, again with the centrifugal couplings to the vibrations evaluated at the equilibrium geometry, the following terms form the rotational part of the nuclear-motion kinetic energy:

$$T_{\text{rot}} = \sum_{i=a,b,c} (J_i^2 / 2I_i).$$

Here,  $I_i$  is the eigenvalue of the moment of inertia tensor:

$$I_{x,x} = \sum_a m_a [ (R_a - R_{\text{CofM}})^2 - (x_a - x_{\text{CofM}})^2 ]$$

$$I_{x,y} = \sum_a m_a [ (x_a - x_{\text{CofM}}) (y_a - y_{\text{CofM}}) ]$$

expressed originally in terms of the cartesian coordinates of the nuclei ( $a$ ) and of the center of mass in an arbitrary molecule-fixed coordinate system (and similarly for  $I_{z,z}$ ,  $I_{y,y}$ ,  $I_{x,z}$  and  $I_{y,z}$ ). The operator  $J_i$  corresponds to the component of the total rotational angular momentum  $\mathbf{J}$  along the direction belonging to the  $i^{\text{th}}$  eigenvector of the moment of inertia tensor.

Molecules for which all three principal moments of inertia (the  $I_i$ 's) are equal are called 'spherical tops'. For these species, the rotational Hamiltonian can be expressed in terms of the square of the total rotational angular momentum  $J^2$  :

$$T_{\text{rot}} = J^2 / 2I,$$

as a consequence of which the rotational energies once again become

$$E_J = \hbar^2 J(J+1) / 2I.$$

However, the  $Y_{J,M}$  are not the corresponding eigenfunctions because the operator  $J^2$  now contains contributions from rotations about three (no longer two) axes (i.e., the three principal axes). The proper rotational eigenfunctions are the  $D^J_{M,K}(\alpha, \beta, \gamma)$  functions known as 'rotation matrices' (see Sections 3.5 and 3.6 of Zare's book on angular momentum) these functions depend on three angles (the three Euler angles needed to describe the orientation of the molecule in space) and three quantum numbers- J, M, and K. The quantum number M labels the projection of the total angular momentum (as  $M\hbar$ ) along the laboratory-fixed z-axis;  $K\hbar$  is the projection along one of the internal principal axes ( in a spherical top molecule, all three axes are equivalent, so it does not matter which axis is chosen).

The energy levels of spherical top molecules are  $(2J+1)^2$  -fold degenerate. Both the M and K quantum numbers run from -J, in steps of unity, to J; because the energy is independent of M and of K, the degeneracy is  $(2J+1)^2$ .

Molecules for which two of the three principal moments of inertia are equal are called symmetric top molecules. Prolate symmetric tops have  $I_a < I_b = I_c$ ; oblate symmetric tops have  $I_a = I_b < I_c$  ( it is convention to order the moments of inertia as  $I_a \geq I_b \geq I_c$  ). The rotational Hamiltonian can now be written in terms of  $J^2$  and the component of  $\mathbf{J}$  along the unique moment of inertia's axis as:

$$T_{\text{rot}} = J_a^2 ( 1/2I_a - 1/2I_b ) + J^2 /2I_b$$

for prolate tops, and

$$T_{\text{rot}} = J_c^2 ( 1/2I_c - 1/2I_b ) + J^2/2I_b$$

for oblate tops. Again, the  $D^J_{M,K}(\alpha, \beta, \gamma)$  are the eigenfunctions, where the quantum number K describes the component of the rotational angular momentum  $\mathbf{J}$  along the unique molecule-fixed axis (i.e., the axis of the unique moment of inertia). The energy levels are now given in terms of J and K as follows:

$$E_{J,K} = \hbar^2 J(J+1)/2I_b + \hbar^2 K^2 (1/2I_a - 1/2I_b)$$

for prolate tops, and

$$E_{J,K} = \hbar^2 J(J+1)/2I_b + \hbar^2 K^2 (1/2I_c - 1/2I_b)$$

for oblate tops.

Because the rotational energies now depend on  $K$  (as well as on  $J$ ), the degeneracies are lower than for spherical tops. In particular, because the energies do not depend on  $M$  and depend on the square of  $K$ , the degeneracies are  $(2J+1)$  for states with  $K=0$  and  $2(2J+1)$  for states with  $|K| > 0$ ; the extra factor of 2 arises for  $|K| > 0$  states because pairs of states with  $K = |K|$  and  $K = -|K|$  are degenerate.

#### IV. Summary

This Chapter has shown how the solution of the Schrödinger equation governing the motions and interparticle potential energies of the nuclei and electrons of an atom or molecule (or ion) can be decomposed into two distinct problems: (i) solution of the electronic Schrödinger equation for the electronic wavefunctions and energies, both of which depend on the nuclear geometry and (ii) solution of the vibration/rotation Schrödinger equation for the motion of the nuclei on any one of the electronic energy surfaces. This decomposition into approximately separable electronic and nuclear-motion problems remains an important point of view in chemistry. It forms the basis of many of our models of molecular structure and our interpretation of molecular spectroscopy. It also establishes how we approach the computational simulation of the energy levels of atoms and molecules; we first compute electronic energy levels at a 'grid' of different positions of the nuclei, and we then solve for the motion of the nuclei on a particular energy surface using this grid of data.

The treatment of electronic motion is treated in detail in Sections 2, 3, and 6 where molecular orbitals and configurations and their computer evaluation is covered. The vibration/rotation motion of molecules on BO surfaces is introduced above, but should be treated in more detail in a subsequent course in molecular spectroscopy.

#### *Section Summary*

This Introductory Section was intended to provide the reader with an overview of the structure of quantum mechanics and to illustrate its application to several exactly solvable model problems. The model problems analyzed play especially important roles in chemistry because they form the basis upon which more sophisticated descriptions of the electronic structure and rotational-vibrational motions of molecules are built. The variational method and perturbation theory constitute the tools needed to make use of solutions of

simpler model problems as starting points in the treatment of Schrödinger equations that are impossible to solve analytically.

In Sections 2, 3, and 6 of this text, the electronic structures of polyatomic molecules, linear molecules, and atoms are examined in some detail. Symmetry, angular momentum methods, wavefunction antisymmetry, and other tools are introduced as needed throughout the text. The application of modern computational chemistry methods to the treatment of molecular electronic structure is included. Given knowledge of the electronic energy surfaces as functions of the internal geometrical coordinates of the molecule, it is possible to treat vibrational-rotational motion on these surfaces. Exercises, problems, and solutions are provided for each Chapter. Readers are strongly encouraged to work these exercises and problems because new material that is used in other Chapters is often developed within this context.

## Section 2 Simple Molecular Orbital Theory

In this section, the conceptual framework of molecular orbital theory is developed. Applications are presented and problems are given and solved within qualitative and semi-empirical models of electronic structure. Ab Initio approaches to these same matters, whose solutions require the use of digital computers, are treated later in Section 6. Semi-empirical methods, most of which also require access to a computer, are treated in this section and in Appendix F.

Unlike most texts on molecular orbital theory and quantum mechanics, this text treats polyatomic molecules before linear molecules before atoms. The finite point-group symmetry (Appendix E provides an introduction to the use of point group symmetry) that characterizes the orbitals and electronic states of non-linear polyatomics is more straightforward to deal with because fewer degeneracies arise. In turn, linear molecules, which belong to an axial rotation group, possess fewer degeneracies (e.g., orbitals or states are no more degenerate than  $s$ ,  $p$ , or  $d$  orbitals or states; all are doubly degenerate) than atomic orbitals and states (e.g.,  $p$  orbitals or states are 3-fold degenerate,  $d$ 's are 5-fold, etc.). Increased orbital degeneracy, in turn, gives rise to more states that can arise from a given orbital occupancy (e.g., the  $2p^2$  configuration of the C atom yields fifteen states, the  $2p^2$  configuration of the NH molecule yields six, and the  $\pi^*$  configuration of ethylene gives four states). For these reasons, it is more straightforward to treat low-symmetry cases (i.e., non-linear polyatomic molecules) first and atoms last.

It is recommended that the reader become familiar with the point-group symmetry tools developed in Appendix E before proceeding with this section. In particular, it is important to know how to label atomic orbitals as well as the various hybrids that can be formed from them according to the irreducible representations of the molecule's point group and how to construct symmetry adapted combinations of atomic, hybrid, and molecular orbitals using projection operator methods. If additional material on group theory is needed, Cotton's book on this subject is very good and provides many excellent chemical applications.

### Chapter 4

*Valence Atomic Orbitals on Neighboring Atoms Combine to Form Bonding, Non-Bonding and Antibonding Molecular Orbitals*

#### I. Atomic Orbitals

In Section 1 the Schrödinger equation for the motion of a single electron moving about a nucleus of charge  $Z$  was explicitly solved. The energies of these orbitals relative to an electron infinitely far from the nucleus with zero kinetic energy were found to depend strongly on  $Z$  and on the principal quantum number  $n$ , as were the radial "sizes" of these hydrogenic orbitals. Closed analytical expressions for the  $r$ ,  $\theta$ , and  $\phi$  dependence of these orbitals are given in Appendix B. The reader is advised to also review this material before undertaking study of this section.

### A. Shapes

*Shapes of atomic orbitals play central roles in governing the types of directional bonds an atom can form.*

All atoms have sets of bound and continuum s,p,d,f,g, etc. orbitals. Some of these orbitals may be unoccupied in the atom's low energy states, but they are still present and able to accept electron density if some physical process (e.g., photon absorption, electron attachment, or Lewis-base donation) causes such to occur. For example, the Hydrogen atom has 1s, 2s, 2p, 3s, 3p, 3d, etc. orbitals. Its negative ion  $H^-$  has states that involve 1s2s, 2p<sup>2</sup>, 3s<sup>2</sup>, 3p<sup>2</sup>, etc. orbital occupancy. Moreover, when an H atom is placed in an external electronic field, its charge density polarizes in the direction of the field. This polarization can be described in terms of the orbitals of the isolated atom being combined to yield distorted orbitals (e.g., the 1s and 2p orbitals can "mix" or combine to yield sp hybrid orbitals, one directed toward increasing field and the other directed in the opposite direction). Thus in many situations it is important to keep in mind that each atom has a full set of orbitals available to it even if some of these orbitals are not occupied in the lowest-energy state of the atom.

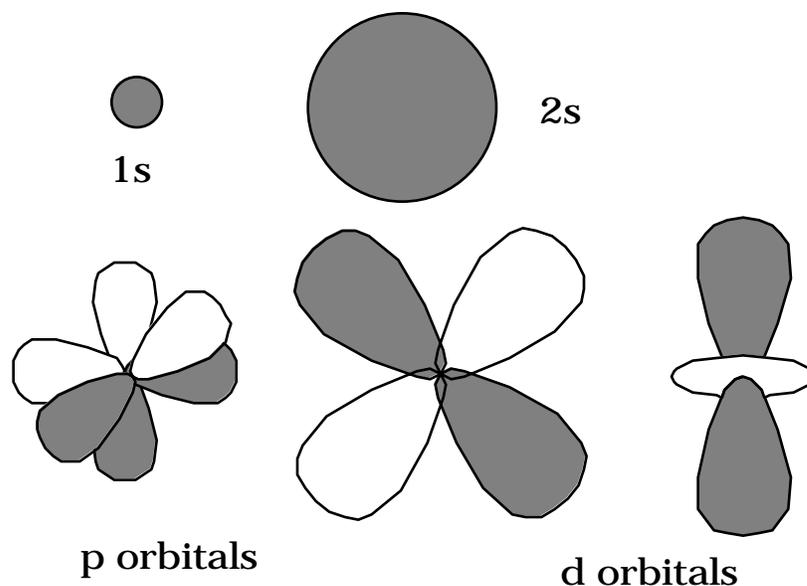
### B. Directions

*Atomic orbital directions also determine what directional bonds an atom will form.*

Each set of p orbitals has three distinct directions or three different angular momentum m-quantum numbers as discussed in Appendix G. Each set of d orbitals has five distinct directions or m-quantum numbers, etc; s orbitals are unidirectional in that they are spherically symmetric, and have only  $m = 0$ . Note that the degeneracy of an orbital  $(2l+1)$ , which is the number of distinct spatial orientations or the number of m-values,

grows with the angular momentum quantum number  $l$  of the orbital without bound.

It is because of the energy degeneracy within a set of orbitals, that these distinct directional orbitals (e.g., x, y, z for p orbitals) may be combined to give new orbitals which no longer possess specific spatial directions but which have specified angular momentum characteristics. The act of combining these degenerate orbitals does not change their energies. For example, the  $2^{-1/2}(p_x + ip_y)$  and  $2^{-1/2}(p_x - ip_y)$  combinations no longer point along the x and y axes, but instead correspond to specific angular momenta ( $+1\hbar$  and  $-1\hbar$ ) about the z axis. The fact that they are angular momentum eigenfunctions can be seen by noting that the x and y orbitals contain dependences of  $\cos(\phi)$  and  $\sin(\phi)$ , respectively. Thus the above combinations contain  $\exp(i\phi)$  and  $\exp(-i\phi)$ , respectively. The sizes, shapes, and directions of a few s, p, and d orbitals are illustrated below (the light and dark areas represent positive and negative values, respectively).



### C. Sizes and Energies

*Orbital energies and sizes go hand-in-hand; small 'tight' orbitals have large electron binding energies (i.e., low energies relative to a detached electron). For orbitals on*

*neighboring atoms to have large (and hence favorable to bond formation) overlap, the two orbitals should be of comparable size and hence of similar electron binding energy.*

The size (e.g., average value or expectation value of the distance from the atomic nucleus to the electron) of an atomic orbital is determined primarily by its principal quantum number  $n$  and by the strength of the potential attracting an electron in this orbital to the atomic center (which has some  $l$ -dependence too). The energy (with negative energies corresponding to bound states in which the electron is attached to the atom with positive binding energy and positive energies corresponding to unbound scattering states) is also determined by  $n$  and by the electrostatic potential produced by the nucleus and by the other electrons. Each atom has an infinite set of orbitals of each  $l$  quantum number ranging from those with low energy and small size to those with higher energy and larger size.

Atomic orbitals are solutions to an orbital-level Schrödinger equation in which an electron moves in a potential energy field provided by the nucleus and all the other electrons. Such one-electron Schrödinger equations are discussed, as they pertain to qualitative and semi-empirical models of electronic structure in Appendix F. The spherical symmetry of the one-electron potential appropriate to atoms and atomic ions is what makes sets of the atomic orbitals degenerate. Such degeneracies arise in molecules too, but the extent of degeneracy is lower because the molecule's nuclear coulomb and electrostatic potential energy has lower symmetry than in the atomic case. As will be seen, it is the symmetry of the potential experienced by an electron moving in the orbital that determines the kind and degree of orbital degeneracy which arises.

Symmetry operators leave the electronic Hamiltonian  $\mathbf{H}$  invariant because the potential and kinetic energies are not changed if one applies such an operator  $\mathbf{R}$  to the coordinates and momenta of all the electrons in the system. Because symmetry operations involve reflections through planes, rotations about axes, or inversions through points, the application of such an operation to a product such as  $\mathbf{H}$  gives the product of the operation applied to each term in the original product. Hence, one can write:

$$\mathbf{R}(\mathbf{H}) = (\mathbf{R}\mathbf{H})(\mathbf{R}).$$

Now using the fact that  $\mathbf{H}$  is invariant to  $\mathbf{R}$ , which means that  $(\mathbf{R}\mathbf{H}) = \mathbf{H}$ , this result reduces to:

$$\mathbf{R}(\mathbf{H}) = \mathbf{H}(\mathbf{R}),$$

which says that  $\mathbf{R}$  commutes with  $\mathbf{H}$ :

$$[\mathbf{R}, \mathbf{H}] = 0.$$

Because symmetry operators commute with the electronic Hamiltonian, the wavefunctions that are eigenstates of  $\mathbf{H}$  can be labeled by the symmetry of the point group of the molecule (i.e., those operators that leave  $\mathbf{H}$  invariant). It is for this reason that one constructs symmetry-adapted atomic basis orbitals to use in forming molecular orbitals.

## II. Molecular Orbitals

*Molecular orbitals (mos) are formed by combining atomic orbitals (aos) of the constituent atoms. This is one of the most important and widely used ideas in quantum chemistry. Much of chemists' understanding of chemical bonding, structure, and reactivity is founded on this point of view.*

When aos are combined to form mos, core, bonding, nonbonding, antibonding, and Rydberg molecular orbitals can result. The mos  $\psi_i$  are usually expressed in terms of the constituent atomic orbitals  $\psi_a$  in the linear-combination-of-atomic-orbital-molecular-orbital (LCAO-MO) manner:

$$\psi_i = \sum_a C_{ia} \psi_a.$$

The orbitals on one atom are orthogonal to one another because they are eigenfunctions of a hermitian operator (the atomic one-electron Hamiltonian) having different eigenvalues. However, those on one atom are not orthogonal to those on another atom because they are eigenfunctions of different operators (the one-electron Hamiltonians of the different atoms). Therefore, in practice, the primitive atomic orbitals must be orthogonalized to preserve maximum identity of each primitive orbital in the resultant orthonormalized orbitals before they can be used in the LCAO-MO process. This is both computationally expedient and conceptually useful. Throughout this book, the atomic orbitals (aos) will be assumed to consist of such orthonormalized primitive orbitals once the nuclei are brought into regions where the "bare" aos interact.

Sets of orbitals that are not orthonormal can be combined to form new orthonormal functions in many ways. One technique that is especially attractive when the original functions are orthonormal in the absence of "interactions" (e.g., at large interatomic

distances in the case of atomic basis orbitals) is the so-called symmetric orthonormalization (SO) method. In this method, one first forms the so-called overlap matrix

$$S_{\mu\nu} = \langle \mu | \nu \rangle$$

for all functions  $\mu$  to be orthonormalized. In the atomic-orbital case, these functions include those on the first atom, those on the second, etc.

Since the orbitals belonging to the individual atoms are themselves orthonormal, the overlap matrix will contain, along its diagonal, blocks of unit matrices, one for each set of individual atomic orbitals. For example, when a carbon and oxygen atom, with their core 1s and valence 2s and 2p orbitals are combined to form CO, the 10x10  $S_{\mu\nu}$  matrix will have two 5x5 unit matrices along its diagonal (representing the overlaps among the carbon and among the oxygen atomic orbitals) and a 5x5 block in its upper right and lower left quadrants. The latter block represents the overlaps  $\langle \mu | \nu \rangle$  among carbon and oxygen atomic orbitals.

After forming the overlap matrix, the new orthonormal functions  $\mu'$  are defined as follows:

$$\mu' = (S^{-1/2})_{\mu}$$

As shown in Appendix A, the matrix  $S^{-1/2}$  is formed by finding the eigenvalues  $\{\epsilon_i\}$  and eigenvectors  $\{V_{i\mu}\}$  of the  $S$  matrix and then constructing:

$$(S^{-1/2})_{\mu} = \sum_i V_{i\mu} V_i (\epsilon_i)^{-1/2}.$$

The new functions  $\{\mu'\}$  have the characteristic that they evolve into the original functions as the "coupling", as represented in the  $S_{\mu\nu}$  matrix's off-diagonal blocks, disappears.

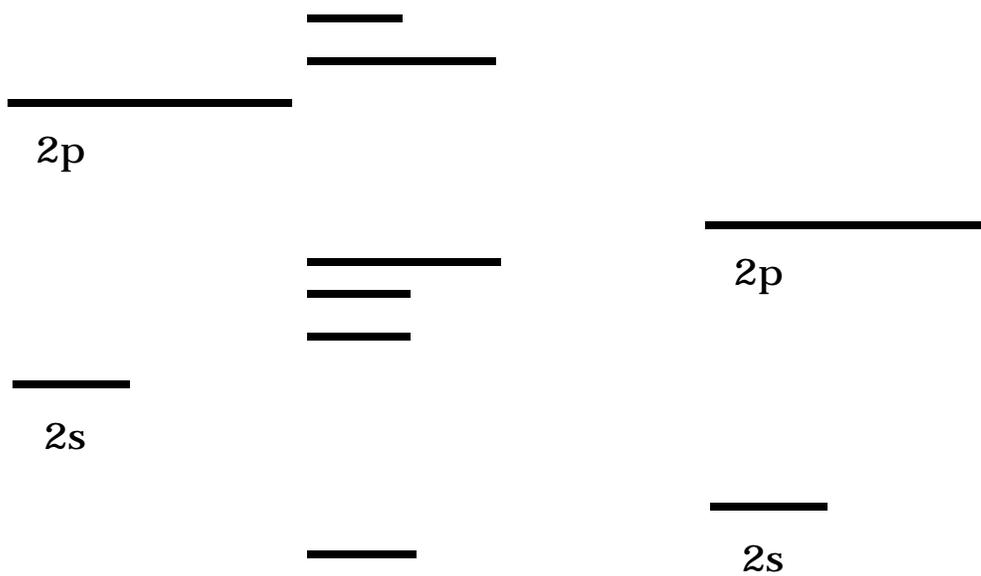
Valence orbitals on neighboring atoms are coupled by changes in the electrostatic potential due to the other atoms (Coulomb attraction to the other nuclei and repulsions from electrons on the other atoms). These coupling potentials vanish when the atoms are far apart and become significant only when the valence orbitals overlap one another. In the most qualitative picture, such interactions are described in terms of off-diagonal Hamiltonian matrix elements ( $h_{ab}$ ; see below and in Appendix F) between pairs of atomic orbitals which interact (the diagonal elements  $h_{aa}$  represent the energies of the various orbitals and are related via Koopmans' theorem (see Section 6, Chapter 18.VII.B) to the ionization energy of the orbital). Such a matrix embodiment of the molecular orbital

problem arises, as developed below and in Appendix F, by using the above LCAO-MO expansion in a variational treatment of the one-electron Schrödinger equation appropriate to the mos  $\{ \psi_i \}$ .

In the simplest two-center, two-valence-orbital case (which could relate, for example, to the  $\text{Li}_2$  molecule's two 2s orbitals), this gives rise to a 2x2 matrix eigenvalue problem  $(h_{11}, h_{12}, h_{22})$  with a low-energy mo  $(E = (h_{11} + h_{22})/2 - 1/2[(h_{11} - h_{22})^2 + 4h_{12}^2]^{1/2})$  and a higher energy mo  $(E = (h_{11} + h_{22})/2 + 1/2[(h_{11} - h_{22})^2 + 4h_{12}^2]^{1/2})$  corresponding to bonding and antibonding orbitals (because their energies lie below and above the lowest and highest interacting atomic orbital energies, respectively). The mos themselves are expressed  $\psi_i = \sum C_{ia} \phi_a$  where the LCAO-MO coefficients  $C_{ia}$  are obtained from the normalized eigenvectors of the  $h_{ab}$  matrix. Note that the bonding-antibonding orbital energy splitting depends on  $h_{ab}^2$  and on the energy difference  $(h_{aa} - h_{bb})$ ; the best bonding (and worst antibonding) occur when two orbitals couple strongly (have large  $h_{ab}$ ) and are similar in energy ( $h_{aa} \approx h_{bb}$ ).

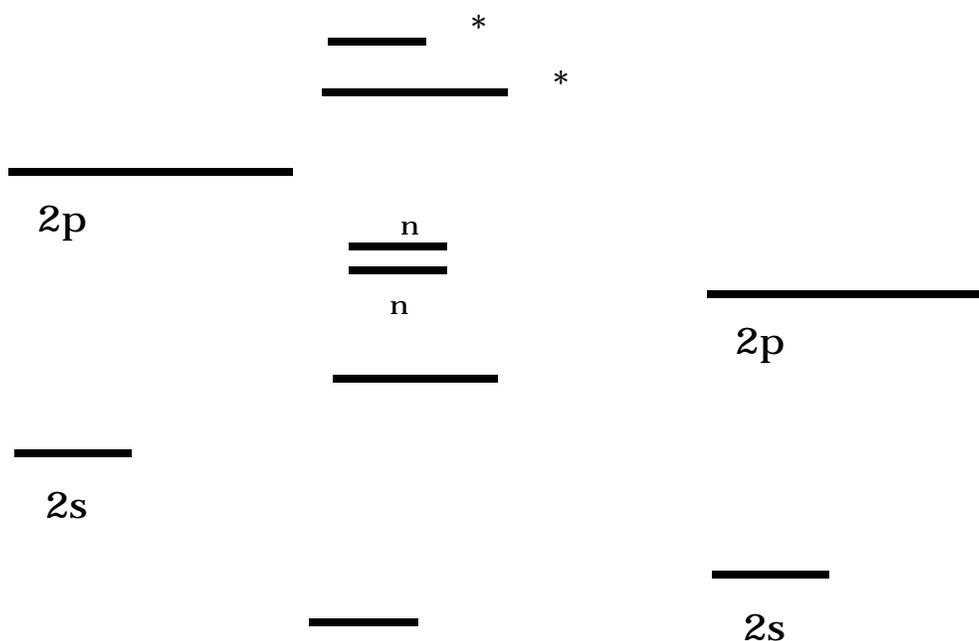


### Homonuclear Bonding With 2s and 2p Orbitals



### Heteronuclear Bonding With 2s and 2p Orbitals

In both the homonuclear and heteronuclear cases depicted above, the energy ordering of the resultant mos depends upon the energy ordering of the constituent aos as well as the strength of the bonding-antibonding interactions among the aos. For example, if the 2s-2p atomic orbital energy splitting is large compared with the interaction matrix elements coupling orbitals on neighboring atoms  $h_{2s,2s}$  and  $h_{2p,2p}$ , then the ordering shown above will result. On the other hand, if the 2s-2p splitting is small, the two 2s and two 2p orbitals can all participate in the formation of the four mos. In this case, it is useful to think of the atomic 2s and 2p orbitals forming sp hybrid orbitals with each atom having one hybrid directed toward the other atom and one hybrid directed away from the other atom. The resultant pattern of four mos will involve one bonding orbital (i.e., an in-phase combination of two sp hybrids), two non-bonding orbitals (those directed away from the other atom) and one antibonding orbital (an out-of-phase combination of two sp hybrids). Their energies will be ordered as shown in the Figure below.



Here  $n$  is used to denote the non-bonding  $\pi$ -type orbitals and  $\sigma$ ,  $\sigma^*$ ,  $\pi$ , and  $\pi^*$  are used to denote bonding and antibonding  $\sigma$ - and  $\pi$ -type orbitals.

Notice that the total number of  $\pi$  orbitals arising from the interaction of the 2s and 2p orbitals is equal to the number of aos that take part in their formation. Notice also that this is true regardless of whether one thinks of the interactions involving bare 2s and 2p

atomic orbitals or hybridized orbitals. The only advantage that the hybrids provide is that they permit one to foresee the fact that two of the four mos must be non-bonding because two of the four hybrids are directed away from all other valence orbitals and hence can not form bonds. In all such qualitative mo analyses, the final results (i.e., how many mos there are of any given symmetry) will not depend on whether one thinks of the interactions involving atomic or hybrid orbitals. However, it is often easier to "guess" the bonding, non-bonding, and antibonding nature of the resultant mos when thought of as formed from hybrids because of the directional properties of the hybrid orbitals.

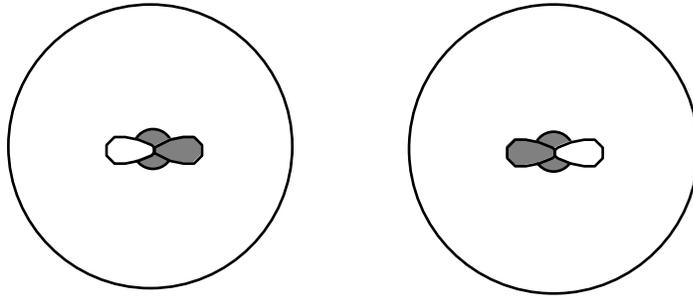
### C. Rydberg Orbitals

*It is essential to keep in mind that all atoms possess 'excited' orbitals that may become involved in bond formation if one or more electrons occupies these orbitals. Whenever aos with principal quantum number one or more unit higher than that of the conventional aos becomes involved in bond formation, Rydberg mos are formed.*

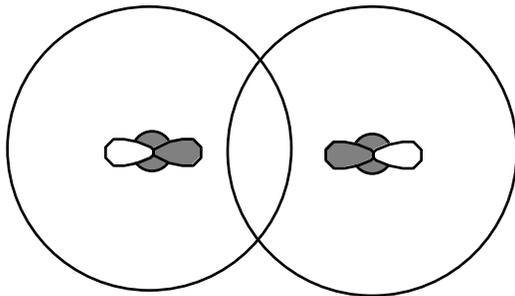
Rydberg orbitals (i.e., very diffuse orbitals having principal quantum numbers higher than the atoms' valence orbitals) can arise in molecules just as they do in atoms. They do not usually give rise to bonding and antibonding orbitals because the valence-orbital interactions bring the atomic centers so close together that the Rydberg orbitals of each atom subsume both atoms. Therefore as the atoms are brought together, the atomic Rydberg orbitals usually pass through the internuclear distance region where they experience (weak) bonding-antibonding interactions all the way to much shorter distances at which they have essentially reached their united-atom limits. As a result, molecular Rydberg orbitals are molecule-centered and display little, if any, bonding or antibonding character. They are usually labeled with principal quantum numbers beginning one higher than the highest n value of the constituent atomic valence orbitals, although they are sometimes labeled by the n quantum number to which they correlate in the united-atom limit.

An example of the interaction of 3s Rydberg orbitals of a molecule whose 2s and 2p orbitals are the valence orbitals and of the evolution of these orbitals into united-atom orbitals is given below.

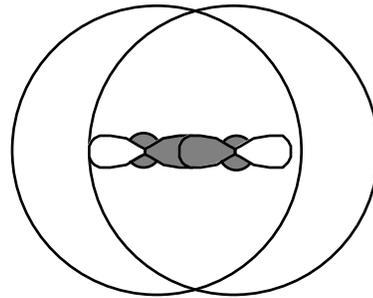




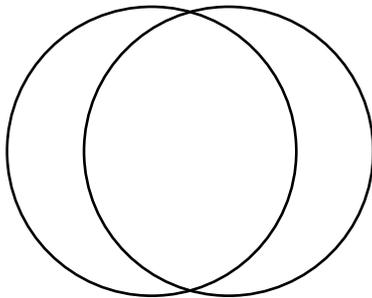
2s and 2p Valence Orbitals and 3s Rydberg Orbitals For Large R Values



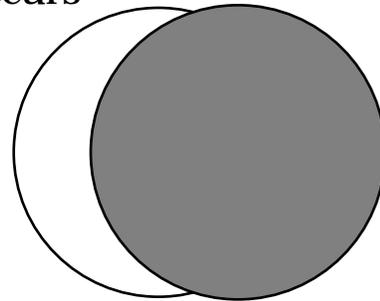
Overlap of the Rydberg Orbitals Begins



Rydberg Overlap is Strong and Bond Formation Occurs



The In-Phase (  $3s + 3s$  ) Combination of Rydberg Orbitals Correlates to an s-type Orbital of the United Atom



The Out-of-Phase Combination of Rydberg Orbitals (  $3s - 3s$  ) Correlates to a p-type United-Atom Orbital

## D. Multicenter Orbitals

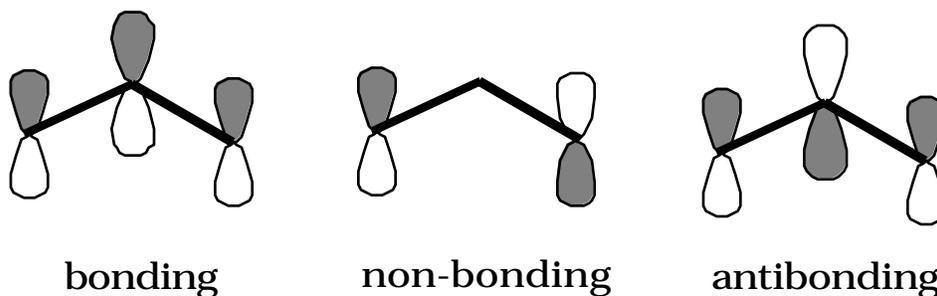
*If aos on one atom overlap aos on more than one neighboring atom, mos that involve amplitudes on three or more atomic centers can be formed. Such mos are termed delocalized or multicenter mos.*

Situations in which more than a pair of orbitals interact can, of course, occur. Three-center bonding occurs in Boron hydrides and in carbonyl bridge bonding in transition metal complexes as well as in delocalized conjugated orbitals common in unsaturated organic hydrocarbons. The three p orbitals on the allyl radical (considered in the absence of the underlying orbitals) can be described qualitatively in terms of three p aos on the three carbon atoms. The couplings  $h_{12}$  and  $h_{23}$  are equal (because the two CC bond lengths are the same) and  $h_{13}$  is approximated as 0 because orbitals 1 and 3 are too far away to interact. The result is a 3x3 secular matrix (see below and in Appendix F):

$$\begin{array}{ccc} h_{11} & h_{12} & 0 \\ h_{21} & h_{22} & h_{23} \\ 0 & h_{32} & h_{33} \end{array}$$

whose eigenvalues give the molecular orbital energies and whose eigenvectors give the LCAO-MO coefficients  $C_{ia}$ .

This 3x3 matrix gives rise to a bonding, a non-bonding and an antibonding orbital (see the Figure below). Since all of the  $h_{aa}$  are equal and  $h_{12} = h_{23}$ , the resultant orbital energies are:  $h_{11} + 2 h_{12}$ ,  $h_{11}$ , and  $h_{11} - 2 h_{12}$ , and the respective LCAO-MO coefficients  $C_{ia}$  are (0.50, 0.707, 0.50), (0.707, 0.00, -0.707), and (0.50, -0.707, 0.50). Notice that the sign (i.e., phase) relations of the bonding orbital are such that overlapping orbitals interact constructively, whereas for the antibonding orbital they interact out of phase. For the nonbonding orbital, there are no interactions because the central C orbital has zero amplitude in this orbital and only  $h_{12}$  and  $h_{23}$  are non-zero.



### Allyl System Orbitals

#### E. Hybrid Orbitals

*It is sometimes convenient to combine aos to form hybrid orbitals that have well defined directional character and to then form mos by combining these hybrid orbitals. This recombination of aos to form hybrids is never necessary and never provides any information that could be achieved in its absence. However, forming hybrids often allows one to focus on those interactions among directed orbitals on neighboring atoms that are most important.*

When atoms combine to form molecules, the molecular orbitals can be thought of as being constructed as linear combinations of the constituent atomic orbitals. This clearly is the only reasonable picture when each atom contributes only one orbital to the particular interactions being considered (e.g., as each Li atom does in  $\text{Li}_2$  and as each C atom does in the orbital aspect of the allyl system). However, when an atom uses more than one of its valence orbitals within particular bonding, non-bonding, or antibonding interactions, it is sometimes useful to combine the constituent atomic orbitals into hybrids and to then use the hybrid orbitals to describe the interactions. As stated above, the directional nature of hybrid orbitals often makes it more straightforward to "guess" the bonding, non-bonding, and antibonding nature of the resultant mos. It should be stressed, however, that exactly the same quantitative results are obtained if one forms mos from primitive aos or from hybrid orbitals; the hybrids span exactly the same space as the original aos and can therefore contain no additional information. This point is illustrated below when the  $\text{H}_2\text{O}$  and  $\text{N}_2$  molecules are treated in both the primitive ao and hybrid orbital bases.

## Chapter 5

### *Molecular Orbitals Possess Specific Topology, Symmetry, and Energy-Level Patterns*

In this chapter the symmetry properties of atomic, hybrid, and molecular orbitals are treated. It is important to keep in mind that both symmetry and characteristics of orbital energetics and bonding "topology", as embodied in the orbital energies themselves and the interactions (i.e.,  $h_{j,k}$  values) among the orbitals, are involved in determining the pattern of molecular orbitals that arise in a particular molecule.

#### I. Orbital Interaction Topology

*The pattern of mo energies can often be 'guessed' by using qualitative information about the energies, overlaps, directions, and shapes of the aos that comprise the mos.*

The orbital interactions determine how many and which mos will have low (bonding), intermediate (non-bonding), and higher (antibonding) energies, with all energies viewed relative to those of the constituent atomic orbitals. The **general patterns** that are observed in most compounds can be summarized as follows:

- i. If the energy splittings among a given atom's aos with the same principal quantum number are small, hybridization can easily occur to produce hybrid orbitals that are directed toward (and perhaps away from) the other atoms in the molecule. In the first-row elements (Li, Be, B, C, N, O, and F), the 2s-2p splitting is small, so hybridization is common. In contrast, for Ca, Ga, Ge, As, and Br it is less common, because the 4s-4p splitting is larger. Orbitals directed toward other atoms can form bonding and antibonding mos; those directed toward no other atoms will form nonbonding mos.
- ii. In attempting to gain a qualitative picture of the electronic structure of any given molecule, it is advantageous to begin by hybridizing the aos of those atoms which contain more than one ao in their valence shell. Only those aos that are not involved in  $\pi$ -orbital interactions should be so hybridized.
- iii. Atomic or hybrid orbitals that are not directed in a  $\pi$ -interaction manner toward other aos or hybrids on neighboring atoms can be involved in  $\sigma$ -interactions or in nonbonding interactions.

iv. Pairs of aos or hybrid orbitals on neighboring atoms directed toward one another interact to produce bonding and antibonding orbitals. The more the bonding orbital lies below the lower-energy ao or hybrid orbital involved in its formation, the higher the antibonding orbital lies above the higher-energy ao or hybrid orbital.

For example, in formaldehyde,  $\text{H}_2\text{CO}$ , one forms  $sp^2$  hybrids on the C atom; on the O atom, either  $sp$  hybrids (with one p orbital "reserved" for use in forming the  $\sigma$  and  $\pi^*$  orbitals and another p orbital to be used as a non-bonding orbital lying in the plane of the molecule) or  $sp^2$  hybrids (with the remaining p orbital reserved for the  $\sigma$  and  $\pi^*$  orbitals) can be used. The H atoms use their 1s orbitals since hybridization is not feasible for them. The C atom clearly uses its  $sp^2$  hybrids to form two CH and one CO  $\sigma$  bonding-antibonding orbital pairs.

The O atom uses one of its  $sp$  or  $sp^2$  hybrids to form the CO  $\sigma$  bond and antibond. When  $sp$  hybrids are used in conceptualizing the bonding, the other  $sp$  hybrid forms a lone pair orbital directed away from the CO bond axis; one of the atomic p orbitals is involved in the CO  $\sigma$  and  $\pi^*$  orbitals, while the other forms an in-plane non-bonding orbital. Alternatively, when  $sp^2$  hybrids are used, the two  $sp^2$  hybrids that do not interact with the C-atom  $sp^2$  orbital form the two non-bonding orbitals. Hence, the final picture of bonding, non-bonding, and antibonding orbitals does not depend on which hybrids one uses as intermediates.

As another example, the 2s and 2p orbitals on the two N atoms of  $\text{N}_2$  can be formed into pairs of  $sp$  hybrids on each N atom plus a pair of p atomic orbitals on each N atom. The  $sp$  hybrids directed toward the other N atom give rise to bonding  $\sigma$  and antibonding  $\sigma^*$  orbitals, and the  $sp$  hybrids directed away from the other N atom yield nonbonding  $\pi$  orbitals. The p orbitals, which consist of 2p orbitals on the N atoms directed perpendicular to the N-N bond axis, produce bonding  $\pi$  and antibonding  $\pi^*$  orbitals.

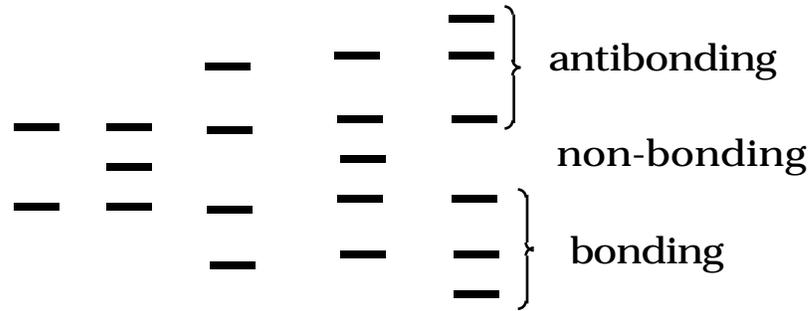
v. In general,  $\sigma$  interactions for a given pair of atoms interacting are stronger than  $\pi$  interactions (which, in turn, are stronger than  $\pi^*$  interactions, etc.) for any given sets (i.e., principal quantum number) of aos that interact. Hence,  $\sigma$  bonding orbitals (originating from a given set of aos) lie below  $\pi$  bonding orbitals, and  $\pi^*$  orbitals lie above  $\pi$  orbitals that arise from the same sets of aos. In the  $\text{N}_2$  example, the  $\sigma$  bonding orbital formed from the two  $sp$  hybrids lies below the  $\pi$  bonding orbital, but the  $\pi^*$  orbital lies below the  $\sigma^*$  orbital. In the  $\text{H}_2\text{CO}$  example, the two CH and the one CO bonding orbitals have low energy; the CO  $\sigma$  bonding orbital has the next lowest energy; the two O-atom non-bonding

orbitals have intermediate energy; the CO \* orbital has somewhat higher energy; and the two CH and one CO antibonding orbitals have the highest energies.

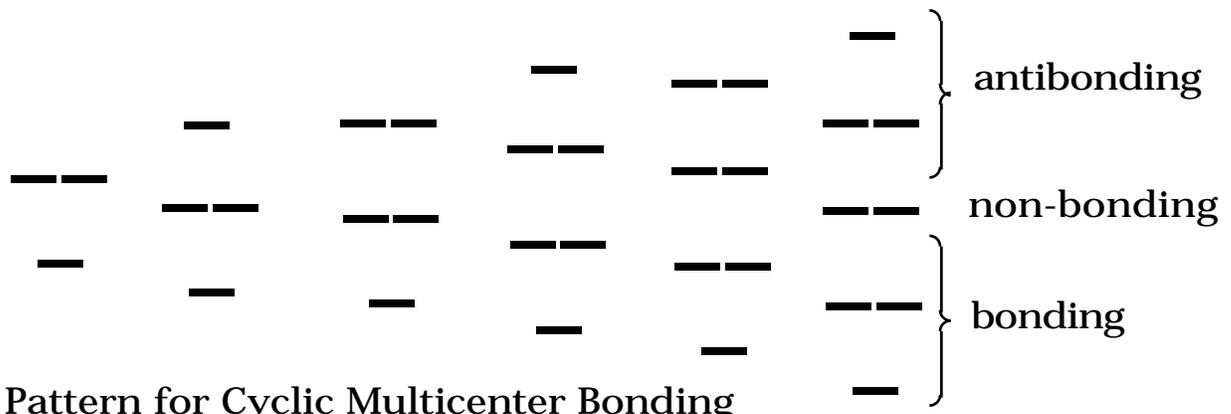
vi. If a given ao or hybrid orbital interacts with or is coupled to orbitals on more than a single neighboring atom, multicenter bonding can occur. For example, in the allyl radical the central carbon atom's p orbital is coupled to the p orbitals on both neighboring atoms; in linear Li<sub>3</sub>, the central Li atom's 2s orbital interacts with the 2s orbitals on both terminal Li atoms; in triangular Cu<sub>3</sub>, the 2s orbitals on each Cu atom couple to each of the other two atoms' 4s orbitals.

vii. Multicenter bonding that involves "linear" chains containing N atoms (e.g., as in conjugated polyenes or in chains of Cu or Na atoms for which the valence orbitals on one atom interact with those of its neighbors on both sides) gives rise to mo energy patterns in which there are N/2 (if N is even) or N/2 -1 non-degenerate bonding orbitals and the same number of antibonding orbitals (if N is odd, there is also a single non-bonding orbital).

viii. Multicenter bonding that involves "cyclic" chains of N atoms (e.g., as in cyclic conjugated polyenes or in rings of Cu or Na atoms for which the valence orbitals on one atom interact with those of its neighbors on both sides and the entire net forms a closed cycle) gives rise to mo energy patterns in which there is a lowest non-degenerate orbital and then a progression of doubly degenerate orbitals. If N is odd, this progression includes (N-1)/2 levels; if N is even, there are (N-2)/2 doubly degenerate levels and a final non-degenerate highest orbital. These patterns and those that appear in linear multicenter bonding are summarized in the Figures shown below.



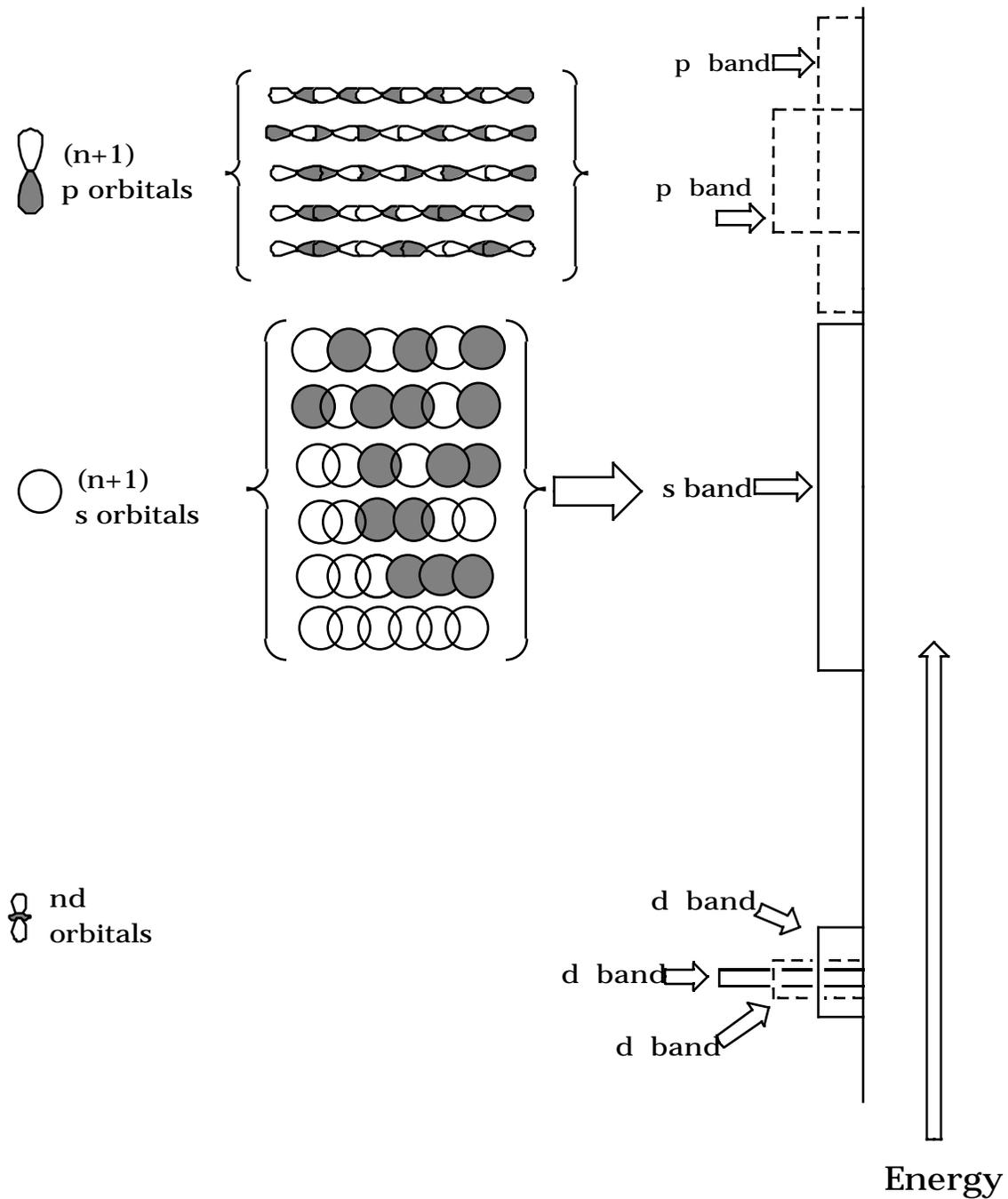
Pattern for Linear Multicenter  
Bonding Situation:  $N=2, 3, \dots, 6$



Pattern for Cyclic Multicenter Bonding  
 $N= 3, 4, 5, \dots, 8$

ix. In extended systems such as solids, atom-based orbitals combine as above to form so-called 'bands' of molecular orbitals. These bands are continuous rather than discrete as in the above cases involving small polyenes. The energy 'spread' within a band depends on the overlap among the atom-based orbitals that form the band; large overlap gives rise to a large band width, while small overlap produces a narrow band. As one moves from the bottom (i.e., the lower energy part) of a band to the top, the number of nodes in the corresponding band orbital increases, as a result of which its bonding nature decreases. In the figure shown below, the bands of a metal such as Ni (with 3d, 4s, and 4p orbitals) is illustrated. The d-orbital band is narrow because the 3d orbitals are small and hence do not overlap appreciably; the 4s and 4p bands are wider because the larger 4s and 4p orbitals overlap to a greater extent. The d-band is split into  $d_{xy}$ ,  $d_{yz}$ , and  $d_{zx}$  components corresponding to the nature of the overlap interactions among the constituent atomic d orbitals. Likewise,

the p-band is split into  $\pi$  and  $\sigma$  components. The widths of the  $\pi$  components of each band are larger than those of the  $\sigma$  components because the corresponding  $\pi$  overlap interactions are stronger. The intensities of the bands at energy  $E$  measure the densities of states at that  $E$ . The total integrated intensity under a given band is a measure of the total number of atomic orbitals that form the band.



## II. Orbital Symmetry

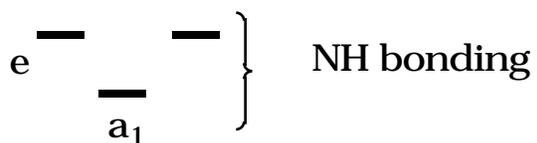
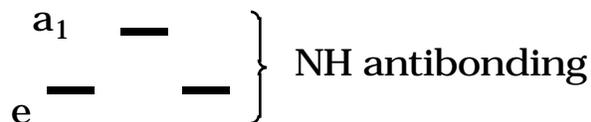
*Symmetry provides additional quantum numbers or labels to use in describing the mos. Each such quantum number further sub-divides the collection of all mos into sets that have vanishing Hamiltonian matrix elements among members belonging to different sets.*

Orbital interaction "topology" as discussed above plays a most- important role in determining the orbital energy level patterns of a molecule. Symmetry also comes into play but in a different manner. Symmetry can be used to characterize the core, bonding, non-bonding, and antibonding molecular orbitals. Much of this chapter is devoted to how this can be carried out in a systematic manner. Once the various mos have been labeled according to symmetry, it may be possible to recognize additional degeneracies that may not have been apparent on the basis of orbital-interaction considerations alone. Thus, topology provides the basic energy ordering pattern and then symmetry enters to identify additional degeneracies.

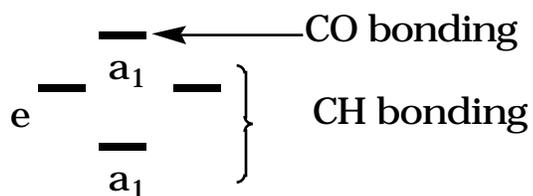
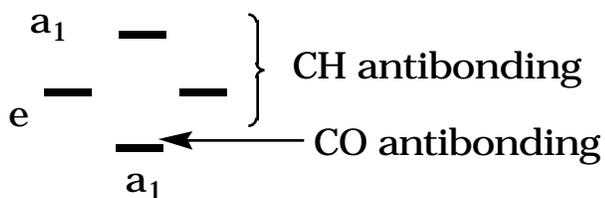
For example, the three NH bonding and three NH antibonding orbitals in  $\text{NH}_3$ , when symmetry adapted within the  $C_{3v}$  point group, cluster into  $a_1$  and  $e$  mos as shown in the Figure below. The N-atom localized non-bonding lone pair orbital and the N-atom  $1s$  core orbital also belong to  $a_1$  symmetry.

In a second example, the three CH bonds, three CH antibonds, CO bond and antibond, and three O-atom non-bonding orbitals of the methoxy radical  $\text{H}_3\text{C-O}$  also cluster into  $a_1$  and  $e$  orbitals as shown below. In these cases, point group symmetry allows one to identify degeneracies that may not have been apparent from the structure of the orbital interactions alone.





### Orbital Character and Symmetry in NH<sub>3</sub>

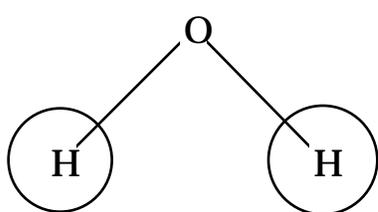


### Orbital Character and Symmetry in H<sub>3</sub>CO Radical

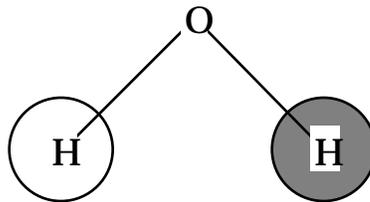


The three resultant molecular orbital energies are, of course, identical to those obtained without symmetry above. The three LCAO-MO coefficients, now expressing the mos in terms of the symmetry adapted orbitals are  $C_{is} = (0.707, 0.707, 0.0)$  for the bonding orbital,  $(0.0, 0.0, 1.00)$  for the nonbonding orbital, and  $(0.707, -0.707, 0.0)$  for the antibonding orbital. These coefficients, when combined with the symmetry adaptation coefficients  $C_{sa}$  given earlier, express the three mos in terms of the three aos as  $\psi_i = \sum_{sa} C_{is} C_{sa} \psi_{sa}$ ; the sum  $\sum_{sa} C_{is} C_{sa}$  gives the LCAO-MO coefficients  $C_{ia}$  which, for example, for the bonding orbital, are  $(0.707^2, 0.707, 0.707^2)$ , in agreement with what was found earlier without using symmetry.

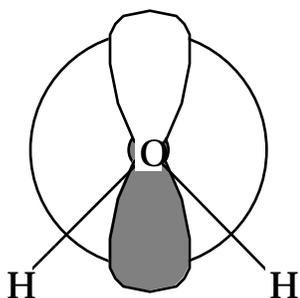
The low energy orbitals of the  $H_2O$  molecule can be used to illustrate the use of symmetry within the primitive ao basis as well as in terms of hybrid orbitals. The 1s orbital on the Oxygen atom is clearly a nonbonding core orbital. The Oxygen 2s orbital and its three 2p orbitals are of valence type, as are the two Hydrogen 1s orbitals. In the absence of symmetry, these six valence orbitals would give rise to a 6x6 secular problem. By combining the two Hydrogen 1s orbitals into  $0.707(1s_L + 1s_R)$  and  $0.707(1s_L - 1s_R)$  symmetry adapted orbitals (labeled  $a_1$  and  $b_2$  within the  $C_{2v}$  point group; see the Figure below), and recognizing that the Oxygen 2s and  $2p_z$  orbitals belong to  $a_1$  symmetry (the z axis is taken as the  $C_2$  rotation axis and the x axis is taken to be perpendicular to the plane in which the three nuclei lie) while the  $2p_x$  orbital is  $b_1$  and the  $2p_y$  orbital is  $b_2$ , allows the 6x6 problem to be decomposed into a 3x3 ( $a_1$ ) secular problem, a 2x2 ( $b_2$ ) secular problem and a 1x1 ( $b_1$ ) problem. These decompositions allow one to conclude that there is one nonbonding  $b_1$  orbital (the Oxygen  $2p_x$  orbital), bonding and antibonding  $b_2$  orbitals (the O-H bond and antibond formed by the Oxygen  $2p_y$  orbital interacting with  $0.707(1s_L - 1s_R)$ ), and, finally, a set of bonding, nonbonding, and antibonding  $a_1$  orbitals (the O-H bond and antibond formed by the Oxygen 2s and  $2p_z$  orbitals interacting with  $0.707(1s_L + 1s_R)$  and the nonbonding orbital formed by the Oxygen 2s and  $2p_z$  orbitals combining to form the "lone pair" orbital directed along the z-axis away from the two Hydrogen atoms).



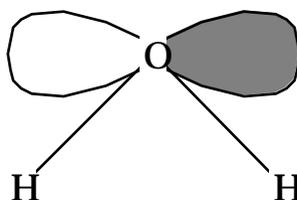
$a_1$  Hydrogen Orbitals



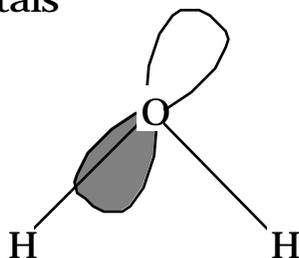
$b_2$  Hydrogen Orbitals



Oxygen  $a_1$  Orbitals



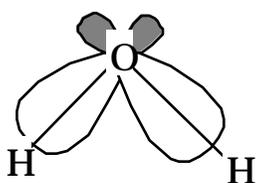
Oxygen  $b_2$  Orbital



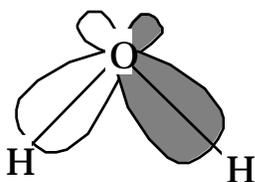
Oxygen  $b_1$  Orbital

Alternatively, to analyze the  $\text{H}_2\text{O}$  molecule in terms of hybrid orbitals, one first combines the Oxygen  $2s$ ,  $2p_z$ ,  $2p_x$  and  $2p_y$  orbitals to form four  $sp^3$  hybrid orbitals. The valence-shell electron-pair repulsion (VSEPR) model of chemical bonding (see R. J. Gillespie and R. S. Nyholm, *Quart. Rev.* 11, 339 (1957) and R. J. Gillespie, *J. Chem. Educ.* 40, 295 (1963)) directs one to involve all of the Oxygen valence orbitals in the hybridization because four  $\sigma$ -bond or nonbonding electron pairs need to be accommodated about the Oxygen center; no  $\pi$  orbital interactions are involved, of course. Having formed the four  $sp^3$  hybrid orbitals, one proceeds as with the primitive aos; one forms symmetry

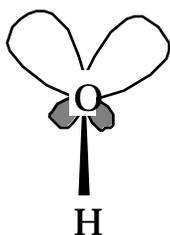
adapted orbitals. In this case, the two Hydrogen  $1s$  orbitals are combined exactly as above to form  $0.707(1s_L + 1s_R)$  and  $0.707(1s_L - 1s_R)$ . The two  $sp^3$  hybrids which lie in the plane of the H and O nuclei (label them L and R) are combined to give symmetry adapted hybrids:  $0.707(L+R)$  and  $0.707(L-R)$ , which are of  $a_1$  and  $b_2$  symmetry, respectively (see the Figure below). The two  $sp^3$  hybrids that lie above and below the plane of the three nuclei (label them T and B) are also symmetry adapted to form  $0.707(T+B)$  and  $0.707(T-B)$ , which are of  $a_1$  and  $b_1$  symmetry, respectively. Once again, one has broken the  $6 \times 6$  secular problem into a  $3 \times 3$   $a_1$  block, a  $2 \times 2$   $b_2$  block and a  $1 \times 1$   $b_1$  block. Although the resulting bonding, nonbonding and antibonding  $a_1$  orbitals, the bonding and antibonding  $b_2$  orbitals and the nonbonding  $b_1$  orbital are now viewed as formed from symmetry adapted Hydrogen orbitals and four Oxygen  $sp^3$  orbitals, they are, of course, exactly the same molecular orbitals as were obtained earlier in terms of the symmetry adapted primitive aos. The formation of hybrid orbitals was an intermediate step which could not alter the final outcome.



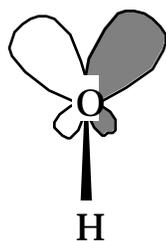
**L + R  $a_1$  Hybrid  
Symmetry Orbital**



**L - R  $b_2$  Hybrid  
Symmetry Orbital**



**T + B Hybrid  
Symmetry Orbital  
Seen From the Side**



**T - B Hybrid  
Symmetry Orbital  
Seen From the Side**

That no degenerate molecular orbitals arose in the above examples is a result of the fact that the  $C_{2v}$  point group to which  $H_2O$  and the allyl system belong (and certainly the

$C_s$  subgroup which was used above in the allyl case) has no degenerate representations. Molecules with higher symmetry such as  $NH_3$ ,  $CH_4$ , and benzene have energetically degenerate orbitals because their molecular point groups have degenerate representations.

## B. Linear Molecules

*Linear molecules belong to the axial rotation group. Their symmetry is intermediate in complexity between nonlinear molecules and atoms.*

For linear molecules, the symmetry of the electrostatic potential provided by the nuclei and the other electrons is described by either the  $C_{\infty v}$  or  $D_{\infty h}$  group. The essential difference between these symmetry groups and the finite point groups which characterize the non-linear molecules lies in the fact that the electrostatic potential which an electron feels is invariant to rotations of any amount about the molecular axis (i.e.,  $V(\theta + \phi) = V(\theta)$ , for any angle increment  $\phi$ ). This means that the operator  $C_{\infty}$  which generates a rotation of the electron's azimuthal angle  $\phi$  by an amount  $\phi$  about the molecular axis commutes with the Hamiltonian  $[H, C_{\infty}] = 0$ .  $C_{\infty}$  can be written in terms of the quantum mechanical operator  $L_z = -i\hbar \frac{\partial}{\partial \phi}$  describing the orbital angular momentum of the electron about the molecular (z) axis:

$$C_{\infty} = \exp(i L_z / \hbar).$$

Because  $C_{\infty}$  commutes with the Hamiltonian and  $C_{\infty}$  can be written in terms of  $L_z$ ,  $L_z$  must commute with the Hamiltonian. As a result, the molecular orbitals of a linear molecule must be eigenfunctions of the z-component of angular momentum  $L_z$ :

$$-i\hbar \frac{\partial}{\partial \phi} \psi = m\hbar \psi.$$

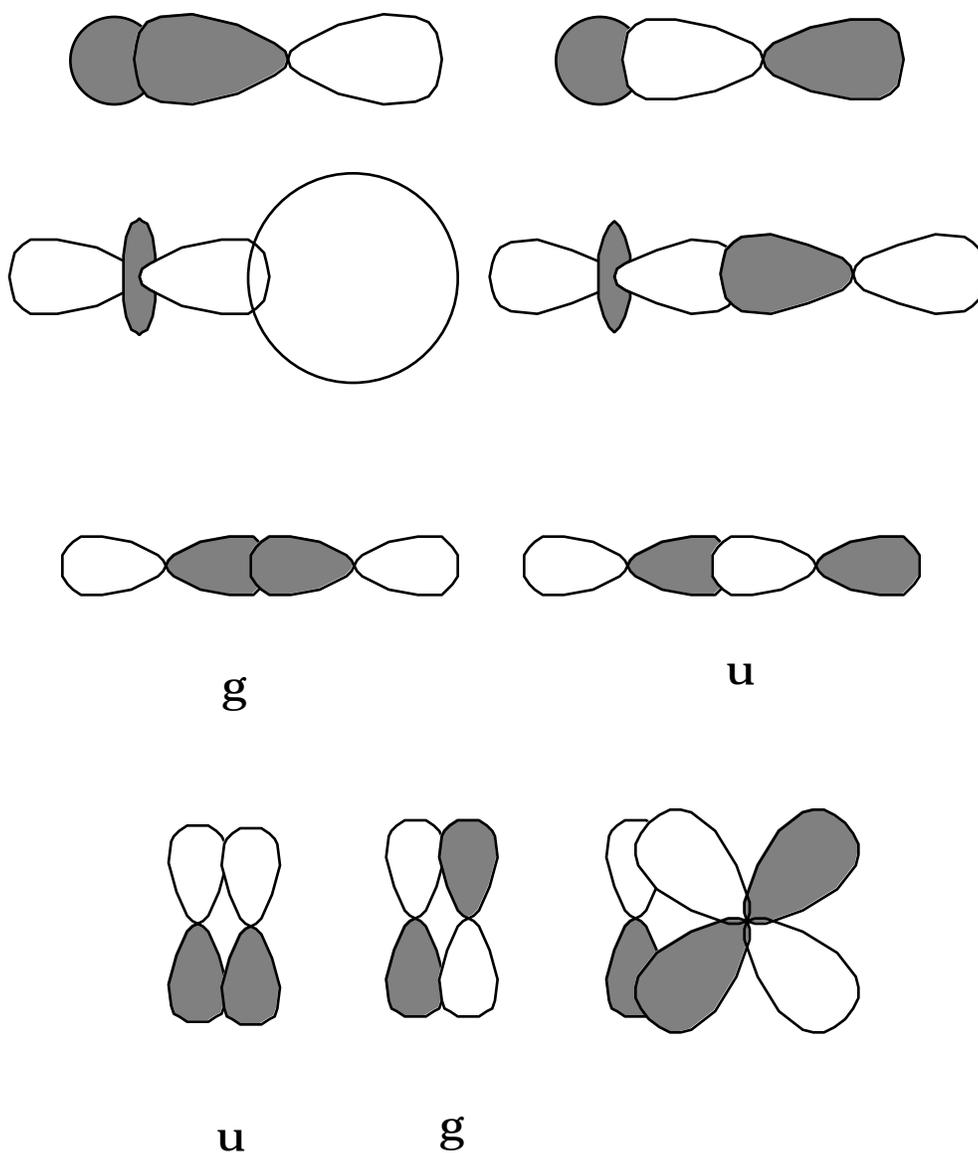
The electrostatic potential is not invariant under rotations of the electron about the x or y axes (those perpendicular to the molecular axis), so  $L_x$  and  $L_y$  do not commute with the Hamiltonian. Therefore, only  $L_z$  provides a "good quantum number" in the sense that the operator  $L_z$  commutes with the Hamiltonian.

In summary, the molecular orbitals of a linear molecule can be labeled by their m quantum number, which plays the same role as the point group labels did for non-linear polyatomic molecules, and which gives the eigenvalue of the angular momentum of the orbital about the molecule's symmetry axis. Because the kinetic energy part of the

Hamiltonian contains  $(\hbar^2/2m_e r^2) \propto 1/r^2$ , whereas the potential energy part is independent of  $\phi$ , the energies of the molecular orbitals depend on the square of the  $m$  quantum number. Thus, pairs of orbitals with  $m = \pm 1$  are energetically degenerate; pairs with  $m = \pm 2$  are degenerate, and so on. The absolute value of  $m$ , which is what the energy depends on, is called the  $|m|$  quantum number. Molecular orbitals with  $|m| = 0$  are called  $\sigma$  orbitals; those with  $|m| = 1$  are  $\pi$  orbitals; and those with  $|m| = 2$  are  $\delta$  orbitals.

Just as in the non-linear polyatomic-molecule case, the atomic orbitals which constitute a given molecular orbital must have the same symmetry as that of the molecular orbital. This means that  $\sigma$ ,  $\pi$ , and  $\delta$  molecular orbitals are formed, via LCAO-MO, from  $m=0$ ,  $m = \pm 1$ , and  $m = \pm 2$  atomic orbitals, respectively. In the diatomic  $N_2$  molecule, for example, the core orbitals are of  $\sigma$  symmetry as are the molecular orbitals formed from the  $2s$  and  $2p_z$  atomic orbitals (or their hybrids) on each Nitrogen atom. The molecular orbitals formed from the atomic  $2p_{-1} = (2p_x - i 2p_y)$  and the  $2p_{+1} = (2p_x + i 2p_y)$  orbitals are of  $\pi$  symmetry and have  $m = -1$  and  $+1$ .

For homonuclear diatomic molecules and other linear molecules which have a center of symmetry, the inversion operation (in which an electron's coordinates are inverted through the center of symmetry of the molecule) is also a symmetry operation. Each resultant molecular orbital can then also be labeled by a quantum number denoting its parity with respect to inversion. The symbols *g* (for gerade or even) and *u* (for ungerade or odd) are used for this label. Again for  $N_2$ , the core orbitals are of *g* and *u* symmetry, and the bonding and antibonding orbitals formed from the 2s and 2p orbitals on the two Nitrogen atoms are of *g* and *u* symmetry.



The bonding molecular orbital pair (with  $m = +1$  and  $-1$ ) is of  $\sigma_u$  symmetry whereas the corresponding antibonding orbital is of  $\sigma_g$  symmetry. Examples of such molecular orbital symmetries are shown above.

The use of hybrid orbitals can be illustrated in the linear-molecule case by considering the  $N_2$  molecule. Because two bonding and antibonding molecular orbital pairs are involved in  $N_2$  (one with  $m = +1$ , one with  $m = -1$ ), VSEPR theory guides one to form  $sp$  hybrid orbitals from each of the Nitrogen atom's  $2s$  and  $2p_z$  (which is also the  $2p$  orbital with  $m = 0$ ) orbitals. Ignoring the core orbitals, which are of  $\sigma_g$  and  $\sigma_u$  symmetry as noted above, one then symmetry adapts the four  $sp$  hybrids (two from each atom) to build one  $\sigma_g$  orbital involving a bonding interaction between two  $sp$  hybrids pointed toward one another, an antibonding  $\sigma_u$  orbital involving the same pair of  $sp$  orbitals but coupled with opposite signs, a nonbonding  $\sigma_g$  orbital composed of two  $sp$  hybrids pointed away from the interatomic region combined with like sign, and a nonbonding  $\sigma_u$  orbital made of the latter two  $sp$  hybrids combined with opposite signs. The two  $2p_m$  orbitals ( $m = +1$  and  $-1$ ) on each Nitrogen atom are then symmetry adapted to produce a pair of bonding  $\sigma_u$  orbitals (with  $m = +1$  and  $-1$ ) and a pair of antibonding  $\sigma_g$  orbitals (with  $m = +1$  and  $-1$ ). This hybridization and symmetry adaptation thereby reduces the  $8 \times 8$  secular problem (which would be  $10 \times 10$  if the core orbitals were included) into a  $2 \times 2$   $\sigma_g$  problem (one bonding and one nonbonding), a  $2 \times 2$   $\sigma_u$  problem (one bonding and one nonbonding), an identical pair of  $1 \times 1$   $\sigma_u$  problems (bonding), and an identical pair of  $1 \times 1$   $\sigma_g$  problems (antibonding).

Another example of the equivalence among various hybrid and atomic orbital points of view is provided by the  $CO$  molecule. Using, for example,  $sp$  hybrid orbitals on  $C$  and  $O$ , one obtains a picture in which there are: two core orbitals corresponding to the  $O$ -atom  $1s$  and  $C$ -atom  $1s$  orbitals; one  $CO$  bonding, two non-bonding, and one  $CO$  antibonding orbitals arising from the four  $sp$  hybrids; a pair of bonding and a pair of antibonding orbitals formed from the two  $p$  orbitals on  $O$  and the two  $p$  orbitals on  $C$ . Alternatively, using  $sp^2$  hybrids on both  $C$  and  $O$ , one obtains: the two core orbitals as above; a  $CO$  bonding and antibonding orbital pair formed from the  $sp^2$  hybrids that are directed along the  $CO$  bond; and a single bonding and antibonding  $\pi^*$  orbital set. The remaining two  $sp^2$  orbitals on  $C$  and the two on  $O$  can then be symmetry adapted by forming  $\pm$  combinations within each pair to yield: an  $a_1$  non-bonding orbital (from the  $+$  combination) on each of  $C$  and  $O$  directed away from the  $CO$  bond axis; and a  $p$  orbital on each of  $C$  and  $O$  that can subsequently overlap to form the second bonding and  $\pi^*$  antibonding orbital pair.

It should be clear from the above examples, that no matter what particular hybrid

orbitals one chooses to utilize in conceptualizing a molecule's orbital interactions, symmetry ultimately returns to force one to form proper symmetry adapted combinations which, in turn, renders the various points of view equivalent. In the above examples and in several earlier examples, symmetry adaptation of, for example,  $sp^2$  orbital pairs (e.g.,  $sp_L^2 \pm sp_R^2$ ) generated orbitals of pure spatial symmetry. In fact, symmetry combining hybrid orbitals in this manner amounts to forming other hybrid orbitals. For example, the above  $\pm$  combinations of  $sp^2$  hybrids directed to the left (L) and right (R) of some bond axis generate a new  $sp$  hybrid directed along the bond axis but opposite to the  $sp^2$  hybrid used to form the bond and a non-hybridized p orbital directed along the L-to-R direction. In the CO example, these combinations of  $sp^2$  hybrids on O and C produce  $sp$  hybrids on O and C and p orbitals on O and C.

### C. Atoms

*Atoms belong to the full rotation symmetry group; this makes their symmetry analysis the most complex to treat.*

In moving from linear molecules to atoms, additional symmetry elements arise. In particular, the potential field experienced by an electron in an orbital becomes invariant to rotations of arbitrary amounts about the x, y, and z axes; in the linear-molecule case, it is invariant only to rotations of the electron's position about the molecule's symmetry axis (the z axis). These invariances are, of course, caused by the spherical symmetry of the potential of any atom. This additional symmetry of the potential causes the Hamiltonian to commute with all three components of the electron's angular momentum:  $[L_x, H] = 0$ ,  $[L_y, H] = 0$ , and  $[L_z, H] = 0$ . It is straightforward to show that H also commutes with the operator  $L^2 = L_x^2 + L_y^2 + L_z^2$ , defined as the sum of the squares of the three individual components of the angular momentum. Because  $L_x$ ,  $L_y$ , and  $L_z$  do not commute with one another, orbitals which are eigenfunctions of H cannot be simultaneous eigenfunctions of all three angular momentum operators. Because  $L_x$ ,  $L_y$ , and  $L_z$  do commute with  $L^2$ , orbitals can be found which are eigenfunctions of H, of  $L^2$  and of any one component of  $\mathbf{L}$ ; it is convention to select  $L_z$  as the operator which, along with H and  $L^2$ , form a mutually commutative operator set of which the orbitals are simultaneous eigenfunctions.

So, for any atom, the orbitals can be labeled by both l and m quantum numbers, which play the role that point group labels did for non-linear molecules and did for linear molecules. Because (i) the kinetic energy operator in the electronic Hamiltonian explicitly contains  $L^2/2m_e r^2$ , (ii) the Hamiltonian does not contain additional  $L_z$ ,  $L_x$ , or  $L_y$  factors,

and (iii) the potential energy part of the Hamiltonian is spherically symmetric (and commutes with  $L^2$  and  $L_z$ ), the energies of atomic orbitals depend upon the  $l$  quantum number and are independent of the  $m$  quantum number. This is the source of the  $2l+1$ -fold degeneracy of atomic orbitals.

The angular part of the atomic orbitals is described in terms of the spherical harmonics  $Y_{l,m}$ ; that is, each atomic orbital can be expressed as

$$\psi_{n,l,m} = Y_{l,m}(\theta, \phi) R_{n,l}(r).$$

The explicit solutions for the  $Y_{l,m}$  and for the radial wavefunctions  $R_{n,l}$  are given in Appendix B. The variables  $r, \theta, \phi$  give the position of the electron in the orbital in spherical coordinates. These angular functions are, as discussed earlier, related to the cartesian (i.e., spatially oriented) orbitals by simple transformations; for example, the orbitals with  $l=2$  and  $m=2,1,0,-1,-2$  can be expressed in terms of the  $d_{xy}, d_{xz}, d_{yz}, d_{xx-yy}$ , and  $d_{zz}$  orbitals. Either set of orbitals is acceptable in the sense that each orbital is an eigenfunction of  $H$ ; transformations within a degenerate set of orbitals do not destroy the Hamiltonian-eigenfunction feature. The orbital set labeled with  $l$  and  $m$  quantum numbers is most useful when one is dealing with isolated atoms (which have spherical symmetry), because  $m$  is then a valid symmetry label, or with an atom in a local environment which is axially symmetric (e.g., in a linear molecule) where the  $m$  quantum number remains a useful symmetry label. The cartesian orbitals are preferred for describing an atom in a local environment which displays lower than axial symmetry (e.g., an atom interacting with a diatomic molecule in  $C_{2v}$  symmetry).

The radial part of the orbital  $R_{n,l}(r)$  as well as the orbital energy  $\epsilon_{n,l}$  depend on  $l$  because the Hamiltonian itself contains  $l(l+1)\hbar^2/2m_e r^2$ ; they are independent of  $m$  because the Hamiltonian has no  $m$ -dependence. For bound orbitals,  $R_{n,l}(r)$  decays exponentially for large  $r$  (as  $\exp(-2r/a_0 n)$ ), and for unbound (scattering) orbitals, it is oscillatory at large  $r$  with an oscillation period related to the deBroglie wavelength of the electron. In  $R_{n,l}(r)$  there are  $(n-l-1)$  radial nodes lying between  $r=0$  and  $r=\infty$ . These nodes provide differential stabilization of low- $l$  orbitals over high- $l$  orbitals of the same principal quantum number  $n$ . That is, penetration of outer shells is greater for low- $l$  orbitals because they have more radial nodes; as a result, they have larger amplitude near the atomic nucleus and thus experience enhanced attraction to the positive nuclear charge. The average size (e.g., average value of  $r$ ;  $\langle r \rangle = \int R_{n,l}^2 r^3 dr$ ) of an orbital depends strongly on  $n$ , weakly on  $l$  and is independent of  $m$ ; it also depends strongly on the nuclear charge and on the potential produced by the other electrons. This potential is often characterized qualitatively in terms

of an effective nuclear charge  $Z_{\text{eff}}$  which is the true nuclear charge of the atom  $Z$  minus a screening component  $Z_{\text{sc}}$  which describes the repulsive effect of the electron density lying radially inside the electron under study. Because, for a given  $n$ , low- $l$  orbitals penetrate closer to the nucleus than do high- $l$  orbitals, they have higher  $Z_{\text{eff}}$  values (i.e., smaller  $Z_{\text{sc}}$  values) and correspondingly smaller average sizes and larger binding energies.



## Section 2 Simple Molecular Orbital Theory

In this section, the conceptual framework of molecular orbital theory is developed. Applications are presented and problems are given and solved within qualitative and semi-empirical models of electronic structure. Ab Initio approaches to these same matters, whose solutions require the use of digital computers, are treated later in Section 6. Semi-empirical methods, most of which also require access to a computer, are treated in this section and in Appendix F.

Unlike most texts on molecular orbital theory and quantum mechanics, this text treats polyatomic molecules before linear molecules before atoms. The finite point-group symmetry (Appendix E provides an introduction to the use of point group symmetry) that characterizes the orbitals and electronic states of non-linear polyatomics is more straightforward to deal with because fewer degeneracies arise. In turn, linear molecules, which belong to an axial rotation group, possess fewer degeneracies (e.g., orbitals or states are no more degenerate than  $s$ ,  $p$ , or  $d$  orbitals or states; all are doubly degenerate) than atomic orbitals and states (e.g.,  $p$  orbitals or states are 3-fold degenerate,  $d$ 's are 5-fold, etc.). Increased orbital degeneracy, in turn, gives rise to more states that can arise from a given orbital occupancy (e.g., the  $2p^2$  configuration of the C atom yields fifteen states, the  $2p^2$  configuration of the NH molecule yields six, and the  $\pi^*$  configuration of ethylene gives four states). For these reasons, it is more straightforward to treat low-symmetry cases (i.e., non-linear polyatomic molecules) first and atoms last.

It is recommended that the reader become familiar with the point-group symmetry tools developed in Appendix E before proceeding with this section. In particular, it is important to know how to label atomic orbitals as well as the various hybrids that can be formed from them according to the irreducible representations of the molecule's point group and how to construct symmetry adapted combinations of atomic, hybrid, and molecular orbitals using projection operator methods. If additional material on group theory is needed, Cotton's book on this subject is very good and provides many excellent chemical applications.

### Chapter 4

*Valence Atomic Orbitals on Neighboring Atoms Combine to Form Bonding, Non-Bonding and Antibonding Molecular Orbitals*

#### I. Atomic Orbitals

In Section 1 the Schrödinger equation for the motion of a single electron moving about a nucleus of charge  $Z$  was explicitly solved. The energies of these orbitals relative to an electron infinitely far from the nucleus with zero kinetic energy were found to depend strongly on  $Z$  and on the principal quantum number  $n$ , as were the radial "sizes" of these hydrogenic orbitals. Closed analytical expressions for the  $r$ ,  $\theta$ , and  $\phi$  dependence of these orbitals are given in Appendix B. The reader is advised to also review this material before undertaking study of this section.

### A. Shapes

*Shapes of atomic orbitals play central roles in governing the types of directional bonds an atom can form.*

All atoms have sets of bound and continuum s,p,d,f,g, etc. orbitals. Some of these orbitals may be unoccupied in the atom's low energy states, but they are still present and able to accept electron density if some physical process (e.g., photon absorption, electron attachment, or Lewis-base donation) causes such to occur. For example, the Hydrogen atom has 1s, 2s, 2p, 3s, 3p, 3d, etc. orbitals. Its negative ion  $H^-$  has states that involve 1s2s, 2p<sup>2</sup>, 3s<sup>2</sup>, 3p<sup>2</sup>, etc. orbital occupancy. Moreover, when an H atom is placed in an external electronic field, its charge density polarizes in the direction of the field. This polarization can be described in terms of the orbitals of the isolated atom being combined to yield distorted orbitals (e.g., the 1s and 2p orbitals can "mix" or combine to yield sp hybrid orbitals, one directed toward increasing field and the other directed in the opposite direction). Thus in many situations it is important to keep in mind that each atom has a full set of orbitals available to it even if some of these orbitals are not occupied in the lowest-energy state of the atom.

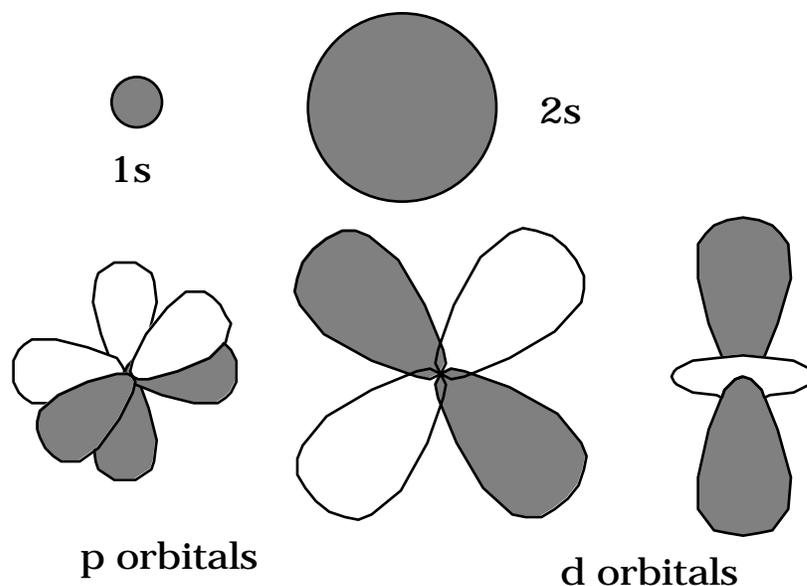
### B. Directions

*Atomic orbital directions also determine what directional bonds an atom will form.*

Each set of p orbitals has three distinct directions or three different angular momentum m-quantum numbers as discussed in Appendix G. Each set of d orbitals has five distinct directions or m-quantum numbers, etc; s orbitals are unidirectional in that they are spherically symmetric, and have only  $m = 0$ . Note that the degeneracy of an orbital  $(2l+1)$ , which is the number of distinct spatial orientations or the number of m-values,

grows with the angular momentum quantum number  $l$  of the orbital without bound.

It is because of the energy degeneracy within a set of orbitals, that these distinct directional orbitals (e.g., x, y, z for p orbitals) may be combined to give new orbitals which no longer possess specific spatial directions but which have specified angular momentum characteristics. The act of combining these degenerate orbitals does not change their energies. For example, the  $2^{-1/2}(p_x + ip_y)$  and  $2^{-1/2}(p_x - ip_y)$  combinations no longer point along the x and y axes, but instead correspond to specific angular momenta ( $+1\hbar$  and  $-1\hbar$ ) about the z axis. The fact that they are angular momentum eigenfunctions can be seen by noting that the x and y orbitals contain dependences of  $\cos(\theta)$  and  $\sin(\theta)$ , respectively. Thus the above combinations contain  $\exp(i\phi)$  and  $\exp(-i\phi)$ , respectively. The sizes, shapes, and directions of a few s, p, and d orbitals are illustrated below (the light and dark areas represent positive and negative values, respectively).



### C. Sizes and Energies

*Orbital energies and sizes go hand-in-hand; small 'tight' orbitals have large electron binding energies (i.e., low energies relative to a detached electron). For orbitals on*

*neighboring atoms to have large (and hence favorable to bond formation) overlap, the two orbitals should be of comparable size and hence of similar electron binding energy.*

The size (e.g., average value or expectation value of the distance from the atomic nucleus to the electron) of an atomic orbital is determined primarily by its principal quantum number  $n$  and by the strength of the potential attracting an electron in this orbital to the atomic center (which has some  $l$ -dependence too). The energy (with negative energies corresponding to bound states in which the electron is attached to the atom with positive binding energy and positive energies corresponding to unbound scattering states) is also determined by  $n$  and by the electrostatic potential produced by the nucleus and by the other electrons. Each atom has an infinite set of orbitals of each  $l$  quantum number ranging from those with low energy and small size to those with higher energy and larger size.

Atomic orbitals are solutions to an orbital-level Schrödinger equation in which an electron moves in a potential energy field provided by the nucleus and all the other electrons. Such one-electron Schrödinger equations are discussed, as they pertain to qualitative and semi-empirical models of electronic structure in Appendix F. The spherical symmetry of the one-electron potential appropriate to atoms and atomic ions is what makes sets of the atomic orbitals degenerate. Such degeneracies arise in molecules too, but the extent of degeneracy is lower because the molecule's nuclear coulomb and electrostatic potential energy has lower symmetry than in the atomic case. As will be seen, it is the symmetry of the potential experienced by an electron moving in the orbital that determines the kind and degree of orbital degeneracy which arises.

Symmetry operators leave the electronic Hamiltonian  $\mathbf{H}$  invariant because the potential and kinetic energies are not changed if one applies such an operator  $\mathbf{R}$  to the coordinates and momenta of all the electrons in the system. Because symmetry operations involve reflections through planes, rotations about axes, or inversions through points, the application of such an operation to a product such as  $\mathbf{H}$  gives the product of the operation applied to each term in the original product. Hence, one can write:

$$\mathbf{R}(\mathbf{H}) = (\mathbf{R}\mathbf{H})(\mathbf{R}).$$

Now using the fact that  $\mathbf{H}$  is invariant to  $\mathbf{R}$ , which means that  $(\mathbf{R}\mathbf{H}) = \mathbf{H}$ , this result reduces to:

$$\mathbf{R}(\mathbf{H}) = \mathbf{H}(\mathbf{R}),$$

which says that  $\mathbf{R}$  commutes with  $\mathbf{H}$ :

$$[\mathbf{R}, \mathbf{H}] = 0.$$

Because symmetry operators commute with the electronic Hamiltonian, the wavefunctions that are eigenstates of  $\mathbf{H}$  can be labeled by the symmetry of the point group of the molecule (i.e., those operators that leave  $\mathbf{H}$  invariant). It is for this reason that one constructs symmetry-adapted atomic basis orbitals to use in forming molecular orbitals.

## II. Molecular Orbitals

*Molecular orbitals (mos) are formed by combining atomic orbitals (aos) of the constituent atoms. This is one of the most important and widely used ideas in quantum chemistry. Much of chemists' understanding of chemical bonding, structure, and reactivity is founded on this point of view.*

When aos are combined to form mos, core, bonding, nonbonding, antibonding, and Rydberg molecular orbitals can result. The mos  $\psi_i$  are usually expressed in terms of the constituent atomic orbitals  $\psi_a$  in the linear-combination-of-atomic-orbital-molecular-orbital (LCAO-MO) manner:

$$\psi_i = \sum_a C_{ia} \psi_a.$$

The orbitals on one atom are orthogonal to one another because they are eigenfunctions of a hermitian operator (the atomic one-electron Hamiltonian) having different eigenvalues. However, those on one atom are not orthogonal to those on another atom because they are eigenfunctions of different operators (the one-electron Hamiltonia of the different atoms). Therefore, in practice, the primitive atomic orbitals must be orthogonalized to preserve maximum identity of each primitive orbital in the resultant orthonormalized orbitals before they can be used in the LCAO-MO process. This is both computationally expedient and conceptually useful. Throughout this book, the atomic orbitals (aos) will be assumed to consist of such orthonormalized primitive orbitals once the nuclei are brought into regions where the "bare" aos interact.

Sets of orbitals that are not orthonormal can be combined to form new orthonormal functions in many ways. One technique that is especially attractive when the original functions are orthonormal in the absence of "interactions" (e.g., at large interatomic

distances in the case of atomic basis orbitals) is the so-called symmetric orthonormalization (SO) method. In this method, one first forms the so-called overlap matrix

$$S_{\mu\nu} = \langle \mu | \nu \rangle$$

for all functions  $\mu$  to be orthonormalized. In the atomic-orbital case, these functions include those on the first atom, those on the second, etc.

Since the orbitals belonging to the individual atoms are themselves orthonormal, the overlap matrix will contain, along its diagonal, blocks of unit matrices, one for each set of individual atomic orbitals. For example, when a carbon and oxygen atom, with their core 1s and valence 2s and 2p orbitals are combined to form CO, the 10x10  $S_{\mu\nu}$  matrix will have two 5x5 unit matrices along its diagonal (representing the overlaps among the carbon and among the oxygen atomic orbitals) and a 5x5 block in its upper right and lower left quadrants. The latter block represents the overlaps  $\langle \mu | \nu \rangle$  among carbon and oxygen atomic orbitals.

After forming the overlap matrix, the new orthonormal functions  $\mu'$  are defined as follows:

$$\mu' = (S^{-1/2})_{\mu}$$

As shown in Appendix A, the matrix  $S^{-1/2}$  is formed by finding the eigenvalues  $\{\epsilon_i\}$  and eigenvectors  $\{V_{i\mu}\}$  of the  $S$  matrix and then constructing:

$$(S^{-1/2})_{\mu} = \sum_i V_{i\mu} V_i (\epsilon_i)^{-1/2}.$$

The new functions  $\{\mu'\}$  have the characteristic that they evolve into the original functions as the "coupling", as represented in the  $S_{\mu\nu}$  matrix's off-diagonal blocks, disappears.

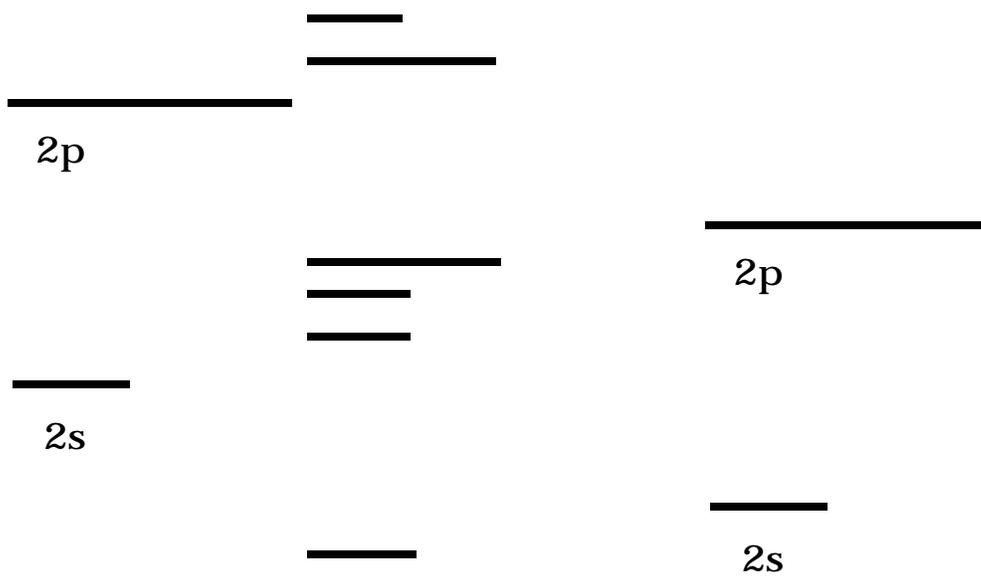
Valence orbitals on neighboring atoms are coupled by changes in the electrostatic potential due to the other atoms (coulomb attraction to the other nuclei and repulsions from electrons on the other atoms). These coupling potentials vanish when the atoms are far apart and become significant only when the valence orbitals overlap one another. In the most qualitative picture, such interactions are described in terms of off-diagonal Hamiltonian matrix elements ( $h_{ab}$ ; see below and in Appendix F) between pairs of atomic orbitals which interact (the diagonal elements  $h_{aa}$  represent the energies of the various orbitals and are related via Koopmans' theorem (see Section 6, Chapter 18.VII.B) to the ionization energy of the orbital). Such a matrix embodiment of the molecular orbital

problem arises, as developed below and in Appendix F, by using the above LCAO-MO expansion in a variational treatment of the one-electron Schrödinger equation appropriate to the mos  $\{ \psi_i \}$ .

In the simplest two-center, two-valence-orbital case (which could relate, for example, to the  $\text{Li}_2$  molecule's two 2s orbitals), this gives rise to a 2x2 matrix eigenvalue problem  $(h_{11}, h_{12}, h_{22})$  with a low-energy mo  $(E = (h_{11} + h_{22})/2 - 1/2[(h_{11} - h_{22})^2 + 4h_{12}^2]^{1/2})$  and a higher energy mo  $(E = (h_{11} + h_{22})/2 + 1/2[(h_{11} - h_{22})^2 + 4h_{12}^2]^{1/2})$  corresponding to bonding and antibonding orbitals (because their energies lie below and above the lowest and highest interacting atomic orbital energies, respectively). The mos themselves are expressed  $\psi_i = \sum C_{ia} \phi_a$  where the LCAO-MO coefficients  $C_{ia}$  are obtained from the normalized eigenvectors of the  $h_{ab}$  matrix. Note that the bonding-antibonding orbital energy splitting depends on  $h_{ab}^2$  and on the energy difference  $(h_{aa} - h_{bb})$ ; the best bonding (and worst antibonding) occur when two orbitals couple strongly (have large  $h_{ab}$ ) and are similar in energy ( $h_{aa} \approx h_{bb}$ ).

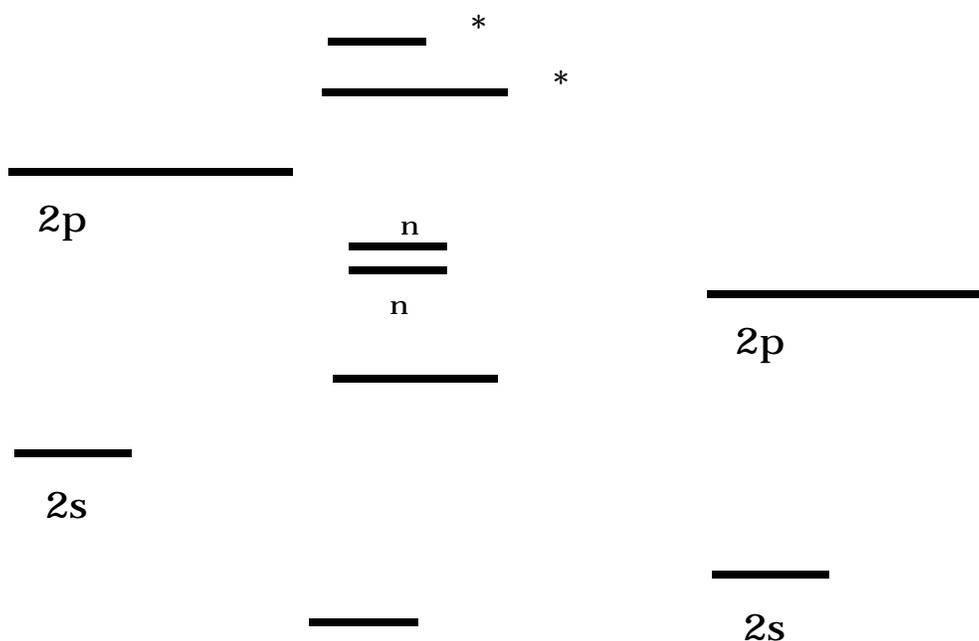


### Homonuclear Bonding With 2s and 2p Orbitals



### Heteronuclear Bonding With 2s and 2p Orbitals

In both the homonuclear and heteronuclear cases depicted above, the energy ordering of the resultant mos depends upon the energy ordering of the constituent aos as well as the strength of the bonding-antibonding interactions among the aos. For example, if the 2s-2p atomic orbital energy splitting is large compared with the interaction matrix elements coupling orbitals on neighboring atoms  $h_{2s,2s}$  and  $h_{2p,2p}$ , then the ordering shown above will result. On the other hand, if the 2s-2p splitting is small, the two 2s and two 2p orbitals can all participate in the formation of the four mos. In this case, it is useful to think of the atomic 2s and 2p orbitals forming sp hybrid orbitals with each atom having one hybrid directed toward the other atom and one hybrid directed away from the other atom. The resultant pattern of four mos will involve one bonding orbital (i.e., an in-phase combination of two sp hybrids), two non-bonding orbitals (those directed away from the other atom) and one antibonding orbital (an out-of-phase combination of two sp hybrids). Their energies will be ordered as shown in the Figure below.



Here  $n$  is used to denote the non-bonding  $\sigma$ -type orbitals and  $\sigma$ ,  $\sigma^*$ ,  $\pi$ , and  $\pi^*$  are used to denote bonding and antibonding  $\sigma$ - and  $\pi$ -type orbitals.

Notice that the total number of  $\sigma$  orbitals arising from the interaction of the 2s and 2p orbitals is equal to the number of aos that take part in their formation. Notice also that this is true regardless of whether one thinks of the interactions involving bare 2s and 2p

atomic orbitals or hybridized orbitals. The only advantage that the hybrids provide is that they permit one to foresee the fact that two of the four mos must be non-bonding because two of the four hybrids are directed away from all other valence orbitals and hence can not form bonds. In all such qualitative mo analyses, the final results (i.e., how many mos there are of any given symmetry) will not depend on whether one thinks of the interactions involving atomic or hybrid orbitals. However, it is often easier to "guess" the bonding, non-bonding, and antibonding nature of the resultant mos when thought of as formed from hybrids because of the directional properties of the hybrid orbitals.

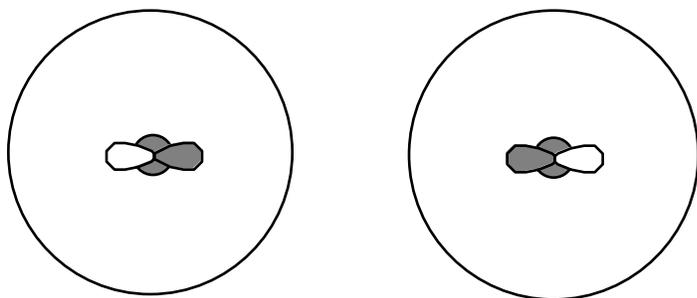
### C. Rydberg Orbitals

*It is essential to keep in mind that all atoms possess 'excited' orbitals that may become involved in bond formation if one or more electrons occupies these orbitals. Whenever aos with principal quantum number one or more unit higher than that of the conventional aos becomes involved in bond formation, Rydberg mos are formed.*

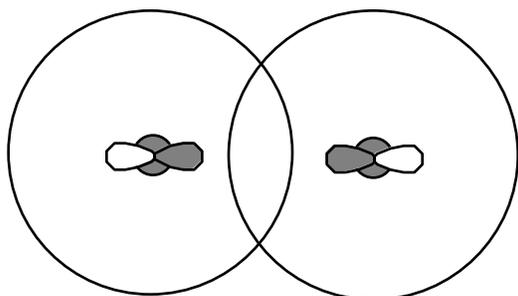
Rydberg orbitals (i.e., very diffuse orbitals having principal quantum numbers higher than the atoms' valence orbitals) can arise in molecules just as they do in atoms. They do not usually give rise to bonding and antibonding orbitals because the valence-orbital interactions bring the atomic centers so close together that the Rydberg orbitals of each atom subsume both atoms. Therefore as the atoms are brought together, the atomic Rydberg orbitals usually pass through the internuclear distance region where they experience (weak) bonding-antibonding interactions all the way to much shorter distances at which they have essentially reached their united-atom limits. As a result, molecular Rydberg orbitals are molecule-centered and display little, if any, bonding or antibonding character. They are usually labeled with principal quantum numbers beginning one higher than the highest  $n$  value of the constituent atomic valence orbitals, although they are sometimes labeled by the  $n$  quantum number to which they correlate in the united-atom limit.

An example of the interaction of 3s Rydberg orbitals of a molecule whose 2s and 2p orbitals are the valence orbitals and of the evolution of these orbitals into united-atom orbitals is given below.

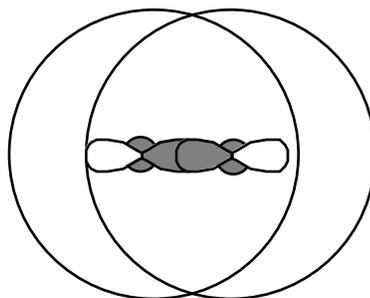




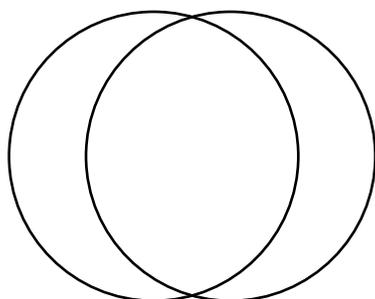
2s and 2p Valence Orbitals and 3s Rydberg Orbitals For Large R Values



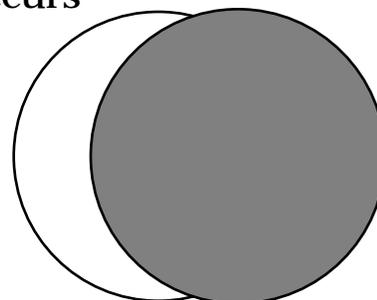
Overlap of the Rydberg Orbitals Begins



Rydberg Overlap is Strong and Bond Formation Occurs



The In-Phase (  $3s + 3s$  ) Combination of Rydberg Orbitals Correlates to an s-type Orbital of the United Atom



The Out-of-Phase (  $3s - 3s$  ) Combination of Rydberg Orbitals Correlates to a p-type United-Atom Orbital

## D. Multicenter Orbitals

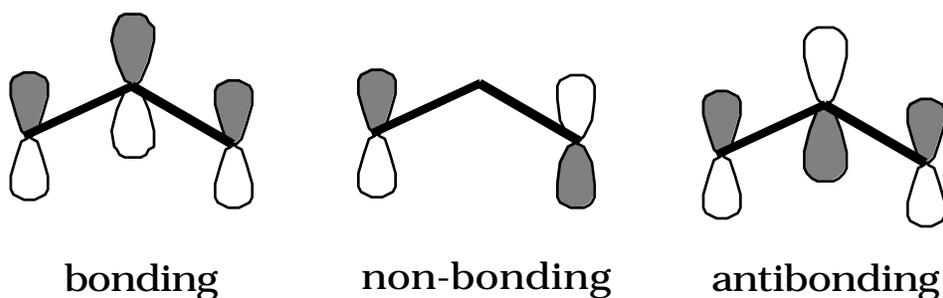
*If aos on one atom overlap aos on more than one neighboring atom, mos that involve amplitudes on three or more atomic centers can be formed. Such mos are termed delocalized or multicenter mos.*

Situations in which more than a pair of orbitals interact can, of course, occur. Three-center bonding occurs in Boron hydrides and in carbonyl bridge bonding in transition metal complexes as well as in delocalized conjugated orbitals common in unsaturated organic hydrocarbons. The three p orbitals on the allyl radical (considered in the absence of the underlying orbitals) can be described qualitatively in terms of three p aos on the three carbon atoms. The couplings  $h_{12}$  and  $h_{23}$  are equal (because the two CC bond lengths are the same) and  $h_{13}$  is approximated as 0 because orbitals 1 and 3 are too far away to interact. The result is a 3x3 secular matrix (see below and in Appendix F):

$$\begin{array}{ccc} h_{11} & h_{12} & 0 \\ h_{21} & h_{22} & h_{23} \\ 0 & h_{32} & h_{33} \end{array}$$

whose eigenvalues give the molecular orbital energies and whose eigenvectors give the LCAO-MO coefficients  $C_{ia}$ .

This 3x3 matrix gives rise to a bonding, a non-bonding and an antibonding orbital (see the Figure below). Since all of the  $h_{aa}$  are equal and  $h_{12} = h_{23}$ , the resultant orbital energies are:  $h_{11} + 2 h_{12}$ ,  $h_{11}$ , and  $h_{11} - 2 h_{12}$ , and the respective LCAO-MO coefficients  $C_{ia}$  are (0.50, 0.707, 0.50), (0.707, 0.00, -0.707), and (0.50, -0.707, 0.50). Notice that the sign (i.e., phase) relations of the bonding orbital are such that overlapping orbitals interact constructively, whereas for the antibonding orbital they interact out of phase. For the nonbonding orbital, there are no interactions because the central C orbital has zero amplitude in this orbital and only  $h_{12}$  and  $h_{23}$  are non-zero.



### Allyl System Orbitals

#### E. Hybrid Orbitals

*It is sometimes convenient to combine aos to form hybrid orbitals that have well defined directional character and to then form mos by combining these hybrid orbitals. This recombination of aos to form hybrids is never necessary and never provides any information that could be achieved in its absence. However, forming hybrids often allows one to focus on those interactions among directed orbitals on neighboring atoms that are most important.*

When atoms combine to form molecules, the molecular orbitals can be thought of as being constructed as linear combinations of the constituent atomic orbitals. This clearly is the only reasonable picture when each atom contributes only one orbital to the particular interactions being considered (e.g., as each Li atom does in  $\text{Li}_2$  and as each C atom does in the orbital aspect of the allyl system). However, when an atom uses more than one of its valence orbitals within particular bonding, non-bonding, or antibonding interactions, it is sometimes useful to combine the constituent atomic orbitals into hybrids and to then use the hybrid orbitals to describe the interactions. As stated above, the directional nature of hybrid orbitals often makes it more straightforward to "guess" the bonding, non-bonding, and antibonding nature of the resultant mos. It should be stressed, however, that exactly the same quantitative results are obtained if one forms mos from primitive aos or from hybrid orbitals; the hybrids span exactly the same space as the original aos and can therefore contain no additional information. This point is illustrated below when the  $\text{H}_2\text{O}$  and  $\text{N}_2$  molecules are treated in both the primitive ao and hybrid orbital bases.

## Chapter 5

### *Molecular Orbitals Possess Specific Topology, Symmetry, and Energy-Level Patterns*

In this chapter the symmetry properties of atomic, hybrid, and molecular orbitals are treated. It is important to keep in mind that both symmetry and characteristics of orbital energetics and bonding "topology", as embodied in the orbital energies themselves and the interactions (i.e.,  $h_{j,k}$  values) among the orbitals, are involved in determining the pattern of molecular orbitals that arise in a particular molecule.

#### I. Orbital Interaction Topology

*The pattern of mo energies can often be 'guessed' by using qualitative information about the energies, overlaps, directions, and shapes of the aos that comprise the mos.*

The orbital interactions determine how many and which mos will have low (bonding), intermediate (non-bonding), and higher (antibonding) energies, with all energies viewed relative to those of the constituent atomic orbitals. The **general patterns** that are observed in most compounds can be summarized as follows:

- i. If the energy splittings among a given atom's aos with the same principal quantum number are small, hybridization can easily occur to produce hybrid orbitals that are directed toward (and perhaps away from) the other atoms in the molecule. In the first-row elements (Li, Be, B, C, N, O, and F), the 2s-2p splitting is small, so hybridization is common. In contrast, for Ca, Ga, Ge, As, and Br it is less common, because the 4s-4p splitting is larger. Orbitals directed toward other atoms can form bonding and antibonding mos; those directed toward no other atoms will form nonbonding mos.
- ii. In attempting to gain a qualitative picture of the electronic structure of any given molecule, it is advantageous to begin by hybridizing the aos of those atoms which contain more than one ao in their valence shell. Only those aos that are not involved in  $\pi$ -orbital interactions should be so hybridized.
- iii. Atomic or hybrid orbitals that are not directed in a  $\pi$ -interaction manner toward other aos or hybrids on neighboring atoms can be involved in  $\sigma$ -interactions or in nonbonding interactions.

iv. Pairs of aos or hybrid orbitals on neighboring atoms directed toward one another interact to produce bonding and antibonding orbitals. The more the bonding orbital lies below the lower-energy ao or hybrid orbital involved in its formation, the higher the antibonding orbital lies above the higher-energy ao or hybrid orbital.

For example, in formaldehyde,  $\text{H}_2\text{CO}$ , one forms  $sp^2$  hybrids on the C atom; on the O atom, either  $sp$  hybrids (with one p orbital "reserved" for use in forming the  $\sigma$  and  $\pi^*$  orbitals and another p orbital to be used as a non-bonding orbital lying in the plane of the molecule) or  $sp^2$  hybrids (with the remaining p orbital reserved for the  $\sigma$  and  $\pi^*$  orbitals) can be used. The H atoms use their 1s orbitals since hybridization is not feasible for them. The C atom clearly uses its  $sp^2$  hybrids to form two CH and one CO  $\sigma$  bonding-antibonding orbital pairs.

The O atom uses one of its  $sp$  or  $sp^2$  hybrids to form the CO  $\sigma$  bond and antibond. When  $sp$  hybrids are used in conceptualizing the bonding, the other  $sp$  hybrid forms a lone pair orbital directed away from the CO bond axis; one of the atomic p orbitals is involved in the CO  $\sigma$  and  $\pi^*$  orbitals, while the other forms an in-plane non-bonding orbital. Alternatively, when  $sp^2$  hybrids are used, the two  $sp^2$  hybrids that do not interact with the C-atom  $sp^2$  orbital form the two non-bonding orbitals. Hence, the final picture of bonding, non-bonding, and antibonding orbitals does not depend on which hybrids one uses as intermediates.

As another example, the 2s and 2p orbitals on the two N atoms of  $\text{N}_2$  can be formed into pairs of  $sp$  hybrids on each N atom plus a pair of p atomic orbitals on each N atom. The  $sp$  hybrids directed toward the other N atom give rise to bonding  $\sigma$  and antibonding  $\sigma^*$  orbitals, and the  $sp$  hybrids directed away from the other N atom yield nonbonding  $\pi$  orbitals. The p orbitals, which consist of 2p orbitals on the N atoms directed perpendicular to the N-N bond axis, produce bonding  $\pi$  and antibonding  $\pi^*$  orbitals.

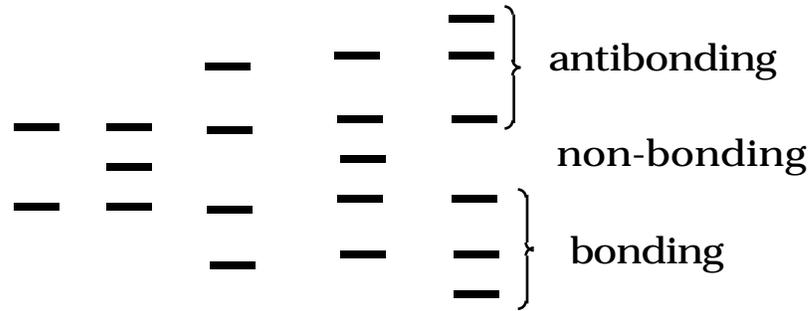
v. In general,  $\sigma$  interactions for a given pair of atoms interacting are stronger than  $\pi$  interactions (which, in turn, are stronger than  $\pi^*$  interactions, etc.) for any given sets (i.e., principal quantum number) of aos that interact. Hence,  $\sigma$  bonding orbitals (originating from a given set of aos) lie below  $\pi$  bonding orbitals, and  $\pi^*$  orbitals lie above  $\pi$  orbitals that arise from the same sets of aos. In the  $\text{N}_2$  example, the  $\sigma$  bonding orbital formed from the two  $sp$  hybrids lies below the  $\pi$  bonding orbital, but the  $\pi^*$  orbital lies below the  $\sigma^*$  orbital. In the  $\text{H}_2\text{CO}$  example, the two CH and the one CO bonding orbitals have low energy; the CO  $\sigma$  bonding orbital has the next lowest energy; the two O-atom non-bonding

orbitals have intermediate energy; the CO \* orbital has somewhat higher energy; and the two CH and one CO antibonding orbitals have the highest energies.

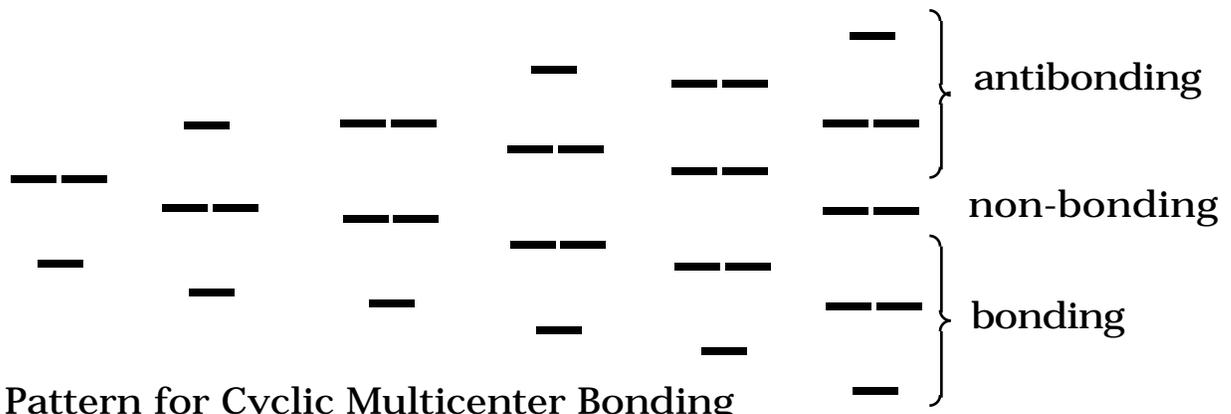
vi. If a given ao or hybrid orbital interacts with or is coupled to orbitals on more than a single neighboring atom, multicenter bonding can occur. For example, in the allyl radical the central carbon atom's p orbital is coupled to the p orbitals on both neighboring atoms; in linear Li<sub>3</sub>, the central Li atom's 2s orbital interacts with the 2s orbitals on both terminal Li atoms; in triangular Cu<sub>3</sub>, the 2s orbitals on each Cu atom couple to each of the other two atoms' 4s orbitals.

vii. Multicenter bonding that involves "linear" chains containing N atoms (e.g., as in conjugated polyenes or in chains of Cu or Na atoms for which the valence orbitals on one atom interact with those of its neighbors on both sides) gives rise to mo energy patterns in which there are N/2 (if N is even) or N/2 -1 non-degenerate bonding orbitals and the same number of antibonding orbitals (if N is odd, there is also a single non-bonding orbital).

viii. Multicenter bonding that involves "cyclic" chains of N atoms (e.g., as in cyclic conjugated polyenes or in rings of Cu or Na atoms for which the valence orbitals on one atom interact with those of its neighbors on both sides and the entire net forms a closed cycle) gives rise to mo energy patterns in which there is a lowest non-degenerate orbital and then a progression of doubly degenerate orbitals. If N is odd, this progression includes (N-1)/2 levels; if N is even, there are (N-2)/2 doubly degenerate levels and a final non-degenerate highest orbital. These patterns and those that appear in linear multicenter bonding are summarized in the Figures shown below.



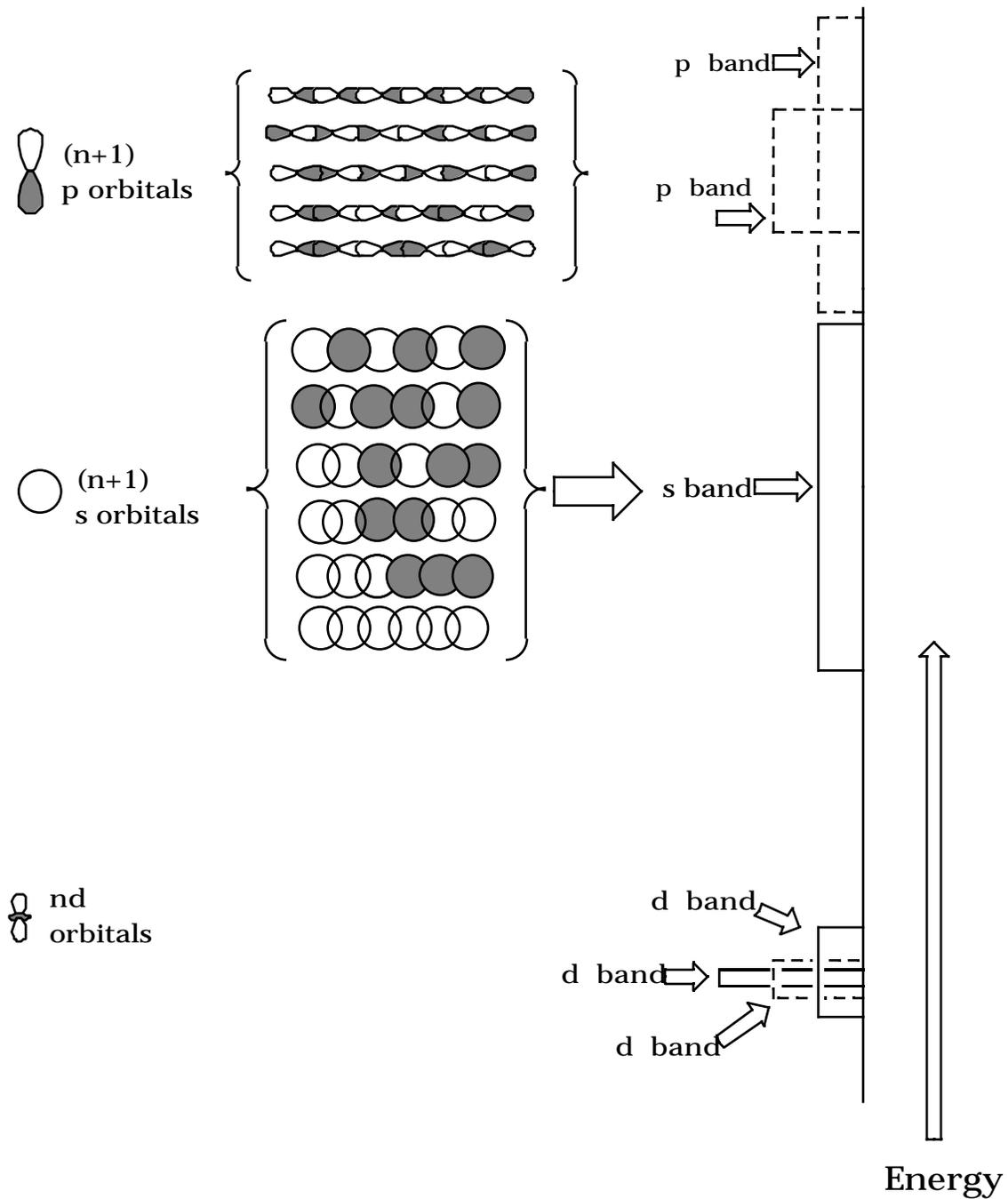
Pattern for Linear Multicenter  
Bonding Situation:  $N=2, 3, \dots, 6$



Pattern for Cyclic Multicenter Bonding  
 $N= 3, 4, 5, \dots, 8$

ix. In extended systems such as solids, atom-based orbitals combine as above to form so-called 'bands' of molecular orbitals. These bands are continuous rather than discrete as in the above cases involving small polyenes. The energy 'spread' within a band depends on the overlap among the atom-based orbitals that form the band; large overlap gives rise to a large band width, while small overlap produces a narrow band. As one moves from the bottom (i.e., the lower energy part) of a band to the top, the number of nodes in the corresponding band orbital increases, as a result of which its bonding nature decreases. In the figure shown below, the bands of a metal such as Ni (with 3d, 4s, and 4p orbitals) is illustrated. The d-orbital band is narrow because the 3d orbitals are small and hence do not overlap appreciably; the 4s and 4p bands are wider because the larger 4s and 4p orbitals overlap to a greater extent. The d-band is split into  $d_{xy}$ ,  $d_{yz}$ , and  $d_{zx}$  components corresponding to the nature of the overlap interactions among the constituent atomic d orbitals. Likewise,

the p-band is split into  $\pi$  and  $\sigma$  components. The widths of the  $\pi$  components of each band are larger than those of the  $\sigma$  components because the corresponding  $\pi$  overlap interactions are stronger. The intensities of the bands at energy  $E$  measure the densities of states at that  $E$ . The total integrated intensity under a given band is a measure of the total number of atomic orbitals that form the band.



## II. Orbital Symmetry

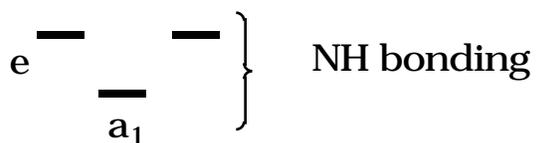
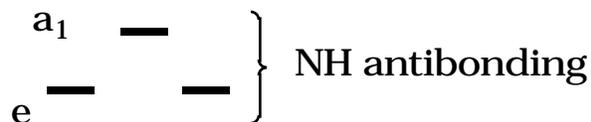
*Symmetry provides additional quantum numbers or labels to use in describing the mos. Each such quantum number further sub-divides the collection of all mos into sets that have vanishing Hamiltonian matrix elements among members belonging to different sets.*

Orbital interaction "topology" as discussed above plays a most- important role in determining the orbital energy level patterns of a molecule. Symmetry also comes into play but in a different manner. Symmetry can be used to characterize the core, bonding, non-bonding, and antibonding molecular orbitals. Much of this chapter is devoted to how this can be carried out in a systematic manner. Once the various mos have been labeled according to symmetry, it may be possible to recognize additional degeneracies that may not have been apparent on the basis of orbital-interaction considerations alone. Thus, topology provides the basic energy ordering pattern and then symmetry enters to identify additional degeneracies.

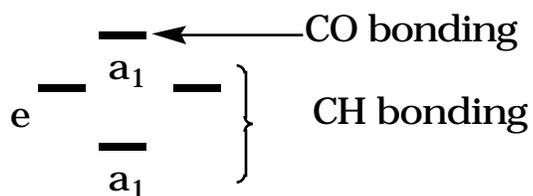
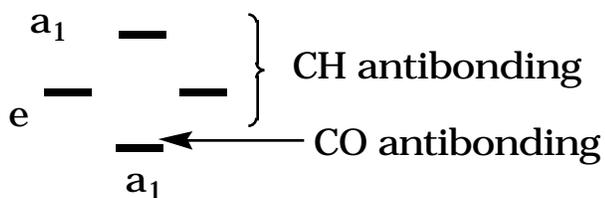
For example, the three NH bonding and three NH antibonding orbitals in  $\text{NH}_3$ , when symmetry adapted within the  $C_{3v}$  point group, cluster into a<sub>1</sub> and e mos as shown in the Figure below. The N-atom localized non-bonding lone pair orbital and the N-atom 1s core orbital also belong to a<sub>1</sub> symmetry.

In a second example, the three CH bonds, three CH antibonds, CO bond and antibond, and three O-atom non-bonding orbitals of the methoxy radical  $\text{H}_3\text{C-O}$  also cluster into a<sub>1</sub> and e orbitals as shown below. In these cases, point group symmetry allows one to identify degeneracies that may not have been apparent from the structure of the orbital interactions alone.





### Orbital Character and Symmetry in NH<sub>3</sub>

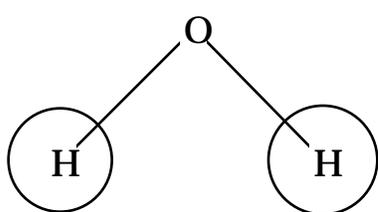


### Orbital Character and Symmetry in H<sub>3</sub>CO Radical

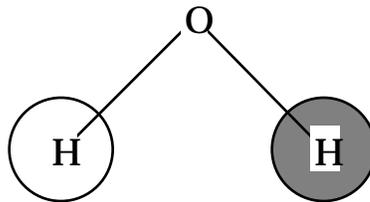


The three resultant molecular orbital energies are, of course, identical to those obtained without symmetry above. The three LCAO-MO coefficients, now expressing the mos in terms of the symmetry adapted orbitals are  $C_{is} = (0.707, 0.707, 0.0)$  for the bonding orbital,  $(0.0, 0.0, 1.00)$  for the nonbonding orbital, and  $(0.707, -0.707, 0.0)$  for the antibonding orbital. These coefficients, when combined with the symmetry adaptation coefficients  $C_{sa}$  given earlier, express the three mos in terms of the three aos as  $\psi_i = \sum_{sa} C_{is} C_{sa} \phi_{sa}$ ; the sum  $\sum_{sa} C_{is} C_{sa}$  gives the LCAO-MO coefficients  $C_{ia}$  which, for example, for the bonding orbital, are  $(0.707^2, 0.707, 0.707^2)$ , in agreement with what was found earlier without using symmetry.

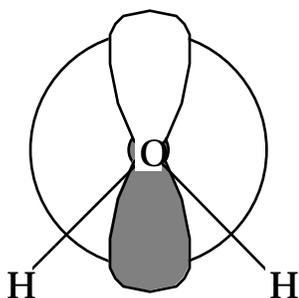
The low energy orbitals of the  $H_2O$  molecule can be used to illustrate the use of symmetry within the primitive ao basis as well as in terms of hybrid orbitals. The 1s orbital on the Oxygen atom is clearly a nonbonding core orbital. The Oxygen 2s orbital and its three 2p orbitals are of valence type, as are the two Hydrogen 1s orbitals. In the absence of symmetry, these six valence orbitals would give rise to a 6x6 secular problem. By combining the two Hydrogen 1s orbitals into  $0.707(1s_L + 1s_R)$  and  $0.707(1s_L - 1s_R)$  symmetry adapted orbitals (labeled  $a_1$  and  $b_2$  within the  $C_{2v}$  point group; see the Figure below), and recognizing that the Oxygen 2s and  $2p_z$  orbitals belong to  $a_1$  symmetry (the z axis is taken as the  $C_2$  rotation axis and the x axis is taken to be perpendicular to the plane in which the three nuclei lie) while the  $2p_x$  orbital is  $b_1$  and the  $2p_y$  orbital is  $b_2$ , allows the 6x6 problem to be decomposed into a 3x3 ( $a_1$ ) secular problem, a 2x2 ( $b_2$ ) secular problem and a 1x1 ( $b_1$ ) problem. These decompositions allow one to conclude that there is one nonbonding  $b_1$  orbital (the Oxygen  $2p_x$  orbital), bonding and antibonding  $b_2$  orbitals (the O-H bond and antibond formed by the Oxygen  $2p_y$  orbital interacting with  $0.707(1s_L - 1s_R)$ ), and, finally, a set of bonding, nonbonding, and antibonding  $a_1$  orbitals (the O-H bond and antibond formed by the Oxygen 2s and  $2p_z$  orbitals interacting with  $0.707(1s_L + 1s_R)$  and the nonbonding orbital formed by the Oxygen 2s and  $2p_z$  orbitals combining to form the "lone pair" orbital directed along the z-axis away from the two Hydrogen atoms).



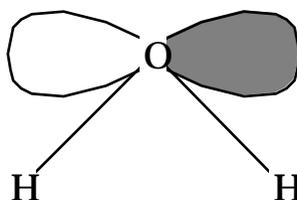
**a<sub>1</sub> Hydrogen Orbitals**



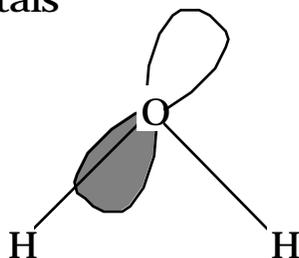
**b<sub>2</sub> Hydrogen Orbitals**



**Oxygen a<sub>1</sub> Orbitals**



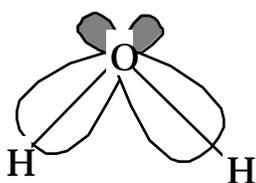
**Oxygen b<sub>2</sub> Orbital**



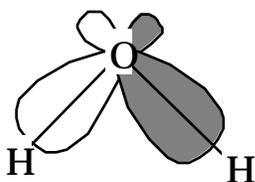
**Oxygen b<sub>1</sub> Orbital**

Alternatively, to analyze the H<sub>2</sub>O molecule in terms of hybrid orbitals, one first combines the Oxygen 2s, 2p<sub>z</sub>, 2p<sub>x</sub> and 2p<sub>y</sub> orbitals to form four sp<sup>3</sup> hybrid orbitals. The valence-shell electron-pair repulsion (VSEPR) model of chemical bonding (see R. J. Gillespie and R. S. Nyholm, *Quart. Rev.* 11, 339 (1957) and R. J. Gillespie, *J. Chem. Educ.* 40, 295 (1963)) directs one to involve all of the Oxygen valence orbitals in the hybridization because four  $\sigma$ -bond or nonbonding electron pairs need to be accommodated about the Oxygen center; no  $\pi$  orbital interactions are involved, of course. Having formed the four sp<sup>3</sup> hybrid orbitals, one proceeds as with the primitive aos; one forms symmetry

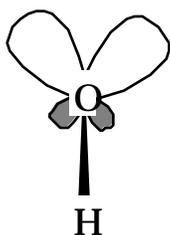
adapted orbitals. In this case, the two Hydrogen  $1s$  orbitals are combined exactly as above to form  $0.707(1s_L + 1s_R)$  and  $0.707(1s_L - 1s_R)$ . The two  $sp^3$  hybrids which lie in the plane of the H and O nuclei (label them L and R) are combined to give symmetry adapted hybrids:  $0.707(L+R)$  and  $0.707(L-R)$ , which are of  $a_1$  and  $b_2$  symmetry, respectively (see the Figure below). The two  $sp^3$  hybrids that lie above and below the plane of the three nuclei (label them T and B) are also symmetry adapted to form  $0.707(T+B)$  and  $0.707(T-B)$ , which are of  $a_1$  and  $b_1$  symmetry, respectively. Once again, one has broken the  $6 \times 6$  secular problem into a  $3 \times 3$   $a_1$  block, a  $2 \times 2$   $b_2$  block and a  $1 \times 1$   $b_1$  block. Although the resulting bonding, nonbonding and antibonding  $a_1$  orbitals, the bonding and antibonding  $b_2$  orbitals and the nonbonding  $b_1$  orbital are now viewed as formed from symmetry adapted Hydrogen orbitals and four Oxygen  $sp^3$  orbitals, they are, of course, exactly the same molecular orbitals as were obtained earlier in terms of the symmetry adapted primitive aos. The formation of hybrid orbitals was an intermediate step which could not alter the final outcome.



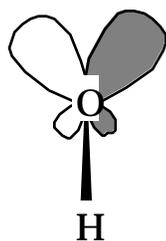
**L + R  $a_1$  Hybrid  
Symmetry Orbital**



**L - R  $b_2$  Hybrid  
Symmetry Orbital**



**T + B Hybrid  
Symmetry Orbital  
Seen From the Side**



**T - B Hybrid  
Symmetry Orbital  
Seen From the Side**

That no degenerate molecular orbitals arose in the above examples is a result of the fact that the  $C_{2v}$  point group to which  $H_2O$  and the allyl system belong (and certainly the

$C_s$  subgroup which was used above in the allyl case) has no degenerate representations. Molecules with higher symmetry such as  $NH_3$ ,  $CH_4$ , and benzene have energetically degenerate orbitals because their molecular point groups have degenerate representations.

## B. Linear Molecules

*Linear molecules belong to the axial rotation group. Their symmetry is intermediate in complexity between nonlinear molecules and atoms.*

For linear molecules, the symmetry of the electrostatic potential provided by the nuclei and the other electrons is described by either the  $C_{\infty v}$  or  $D_{\infty h}$  group. The essential difference between these symmetry groups and the finite point groups which characterize the non-linear molecules lies in the fact that the electrostatic potential which an electron feels is invariant to rotations of any amount about the molecular axis (i.e.,  $V(\theta + \phi) = V(\theta)$ , for any angle increment  $\phi$ ). This means that the operator  $C_{\infty}$  which generates a rotation of the electron's azimuthal angle  $\phi$  by an amount  $\phi$  about the molecular axis commutes with the Hamiltonian  $[H, C_{\infty}] = 0$ .  $C_{\infty}$  can be written in terms of the quantum mechanical operator  $L_z = -i\hbar \frac{\partial}{\partial \phi}$  describing the orbital angular momentum of the electron about the molecular (z) axis:

$$C_{\infty} = \exp(i L_z / \hbar).$$

Because  $C_{\infty}$  commutes with the Hamiltonian and  $C_{\infty}$  can be written in terms of  $L_z$ ,  $L_z$  must commute with the Hamiltonian. As a result, the molecular orbitals of a linear molecule must be eigenfunctions of the z-component of angular momentum  $L_z$ :

$$-i\hbar \frac{\partial}{\partial \phi} \psi = m\hbar \psi.$$

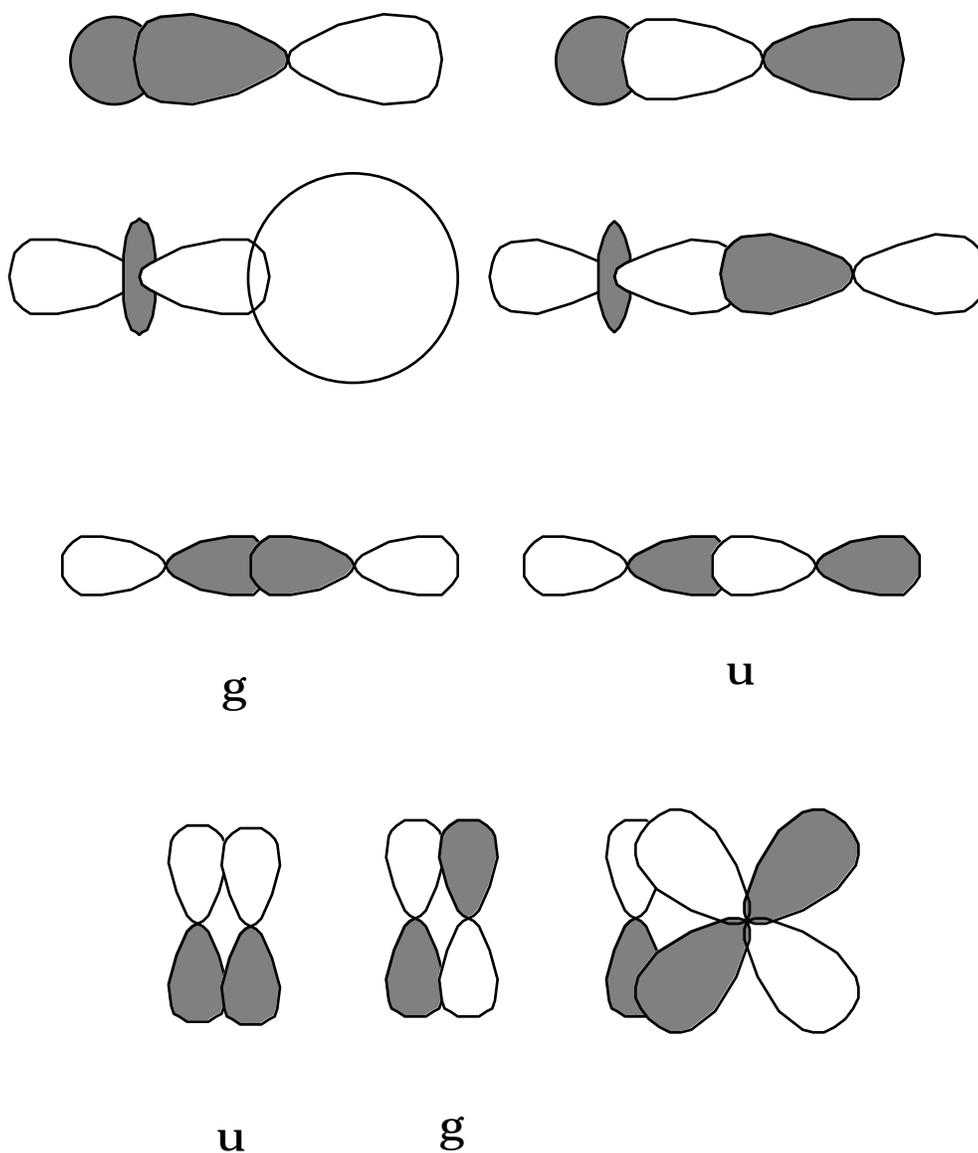
The electrostatic potential is not invariant under rotations of the electron about the x or y axes (those perpendicular to the molecular axis), so  $L_x$  and  $L_y$  do not commute with the Hamiltonian. Therefore, only  $L_z$  provides a "good quantum number" in the sense that the operator  $L_z$  commutes with the Hamiltonian.

In summary, the molecular orbitals of a linear molecule can be labeled by their m quantum number, which plays the same role as the point group labels did for non-linear polyatomic molecules, and which gives the eigenvalue of the angular momentum of the orbital about the molecule's symmetry axis. Because the kinetic energy part of the

Hamiltonian contains  $(\hbar^2/2m_e r^2) \propto 1/r^2$ , whereas the potential energy part is independent of  $\theta$ , the energies of the molecular orbitals depend on the square of the  $m$  quantum number. Thus, pairs of orbitals with  $m = \pm 1$  are energetically degenerate; pairs with  $m = \pm 2$  are degenerate, and so on. The absolute value of  $m$ , which is what the energy depends on, is called the  $|m|$  quantum number. Molecular orbitals with  $|m| = 0$  are called  $\sigma$  orbitals; those with  $|m| = 1$  are  $\pi$  orbitals; and those with  $|m| = 2$  are  $\delta$  orbitals.

Just as in the non-linear polyatomic-molecule case, the atomic orbitals which constitute a given molecular orbital must have the same symmetry as that of the molecular orbital. This means that  $\sigma$ ,  $\pi$ , and  $\delta$  molecular orbitals are formed, via LCAO-MO, from  $m=0$ ,  $m = \pm 1$ , and  $m = \pm 2$  atomic orbitals, respectively. In the diatomic  $N_2$  molecule, for example, the core orbitals are of  $\sigma$  symmetry as are the molecular orbitals formed from the  $2s$  and  $2p_z$  atomic orbitals (or their hybrids) on each Nitrogen atom. The molecular orbitals formed from the atomic  $2p_{-1} = (2p_x - i 2p_y)$  and the  $2p_{+1} = (2p_x + i 2p_y)$  orbitals are of  $\pi$  symmetry and have  $m = -1$  and  $+1$ .

For homonuclear diatomic molecules and other linear molecules which have a center of symmetry, the inversion operation (in which an electron's coordinates are inverted through the center of symmetry of the molecule) is also a symmetry operation. Each resultant molecular orbital can then also be labeled by a quantum number denoting its parity with respect to inversion. The symbols *g* (for gerade or even) and *u* (for ungerade or odd) are used for this label. Again for  $N_2$ , the core orbitals are of *g* and *u* symmetry, and the bonding and antibonding orbitals formed from the 2s and 2p orbitals on the two Nitrogen atoms are of *g* and *u* symmetry.



The bonding molecular orbital pair (with  $m = +1$  and  $-1$ ) is of  $\sigma_u$  symmetry whereas the corresponding antibonding orbital is of  $\sigma_g$  symmetry. Examples of such molecular orbital symmetries are shown above.

The use of hybrid orbitals can be illustrated in the linear-molecule case by considering the  $N_2$  molecule. Because two bonding and antibonding molecular orbital pairs are involved in  $N_2$  (one with  $m = +1$ , one with  $m = -1$ ), VSEPR theory guides one to form  $sp$  hybrid orbitals from each of the Nitrogen atom's  $2s$  and  $2p_z$  (which is also the  $2p$  orbital with  $m = 0$ ) orbitals. Ignoring the core orbitals, which are of  $\sigma_g$  and  $\sigma_u$  symmetry as noted above, one then symmetry adapts the four  $sp$  hybrids (two from each atom) to build one  $\sigma_g$  orbital involving a bonding interaction between two  $sp$  hybrids pointed toward one another, an antibonding  $\sigma_u$  orbital involving the same pair of  $sp$  orbitals but coupled with opposite signs, a nonbonding  $\sigma_g$  orbital composed of two  $sp$  hybrids pointed away from the interatomic region combined with like sign, and a nonbonding  $\sigma_u$  orbital made of the latter two  $sp$  hybrids combined with opposite signs. The two  $2p_m$  orbitals ( $m = +1$  and  $-1$ ) on each Nitrogen atom are then symmetry adapted to produce a pair of bonding  $\pi_u$  orbitals (with  $m = +1$  and  $-1$ ) and a pair of antibonding  $\pi_g$  orbitals (with  $m = +1$  and  $-1$ ). This hybridization and symmetry adaptation thereby reduces the  $8 \times 8$  secular problem (which would be  $10 \times 10$  if the core orbitals were included) into a  $2 \times 2$   $\sigma_g$  problem (one bonding and one nonbonding), a  $2 \times 2$   $\sigma_u$  problem (one bonding and one nonbonding), an identical pair of  $1 \times 1$   $\pi_u$  problems (bonding), and an identical pair of  $1 \times 1$   $\pi_g$  problems (antibonding).

Another example of the equivalence among various hybrid and atomic orbital points of view is provided by the  $CO$  molecule. Using, for example,  $sp$  hybrid orbitals on  $C$  and  $O$ , one obtains a picture in which there are: two core orbitals corresponding to the  $O$ -atom  $1s$  and  $C$ -atom  $1s$  orbitals; one  $CO$  bonding, two non-bonding, and one  $CO$  antibonding orbitals arising from the four  $sp$  hybrids; a pair of bonding and a pair of antibonding orbitals formed from the two  $p$  orbitals on  $O$  and the two  $p$  orbitals on  $C$ . Alternatively, using  $sp^2$  hybrids on both  $C$  and  $O$ , one obtains: the two core orbitals as above; a  $CO$  bonding and antibonding orbital pair formed from the  $sp^2$  hybrids that are directed along the  $CO$  bond; and a single bonding and antibonding  $\pi^*$  orbital set. The remaining two  $sp^2$  orbitals on  $C$  and the two on  $O$  can then be symmetry adapted by forming  $\pm$  combinations within each pair to yield: an  $a_1$  non-bonding orbital (from the  $+$  combination) on each of  $C$  and  $O$  directed away from the  $CO$  bond axis; and a  $p$  orbital on each of  $C$  and  $O$  that can subsequently overlap to form the second bonding and  $\pi^*$  antibonding orbital pair.

It should be clear from the above examples, that no matter what particular hybrid

orbitals one chooses to utilize in conceptualizing a molecule's orbital interactions, symmetry ultimately returns to force one to form proper symmetry adapted combinations which, in turn, renders the various points of view equivalent. In the above examples and in several earlier examples, symmetry adaptation of, for example,  $sp^2$  orbital pairs (e.g.,  $sp_L^2 \pm sp_R^2$ ) generated orbitals of pure spatial symmetry. In fact, symmetry combining hybrid orbitals in this manner amounts to forming other hybrid orbitals. For example, the above  $\pm$  combinations of  $sp^2$  hybrids directed to the left (L) and right (R) of some bond axis generate a new  $sp$  hybrid directed along the bond axis but opposite to the  $sp^2$  hybrid used to form the bond and a non-hybridized p orbital directed along the L-to-R direction. In the CO example, these combinations of  $sp^2$  hybrids on O and C produce  $sp$  hybrids on O and C and p orbitals on O and C.

### C. Atoms

*Atoms belong to the full rotation symmetry group; this makes their symmetry analysis the most complex to treat.*

In moving from linear molecules to atoms, additional symmetry elements arise. In particular, the potential field experienced by an electron in an orbital becomes invariant to rotations of arbitrary amounts about the x, y, and z axes; in the linear-molecule case, it is invariant only to rotations of the electron's position about the molecule's symmetry axis (the z axis). These invariances are, of course, caused by the spherical symmetry of the potential of any atom. This additional symmetry of the potential causes the Hamiltonian to commute with all three components of the electron's angular momentum:  $[L_x, H] = 0$ ,  $[L_y, H] = 0$ , and  $[L_z, H] = 0$ . It is straightforward to show that H also commutes with the operator  $L^2 = L_x^2 + L_y^2 + L_z^2$ , defined as the sum of the squares of the three individual components of the angular momentum. Because  $L_x$ ,  $L_y$ , and  $L_z$  do not commute with one another, orbitals which are eigenfunctions of H cannot be simultaneous eigenfunctions of all three angular momentum operators. Because  $L_x$ ,  $L_y$ , and  $L_z$  do commute with  $L^2$ , orbitals can be found which are eigenfunctions of H, of  $L^2$  and of any one component of  $\mathbf{L}$ ; it is convention to select  $L_z$  as the operator which, along with H and  $L^2$ , form a mutually commutative operator set of which the orbitals are simultaneous eigenfunctions.

So, for any atom, the orbitals can be labeled by both l and m quantum numbers, which play the role that point group labels did for non-linear molecules and did for linear molecules. Because (i) the kinetic energy operator in the electronic Hamiltonian explicitly contains  $L^2/2m_e r^2$ , (ii) the Hamiltonian does not contain additional  $L_z$ ,  $L_x$ , or  $L_y$  factors,

and (iii) the potential energy part of the Hamiltonian is spherically symmetric (and commutes with  $L^2$  and  $L_z$ ), the energies of atomic orbitals depend upon the  $l$  quantum number and are independent of the  $m$  quantum number. This is the source of the  $2l+1$ -fold degeneracy of atomic orbitals.

The angular part of the atomic orbitals is described in terms of the spherical harmonics  $Y_{l,m}$ ; that is, each atomic orbital can be expressed as

$$\psi_{n,l,m} = Y_{l,m}(\theta, \phi) R_{n,l}(r).$$

The explicit solutions for the  $Y_{l,m}$  and for the radial wavefunctions  $R_{n,l}$  are given in Appendix B. The variables  $r, \theta, \phi$  give the position of the electron in the orbital in spherical coordinates. These angular functions are, as discussed earlier, related to the cartesian (i.e., spatially oriented) orbitals by simple transformations; for example, the orbitals with  $l=2$  and  $m=2,1,0,-1,-2$  can be expressed in terms of the  $d_{xy}, d_{xz}, d_{yz}, d_{xx-yy}$ , and  $d_{zz}$  orbitals. Either set of orbitals is acceptable in the sense that each orbital is an eigenfunction of  $H$ ; transformations within a degenerate set of orbitals do not destroy the Hamiltonian-eigenfunction feature. The orbital set labeled with  $l$  and  $m$  quantum numbers is most useful when one is dealing with isolated atoms (which have spherical symmetry), because  $m$  is then a valid symmetry label, or with an atom in a local environment which is axially symmetric (e.g., in a linear molecule) where the  $m$  quantum number remains a useful symmetry label. The cartesian orbitals are preferred for describing an atom in a local environment which displays lower than axial symmetry (e.g., an atom interacting with a diatomic molecule in  $C_{2v}$  symmetry).

The radial part of the orbital  $R_{n,l}(r)$  as well as the orbital energy  $\epsilon_{n,l}$  depend on  $l$  because the Hamiltonian itself contains  $l(l+1)\hbar^2/2m_e r^2$ ; they are independent of  $m$  because the Hamiltonian has no  $m$ -dependence. For bound orbitals,  $R_{n,l}(r)$  decays exponentially for large  $r$  (as  $\exp(-2r/a_0 n)$ ), and for unbound (scattering) orbitals, it is oscillatory at large  $r$  with an oscillation period related to the deBroglie wavelength of the electron. In  $R_{n,l}(r)$  there are  $(n-l-1)$  radial nodes lying between  $r=0$  and  $r=\infty$ . These nodes provide differential stabilization of low- $l$  orbitals over high- $l$  orbitals of the same principal quantum number  $n$ . That is, penetration of outer shells is greater for low- $l$  orbitals because they have more radial nodes; as a result, they have larger amplitude near the atomic nucleus and thus experience enhanced attraction to the positive nuclear charge. The average size (e.g., average value of  $r$ ;  $\langle r \rangle = \int R_{n,l}^2 r^3 dr$ ) of an orbital depends strongly on  $n$ , weakly on  $l$  and is independent of  $m$ ; it also depends strongly on the nuclear charge and on the potential produced by the other electrons. This potential is often characterized qualitatively in terms

of an effective nuclear charge  $Z_{\text{eff}}$  which is the true nuclear charge of the atom  $Z$  minus a screening component  $Z_{\text{sc}}$  which describes the repulsive effect of the electron density lying radially inside the electron under study. Because, for a given  $n$ , low- $l$  orbitals penetrate closer to the nucleus than do high- $l$  orbitals, they have higher  $Z_{\text{eff}}$  values (i.e., smaller  $Z_{\text{sc}}$  values) and correspondingly smaller average sizes and larger binding energies.



## Chapter 6

*Along "Reaction Paths", Orbitals Can be Connected One-to-One According to Their Symmetries and Energies. This is the Origin of the Woodward-Hoffmann Rules*

### I. Reduction in Symmetry

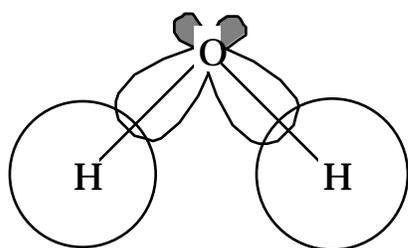
*As fragments are brought together to form a larger molecule, the symmetry of the nuclear framework (recall the symmetry of the coulombic potential experienced by electrons depends on the locations of the nuclei) changes. However, in some cases, certain symmetry elements persist throughout the path connecting the fragments and the product molecule. These preserved symmetry elements can be used to label the orbitals throughout the 'reaction'.*

The point-group, axial- and full-rotation group symmetries which arise in non-linear molecules, linear molecules, and atoms, respectively, are seen to provide quantum numbers or symmetry labels which can be used to characterize the orbitals appropriate for each such species. In a physical event such as interaction with an external electric or magnetic field or a chemical process such as collision or reaction with another species, the atom or molecule can experience a change in environment which causes the electrostatic potential which its orbitals experience to be of lower symmetry than that of the isolated atom or molecule. For example, when an atom interacts with another atom to form a diatomic molecule or simply to exchange energy during a collision, each atom's environment changes from being spherically symmetric to being axially symmetric. When the formaldehyde molecule undergoes unimolecular decomposition to produce CO + H<sub>2</sub> along a path that preserves C<sub>2v</sub> symmetry, the orbitals of the CO moiety evolve from C<sub>2v</sub> symmetry to axial symmetry.

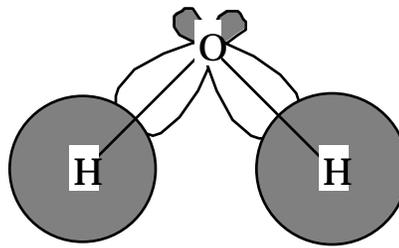
It is important, therefore to be able to label the orbitals of atoms, linear, and non-linear molecules in terms of their full symmetries as well in terms of the groups appropriate to lower-symmetry situations. This can be done by knowing how the representations of a higher symmetry group decompose into representations of a lower group. For example, the Y<sub>l,m</sub> functions appropriate for spherical symmetry, which belong to a 2l+1 fold degenerate set in this higher symmetry, decompose into doubly degenerate pairs of functions Y<sub>l,1</sub>, Y<sub>l,-1</sub>; Y<sub>l,1-1</sub>, Y<sub>l,-1+1</sub>; etc., plus a single non-degenerate function Y<sub>l,0</sub>, in axial symmetry. Moreover, because L<sup>2</sup> no longer commutes with the Hamiltonian whereas L<sub>z</sub> does, orbitals with different l-values but the same m-values can be coupled. As the N<sub>2</sub> molecule is formed from two N atoms, the 2s and 2p<sub>z</sub> orbitals, both of which belong to the same ( ) symmetry in the axial rotation group but which are of different symmetry in the isolated-atom

spherical symmetry, can mix to form the  $\sigma_g$  bonding orbital, the  $\sigma_u^*$  antibonding, as well as the  $\pi_g$  and  $\pi_u$  nonbonding lone-pair orbitals. The fact that 2s and 2p have different l-values no longer uncouples these orbitals as it did for the isolated atoms, because l is no longer a "good" quantum number.

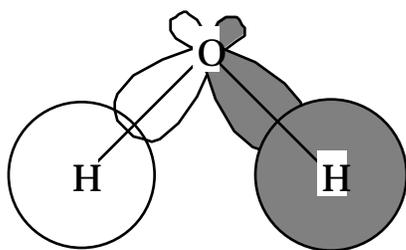
Another example of reduced symmetry is provided by the changes that occur as  $H_2O$  fragments into OH and H. The  $\sigma$  bonding orbitals ( $a_1$  and  $b_2$ ) and in-plane lone pair ( $a_1$ ) and the  $\pi^*$  antibonding ( $a_1$  and  $b_2$ ) of  $H_2O$  become  $a'$  orbitals (see the Figure below); the out-of-plane  $b_1$  lone pair orbital becomes  $a''$  (in Appendix IV of Electronic Spectra and Electronic Structure of Polyatomic Molecules, G. Herzberg, Van Nostrand Reinhold Co., New York, N.Y. (1966) tables are given which allow one to determine how particular symmetries of a higher group evolve into symmetries of a lower group).



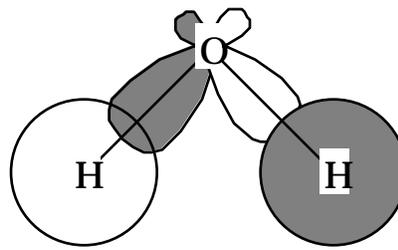
$a_1$  bonding orbital



$a_1^*$  antibonding orbital



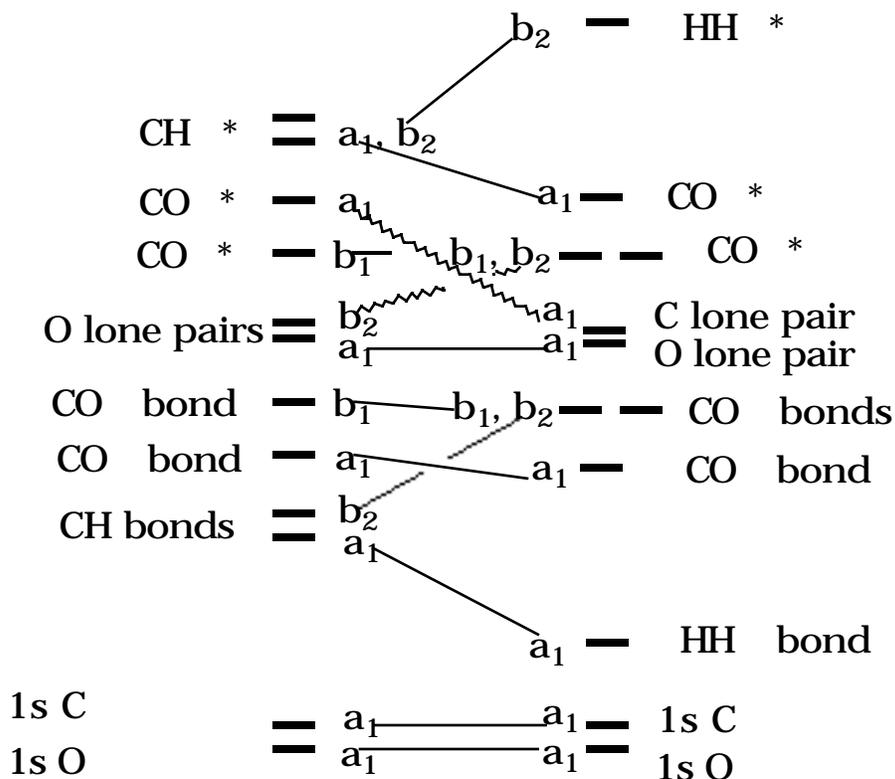
$b_2$  bonding orbital



$b_2^*$  antibonding orbital

To further illustrate these points dealing with orbital symmetry, consider the insertion of CO into  $H_2$  along a path which preserves  $C_{2v}$  symmetry. As the insertion occurs, the degenerate  $\pi$  bonding orbitals of CO become  $b_1$  and  $b_2$  orbitals. The antibonding  $\pi^*$  orbitals of CO also become  $b_1$  and  $b_2$ . The  $\sigma_g$  bonding orbital of  $H_2$  becomes  $a_1$ , and the  $\sigma_u^*$  antibonding  $H_2$  orbital becomes  $b_2$ . The orbitals of the reactant

H<sub>2</sub>CO are energy-ordered and labeled according to C<sub>2v</sub> symmetry in the Figure shown below as are the orbitals of the product H<sub>2</sub> + CO.



### H<sub>2</sub>CO ==> H<sub>2</sub> + CO Orbital Correlation Diagram in C<sub>2v</sub> Symmetry

When these orbitals are connected according to their symmetries as shown above, one reactant orbital to one product orbital starting with the low-energy orbitals and working to increasing energy, an orbital correlation diagram (OCD) is formed. These diagrams play essential roles in analyzing whether reactions will have symmetry-imposed energy barriers on their potential energy surfaces along the reaction path considered in the symmetry analysis. The essence of this analysis, which is covered in detail in Chapter 12, can be understood by noticing that the sixteen electrons of ground-state H<sub>2</sub>CO do not occupy their orbitals with the same occupancy pattern, symmetry-by-symmetry, as do the sixteen electrons of ground-state H<sub>2</sub> + CO. In particular, H<sub>2</sub>CO places a pair of electrons in the second b<sub>2</sub> orbital while H<sub>2</sub> + CO does not; on the other hand, H<sub>2</sub> + CO places two electrons in the sixth a<sub>1</sub> orbital while H<sub>2</sub>CO does not. The mismatch of the orbitals near the 5a<sub>1</sub>, 6a<sub>1</sub>, and 2b<sub>2</sub> orbitals is the source of the mismatch in the electronic configurations of the ground-states of H<sub>2</sub>CO and H<sub>2</sub> + CO. These mismatches give rise, as shown in

Chapter 12, to symmetry-caused energy barriers on the  $\text{H}_2\text{CO} \rightleftharpoons \text{H}_2 + \text{CO}$  reaction potential energy surface.

## II. Orbital Correlation Diagrams

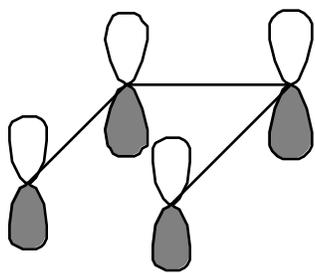
*Connecting the energy-ordered orbitals of reactants to those of products according to symmetry elements that are preserved throughout the reaction produces an orbital correlation diagram.*

In each of the examples cited above, symmetry reduction occurred as a molecule or atom approached and interacted with another species. The "path" along which this approach was thought to occur was characterized by symmetry in the sense that it preserved certain symmetry elements while destroying others. For example, the collision of two Nitrogen atoms to produce  $\text{N}_2$  clearly occurs in a way which destroys spherical symmetry but preserves axial symmetry. In the other example used above, the formaldehyde molecule was postulated to decompose along a path which preserves  $\text{C}_{2v}$  symmetry while destroying the axial symmetries of CO and  $\text{H}_2$ . The actual decomposition of formaldehyde may occur along some other path, but if it were to occur along the proposed path, then the symmetry analysis presented above would be useful.

The symmetry reduction analysis outlined above allows one to see new orbital interactions that arise (e.g., the  $2s$  and  $2p_z$  interactions in the  $\text{N} + \text{N} \rightleftharpoons \text{N}_2$  example) as the interaction increases. It also allows one to construct orbital correlation diagrams (OCD's) in which the orbitals of the "reactants" and "products" are energy ordered and labeled by the symmetries which are preserved throughout the "path", and the orbitals are then correlated by drawing lines connecting the orbitals of a given symmetry, one-by-one in increasing energy, from the reactants side of the diagram to the products side. As noted above, such orbital correlation diagrams play a central role in using symmetry to predict whether photochemical and thermal chemical reactions will experience activation barriers along proposed reaction paths (this subject is treated in Chapter 12).

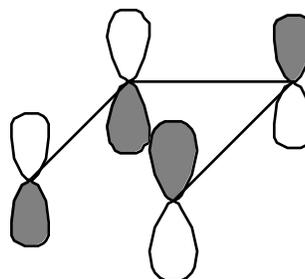
To again illustrate the construction of an OCD, consider the orbitals of 1,3-butadiene as the molecule undergoes disrotatory closing (notice that this is where a particular path is postulated; the actual reaction may or may not occur along such a path) to form cyclobutene. Along this path, the plane of symmetry which bisects and is perpendicular to the  $\text{C}_2\text{-C}_3$  bond is preserved, so the orbitals of the reactant and product are labeled as being even-e or odd-o under reflection through this plane. It is not proper to label the orbitals with respect to their symmetry under the plane containing the four C atoms;

although this plane is indeed a symmetry operation for the reactants and products, it does not remain a valid symmetry throughout the reaction path.

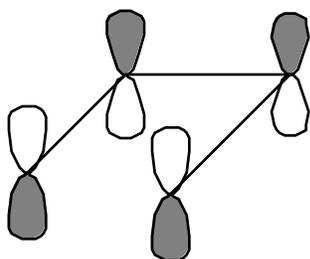


Lowest orbital of  
1,3- butadiene denoted

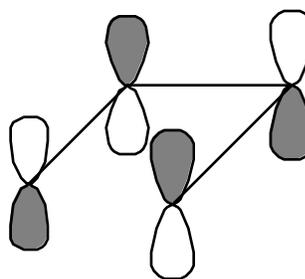
1



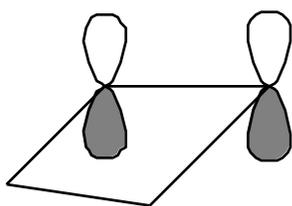
2



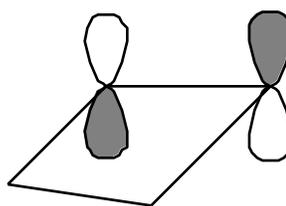
3



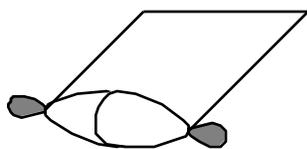
4



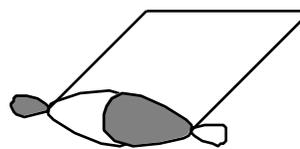
orbital of  
cyclobutene



\* orbital of  
cyclobutene

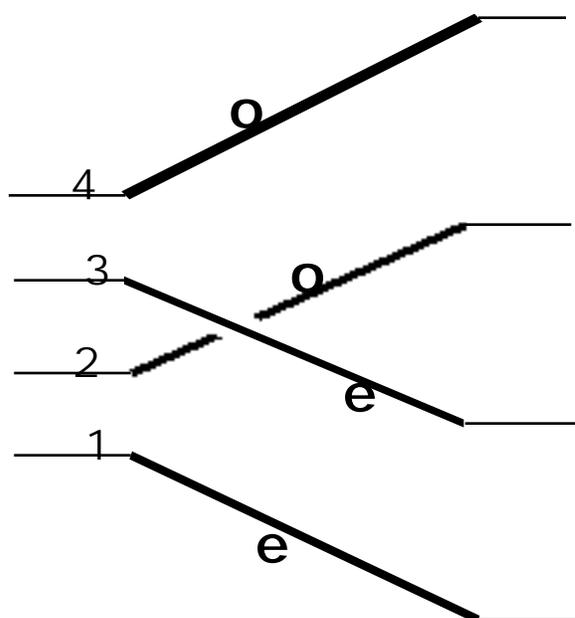


orbital of  
cyclobutene



\* orbital of  
cyclobutene

The four orbitals of 1,3-butadiene are of the following symmetries under the preserved plane (see the orbitals in the Figure above):  $\psi_1 = e$ ,  $\psi_2 = o$ ,  $\psi_3 = e$ ,  $\psi_4 = o$ . The  $\psi_1$  and  $\psi_3$  and  $\psi_2$  and  $\psi_4$  orbitals of cyclobutane which evolve from the four active orbitals of the 1,3-butadiene are of the following symmetry and energy order:  $\psi_1 = e$ ,  $\psi_2 = e$ ,  $\psi_3 = o$ ,  $\psi_4 = o$ . Connecting these orbitals by symmetry, starting with the lowest energy orbital and going through the highest energy orbital, gives the following OCD:



The fact that the lowest two orbitals of the reactants, which are those occupied by the four electrons of the reactant, do not correlate to the lowest two orbitals of the products, which are the orbitals occupied by the two and two electrons of the products, will be shown later in Chapter 12 to be the origin of the activation barrier for the thermal disrotatory rearrangement (in which the four active electrons occupy these lowest two orbitals) of 1,3-butadiene to produce cyclobutene.

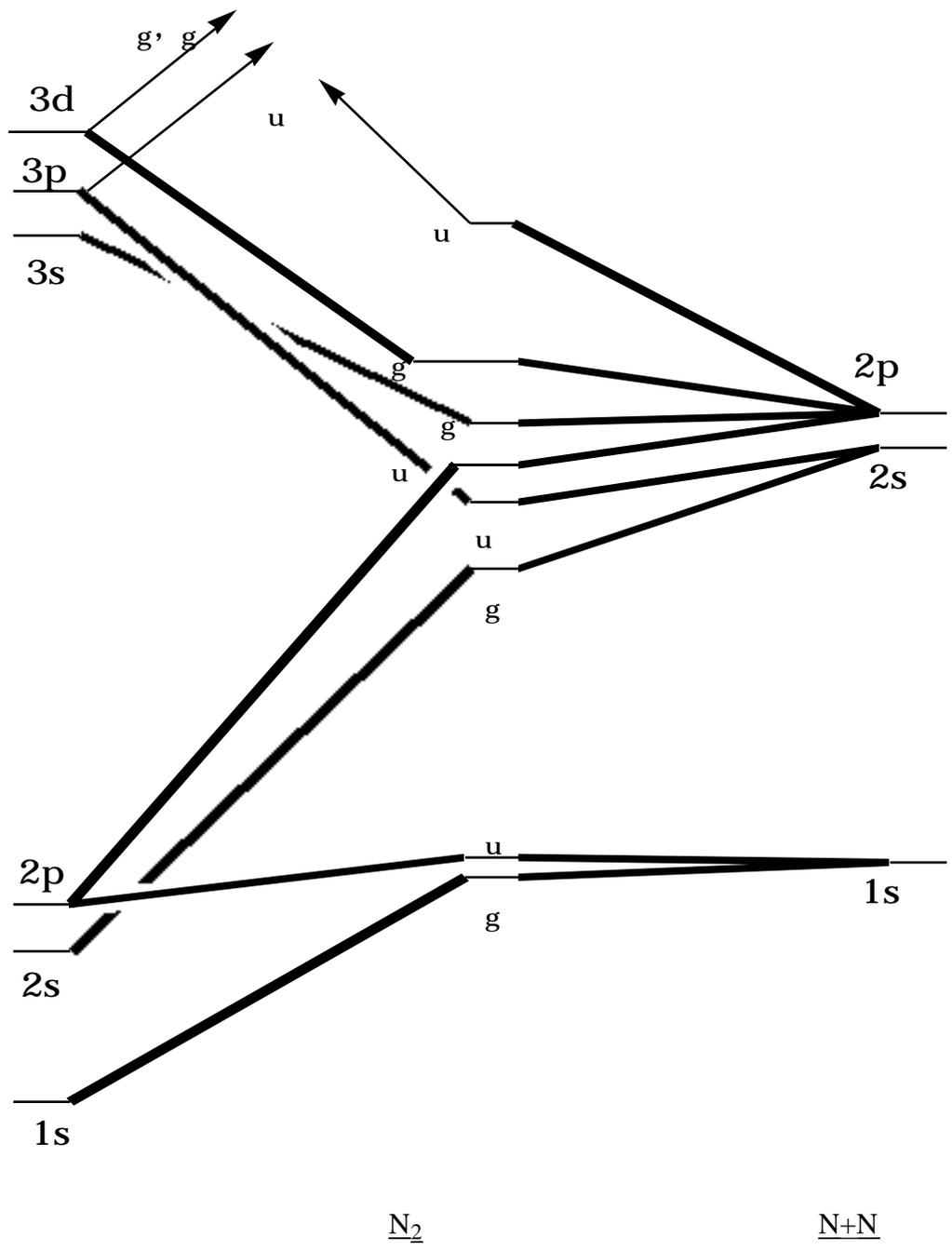
If the reactants could be prepared, for example by photolysis, in an excited state having orbital occupancy  $\psi_1^2 \psi_2^1 \psi_3^1$ , then reaction along the path considered would not have any symmetry-imposed barrier because this singly excited configuration correlates to a singly-excited configuration  $\psi_2^2 \psi_1^1 \psi_3^1$  of the products. The fact that the reactant and product configurations are of equivalent excitation level causes there to be no symmetry constraints

on the photochemically induced reaction of 1,3-butadiene to produce cyclobutene. In contrast, the thermal reaction considered first above has a symmetry-imposed barrier because the orbital occupancy is forced to rearrange (by the occupancy of two electrons) from the ground-state wavefunction of the reactant to smoothly evolve into that of the product.

It should be stressed that although these symmetry considerations may allow one to anticipate barriers on reaction potential energy surfaces, they have nothing to do with the thermodynamic energy differences of such reactions. Symmetry says whether there will be symmetry-imposed barriers above and beyond any thermodynamic energy differences. The enthalpies of formation of reactants and products contain the information about the reaction's overall energy balance.

As another example of an OCD, consider the  $N + N \Rightarrow N_2$  recombination reaction mentioned above. The orbitals of the atoms must first be labeled according to the axial rotation group (including the inversion operation because this is a homonuclear molecule). The core 1s orbitals are symmetry adapted to produce  $1_g$  and  $1_u$  orbitals (the number 1 is used to indicate that these are the lowest energy orbitals of their respective symmetries); the 2s orbitals generate  $2_g$  and  $2_u$  orbitals; the 2p orbitals combine to yield  $3_g$ , a pair of  $1_u$  orbitals, a pair of  $1_g$  orbitals, and the  $3_u$  orbital, whose bonding, nonbonding, and antibonding nature was detailed earlier. In the two separated Nitrogen atoms, the two orbitals derived from the 2s atomic orbitals are degenerate, and the six orbitals derived from the Nitrogen atoms' 2p orbitals are degenerate. At the equilibrium geometry of the  $N_2$  molecule, these degeneracies are lifted, Only the degeneracies of the  $1_u$  and  $1_g$  orbitals, which are dictated by the degeneracy of +m and -m orbitals within the axial rotation group, remain.

As one proceeds inward past the equilibrium bond length of  $N_2$ , toward the united-atom limit in which the two Nitrogen nuclei are fused to produce a Silicon nucleus, the energy ordering of the orbitals changes. Labeling the orbitals of the Silicon atom according to the axial rotation group, one finds the 1s is  $g$ , the 2s is  $g$ , the 2p orbitals are  $u$  and  $u$ , the 3s orbital is  $g$ , the 3p orbitals are  $u$  and  $u$ , and the 3d orbitals are  $g$ ,  $g$ , and  $g$ . The following OCD is obtained when one connects the orbitals of the two separated Nitrogen atoms (properly symmetry adapted) to those of the  $N_2$  molecule and eventually to those of the Silicon atom.



The fact that the separated-atom and united-atom limits involve several crossings in the OCD can be used to explain barriers in the potential energy curves of such diatomic molecules which occur at short internuclear distances. It should be noted that the Silicon atom's 3p orbitals of  $\sigma_u$  symmetry and its 3d orbitals of  $\sigma_g$  and  $\sigma_u$  symmetry correlate with higher energy orbitals of  $N_2$  not with the valence orbitals of this molecule, and that the  $3\sigma_u$  antibonding orbital of  $N_2$  correlates with a higher energy orbital of Silicon (in particular, its 4p orbital).

## Chapter 7

### *The Most Elementary Molecular Orbital Models Contain Symmetry, Nodal Pattern, and Approximate Energy Information*

#### I. The LCAO-MO Expansion and the Orbital-Level Schrödinger Equation

In the simplest picture of chemical bonding, the valence molecular orbitals  $\psi_i$  are constructed as linear combinations of valence atomic orbitals  $\phi_\mu$  according to the LCAO-MO formula:

$$\psi_i = \sum_{\mu} C_{i\mu} \phi_{\mu}$$

The core electrons are not explicitly included in such a treatment, although their effects are felt through an electrostatic potential  $V$  that has the following properties:

i.  $V$  contains contributions from all of the nuclei in the molecule exerting coulombic attractions on the electron, as well as coulombic repulsions and exchange interactions exerted by the other electrons on this electron;

ii. As a result of the (assumed) cancellation of attractions from distant nuclei and repulsions from the electron clouds (i.e., the core, lone-pair, and valence orbitals) that surround these distant nuclei, the effect of  $V$  on any particular molecular orbital  $\psi_i$  depends primarily on the atomic charges and local bond polarities of the atoms over which  $\psi_i$  is delocalized.

As a result of these assumptions, qualitative molecular orbital models can be developed in which one assumes that each molecular orbital  $\psi_i$  obeys a one-electron Schrödinger equation

$$\hat{h} \psi_i = \epsilon_i \psi_i$$

Here the orbital-level hamiltonian  $\hat{h}$  contains the kinetic energy of motion of the electron and the potential  $V$  mentioned above:

$$\left[ -\frac{\hbar^2}{2m_e} \nabla^2 + V \right] \psi_i = \epsilon_i \psi_i$$

Expanding the  $\psi_i$  in the LCAO-MO manner, substituting this expansion into the above Schrödinger equation, multiplying on the left by  $\psi_i^*$ , and integrating over the coordinates of the electron generates the following orbital-level eigenvalue problem:

$$\int \psi_i^* \left[ -\frac{\hbar^2}{2m_e} \nabla^2 + V \right] \sum_{\mu} C_{i\mu} \psi_{\mu} = \epsilon_i \int \psi_i^* \sum_{\mu} C_{i\mu} \psi_{\mu}.$$

If the constituent atomic orbitals  $\{\psi_{\mu}\}$  have been orthonormalized as discussed earlier in this chapter, the overlap integrals  $\int \psi_i^* \psi_{\mu}$  reduce to  $\delta_{i\mu}$ .

## II. Determining the Effective Potential $V$

In the most elementary models of orbital structure, the quantities that explicitly define the potential  $V$  are not computed from first principles as they are in so-called ab initio methods (see Section 6). Rather, either experimental data or results of ab initio calculations are used to determine the parameters in terms of which  $V$  is expressed. The resulting empirical or semi-empirical methods discussed below differ in the sophistication used to include electron-electron interactions as well as in the manner experimental data or ab initio computational results are used to specify  $V$ .

If experimental data is used to parameterize a semi-empirical model, then the model should not be extended beyond the level at which it has been parameterized. For example, experimental bond energies, excitation energies, and ionization energies may be used to determine molecular orbital energies which, in turn, are summed to compute total energies. In such a parameterization it would be incorrect to subsequently use these mos to form a wavefunction, as in Sections 3 and 6, that goes beyond the simple 'product of orbitals' description. To do so would be inconsistent because the more sophisticated wavefunction would duplicate what using the experimental data (which already contains mother nature's electronic correlations) to determine the parameters had accomplished.

Alternatively, if results of ab initio theory at the single-configuration orbital-product wavefunction level are used to define the parameters of a semi-empirical model, it would then be proper to use the semi-empirical orbitals in a subsequent higher-level treatment of electronic structure as done in Section 6.

### A. The Hückel Parameterization of $V$

In the most simplified embodiment of the above orbital-level model, the following additional approximations are introduced:

1. The diagonal values  $\langle \mu | -\hbar^2/2m_e \nabla^2 + V | \mu \rangle$ , which are usually denoted  $\epsilon_\mu$ , are taken to be equal to the energy of an electron in the atomic orbital  $\mu$  and, as such, are evaluated in terms of atomic ionization energies (IP's) and electron affinities (EA's):

$$\langle \mu | -\hbar^2/2m_e \nabla^2 + V | \mu \rangle = -IP_\mu,$$

for atomic orbitals that are occupied in the atom, and

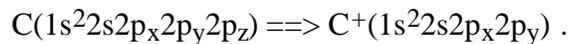
$$\langle \mu | -\hbar^2/2m_e \nabla^2 + V | \mu \rangle = -EA_\mu,$$

for atomic orbitals that are not occupied in the atom.

These approximations assume that contributions in  $V$  arising from coulombic attraction to nuclei other than the one on which  $\mu$  is located, and repulsions from the core, lone-pair, and valence electron clouds surrounding these other nuclei cancel to an extent that

$\langle \mu | V | \mu \rangle$  contains only potentials from the atom on which  $\mu$  sits.

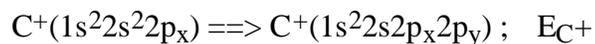
It should be noted that the IP's and EA's of valence-state orbitals are not identical to the experimentally measured IP's and EA's of the corresponding atom, but can be obtained from such information. For example, the 2p valence-state IP (VSIP) for a Carbon atom is the energy difference associated with the hypothetical process



If the energy differences for the "promotion" of C



and for the promotion of  $C^+$



are known, the desired VSIP is given by:

$$IP_{2p_z} = IP_C + E_{C^+} - E_C .$$

The EA of the 2p orbital is obtained from the

$$C(1s^2 2s^2 2p_x 2p_y) \implies C^-(1s^2 2s^2 2p_x 2p_y 2p_z)$$

energy gap, which means that  $EA_{2p_z} = EA_C$ . Some common IP's of valence 2p orbitals in eV are as follows: C (11.16), N (14.12), N<sup>+</sup> (28.71), O (17.70), O<sup>+</sup> (31.42), F<sup>+</sup> (37.28).

2. The off-diagonal elements  $\langle \mu | -\hbar^2/2m_e \nabla^2 + V | \nu \rangle$  are taken as zero if  $\mu$  and  $\nu$  belong to the same atom because the atomic orbitals are assumed to have been constructed to diagonalize the one-electron hamiltonian appropriate to an electron moving in that atom. They are set equal to a parameter denoted  $\mu_{\nu}$ , if  $\mu$  and  $\nu$  reside on neighboring atoms that are chemically bonded. If  $\mu$  and  $\nu$  reside on atoms that are not bonded neighbors, then the off-diagonal matrix element is set equal to zero.

3. The geometry dependence of the  $\mu_{\nu}$  parameters is often approximated by assuming that  $\mu_{\nu}$  is proportional to the overlap  $S_{\mu\nu}$  between the corresponding atomic orbitals:

$$\mu_{\nu} = \mu_0 S_{\mu\nu}.$$

Here  $\mu_0$  is a constant (having energy units) characteristic of the bonding interaction between  $\mu$  and  $\nu$ ; its value is usually determined by forcing the molecular orbital energies obtained from such a qualitative orbital treatment to yield experimentally correct ionization potentials, bond dissociation energies, or electronic transition energies.

The particular approach described thus far forms the basis of the so-called Hückel model. Its implementation requires knowledge of the atomic  $\mu$  and  $\mu_0$  values, which are eventually expressed in terms of experimental data, as well as a means of calculating the geometry dependence of the  $\mu_{\nu}$ 's (e.g., some method for computing overlap matrices  $S_{\mu\nu}$ ).

## B. The Extended Hückel Method

It is well known that bonding and antibonding orbitals are formed when a pair of atomic orbitals from neighboring atoms interact. The energy splitting between the bonding

and antibonding orbitals depends on the overlap between the pair of atomic orbitals. Also, the energy of the antibonding orbital lies higher above the arithmetic mean  $E_{ave} = E_A + E_B$  of the energies of the constituent atomic orbitals ( $E_A$  and  $E_B$ ) than the bonding orbital lies below  $E_{ave}$ . If overlap is ignored, as in conventional Hückel theory (except in parameterizing the geometry dependence of  $S_{\mu, \nu}$ ), the differential destabilization of antibonding orbitals compared to stabilization of bonding orbitals can not be accounted for.

By parameterizing the off-diagonal Hamiltonian matrix elements in the following overlap-dependent manner:

$$h_{\nu, \mu} = \langle \nu | -\hbar^2/2m_e \nabla^2 + V | \mu \rangle = 0.5 K (h_{\mu, \mu} + h_{\nu, \nu}) S_{\mu, \nu},$$

and explicitly treating the overlaps among the constituent atomic orbitals  $\{ \mu \}$  in solving the orbital-level Schrödinger equation

$$\mu \langle \nu | -\hbar^2/2m_e \nabla^2 + V | \mu \rangle C_{i\mu} = \epsilon_i \mu \langle \nu | \mu \rangle C_{i\mu},$$

Hoffmann introduced the so-called extended Hückel method. He found that a value for  $K=1.75$  gave optimal results when using Slater-type orbitals as a basis (and for calculating the  $S_{\mu, \nu}$ ). The diagonal  $h_{\mu, \mu}$  elements are given, as in the conventional Hückel method, in terms of valence-state IP's and EA's. Cusachs later proposed a variant of this parameterization of the off-diagonal elements:

$$h_{\nu, \mu} = 0.5 K (h_{\mu, \mu} + h_{\nu, \nu}) S_{\mu, \nu} (2 - |S_{\mu, \nu}|).$$

For first- and second-row atoms, the 1s or (2s, 2p) or (3s, 3p, 3d) valence-state ionization energies ( $\mu$ 's), the number of valence electrons (#Elec.) as well as the orbital exponents ( $e_s$ ,  $e_p$  and  $e_d$ ) of Slater-type orbitals used to calculate the overlap matrix elements  $S_{\mu, \nu}$ , corresponding are given below.

Atom	# Elec.	$e_s=e_p$	$e_d$	s(eV)	p(eV)	d(eV)
H	1	1.3		-13.6		
Li	1	0.650		-5.4	-3.5	
Be	2	0.975		-10.0	-6.0	
B	3	1.300		-15.2	-8.5	
C	4	1.625		-21.4	-11.4	
N	5	1.950		-26.0	-13.4	
O	6	2.275		-32.3	-14.8	
F	7	2.425		-40.0	-18.1	
Na	1	0.733		-5.1	-3.0	
Mg	2	0.950		-9.0	-4.5	
Al	3	1.167		-12.3	-6.5	
Si	4	1.383	1.383	-17.3	-9.2	-6.0
P	5	1.600	1.400	-18.6	-14.0	-7.0
S	6	1.817	1.500	-20.0	-13.3	-8.0
Cl	7	2.033	2.033	-30.0	-15.0	-9.0

In the Hückel or extended Hückel methods no explicit reference is made to electron-electron interactions although such contributions are absorbed into the V potential, and hence into the  $\mu$  and  $\mu$ , parameters of Hückel theory or the  $h_{\mu,\mu}$  and  $h_{\mu}$  parameters of extended Hückel theory. As electron density flows from one atom to another (due to electronegativity differences), the electron-electron repulsions in various atomic orbitals changes. To account for such charge-density-dependent coulombic energies, one must use an approach that includes explicit reference to inter-orbital coulomb and exchange interactions. There exists a large family of semi-empirical methods that permit explicit treatment of electronic interactions; some of the more commonly used approaches are discussed in Appendix F.

## Section 3 Electronic Configurations, Term Symbols, and States

**Introductory Remarks-** The Orbital, Configuration, and State Pictures of Electronic Structure

*One of the goals of quantum chemistry is to allow practicing chemists to use knowledge of the electronic states of fragments (atoms, radicals, ions, or molecules) to predict and understand the behavior (i.e., electronic energy levels, geometries, and reactivities) of larger molecules. In the preceding Section, orbital correlation diagrams were introduced to connect the orbitals of the fragments along a 'reaction path' leading to the orbitals of the products. In this Section, analogous connections are made among the fragment and product electronic states, again labeled by appropriate symmetries. To realize such connections, one must first write down  $N$ -electron wavefunctions that possess the appropriate symmetry; this task requires combining symmetries of the occupied orbitals to obtain the symmetries of the resulting states.*

### Chapter 8

*Electrons are Placed into Orbitals to Form Configurations, Each of Which Can be Labeled by its Symmetry. The Configurations May "Interact" Strongly if They Have Similar Energies.*

#### I. Orbitals Do Not Provide the Complete Picture; Their Occupancy By the $N$ Electrons Must Be Specified

Knowing the orbitals of a particular species provides one information about the sizes, shapes, directions, symmetries, and energies of those regions of space that are available to the electrons (i.e., the complete set of orbitals that are available). This knowledge does not determine into which orbitals the electrons are placed. It is by describing the electronic configurations (i.e., orbital occupancies such as  $1s^2 2s^2 2p^2$  or  $1s^2 2s^2 2p^1 3s^1$ ) appropriate to the energy range under study that one focuses on how the electrons occupy the orbitals. Moreover, a given configuration may give rise to several energy levels whose energies differ by chemically important amounts. For example, the  $1s^2 2s^2 2p^2$  configuration of the Carbon atom produces nine degenerate  $^3P$  states, five degenerate  $^1D$  states, and a single  $^1S$  state. These three energy levels differ in energy by 1.5 eV and 1.2 eV, respectively.

## II. Even N-Electron Configurations Are Not Mother Nature's True Energy States

Moreover, even single-configuration descriptions of atomic and molecular structure (e.g.,  $1s^2 2s^2 2p^4$  for the Oxygen atom) do not provide fully correct or highly accurate representations of the respective electronic wavefunctions. As will be shown in this Section and in more detail in Section 6, the picture of N electrons occupying orbitals to form a configuration is based on a so-called "mean field" description of the coulomb interactions among electrons. In such models, an electron at  $\mathbf{r}$  is viewed as interacting with an "averaged" charge density arising from the N-1 remaining electrons:

$$V_{\text{mean field}} = \int \rho_{N-1}(\mathbf{r}') \frac{e^2}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' .$$

Here  $\rho_{N-1}(\mathbf{r}')$  represents the probability density for finding electrons at  $\mathbf{r}'$ , and  $e^2/|\mathbf{r}-\mathbf{r}'|$  is the mutual coulomb repulsion between electron density at  $\mathbf{r}$  and  $\mathbf{r}'$ . Analogous mean-field models arise in many areas of chemistry and physics, including electrolyte theory (e.g., the Debye-Hückel theory), statistical mechanics of dense gases (e.g., where the Mayer-Mayer cluster expansion is used to improve the ideal-gas mean field model), and chemical dynamics (e.g., the vibrationally averaged potential of interaction).

In each case, the mean-field model forms only a starting point from which one attempts to build a fully correct theory by effecting systematic corrections (e.g., using perturbation theory) to the mean-field model. The ultimate value of any particular mean-field model is related to its accuracy in describing experimental phenomena. If predictions of the mean-field model are far from the experimental observations, then higher-order corrections (which are usually difficult to implement) must be employed to improve its predictions. In such a case, one is motivated to search for a better model to use as a starting point so that lower-order perturbative (or other) corrections can be used to achieve chemical accuracy (e.g.,  $\pm 1$  kcal/mole).

In electronic structure theory, the single-configuration picture (e.g., the  $1s^2 2s^2 2p^4$  description of the Oxygen atom) forms the mean-field starting point; the configuration interaction (CI) or perturbation theory techniques are then used to systematically improve this level of description.

The single-configuration mean-field theories of electronic structure neglect correlations among the electrons. That is, in expressing the interaction of an electron at  $\mathbf{r}$

with the N-1 other electrons, they use a probability density  $\rho_{N-1}(\mathbf{r}')$  that is independent of the fact that another electron resides at  $\mathbf{r}$ . In fact, the so-called conditional probability density for finding one of N-1 electrons at  $\mathbf{r}'$ , given that an electron is at  $\mathbf{r}$  certainly depends on  $\mathbf{r}$ . As a result, the mean-field coulomb potential felt by a  $2p_x$  orbital's electron in the  $1s^2 2s^2 2p_x 2p_y$  single-configuration description of the Carbon atom is:

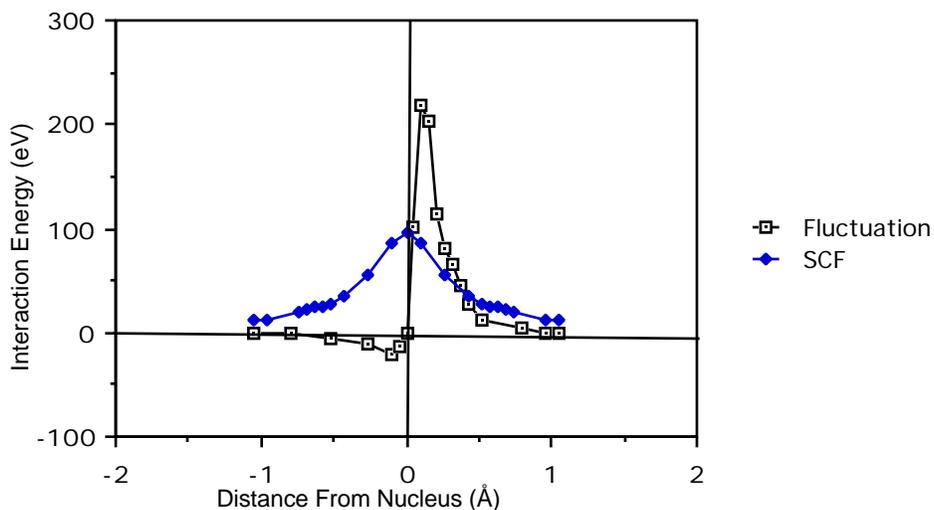
$$\begin{aligned}
 V_{\text{mean field}} = & 2 \int |1s(\mathbf{r}')|^2 \frac{e^2}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' \\
 & + 2 \int |2s(\mathbf{r}')|^2 \frac{e^2}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' \\
 & + \int |2p_y(\mathbf{r}')|^2 \frac{e^2}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}' .
 \end{aligned}$$

In this example, the density  $\rho_{N-1}(\mathbf{r}')$  is the sum of the charge densities of the orbitals occupied by the five other electrons  $2 |1s(\mathbf{r}')|^2 + 2 |2s(\mathbf{r}')|^2 + |2p_y(\mathbf{r}')|^2$ , and is not dependent on the fact that an electron resides at  $\mathbf{r}$ .

### III. Mean-Field Models

*The Mean-Field Model, Which Forms the Basis of Chemists' Pictures of Electronic Structure of Molecules, Is Not Very Accurate*

The magnitude and "shape" of such a mean-field potential is shown below for the Beryllium atom. In this figure, the nucleus is at the origin, and one electron is placed at a distance from the nucleus equal to the maximum of the 1s orbital's radial probability density (near 0.13 Å). The radial coordinate of the second is plotted as the abscissa; this second electron is arbitrarily constrained to lie on the line connecting the nucleus and the first electron (along this direction, the inter-electronic interactions are largest). On the ordinate, there are two quantities plotted: (i) the Self-Consistent Field (SCF) mean-field potential  $\int |1s(\mathbf{r}')|^2 \frac{e^2}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}'$ , and (ii) the so-called Fluctuation potential (F), which is the true coulombic  $\frac{e^2}{|\mathbf{r}-\mathbf{r}'|}$  interaction potential minus the SCF potential.



As a function of the inter-electron distance, the fluctuation potential decays to zero more rapidly than does the SCF potential. For this reason, approaches in which  $F$  is treated as a perturbation and corrections to the mean-field picture are computed perturbatively might be expected to be rapidly convergent (whenever perturbations describing long-range interactions arise, convergence of perturbation theory is expected to be slow or not successful). However, the magnitude of  $F$  is quite large and remains so over an appreciable range of inter-electron distances.

The resultant corrections to the SCF picture are therefore quite large when measured in kcal/mole. For example, the differences  $E$  between the true (state-of-the-art quantum chemical calculation) energies of interaction among the four electrons in Be and the SCF mean-field estimates of these interactions are given in the table shown below in eV (recall that  $1 \text{ eV} = 23.06 \text{ kcal/mole}$ ).

Orb. Pair	1s 1s	1s 2s	1s 2s	1s 2s	1s 2s	2s 2s
E in eV	1.126	0.022	0.058	0.058	0.022	1.234

To provide further insight why the SCF mean-field model in electronic structure theory is of limited accuracy, it can be noted that the average value of the kinetic energy plus the attraction to the Be nucleus plus the SCF interaction potential for one of the 2s orbitals of Be with the three remaining electrons in the  $1s^2 2s^2$  configuration is:

$$\langle 2s | -\hbar^2/2m_e \nabla^2 - 4e^2/r + V_{\text{SCF}} | 2s \rangle = -15.4 \text{ eV};$$

the analogous quantity for the 2p orbital in the  $1s^2 2s 2p$  configuration is:

$$\langle 2p | -\hbar^2/2m_e \nabla^2 - 4e^2/r + V'_{\text{SCF}} | 2p \rangle = -12.28 \text{ eV};$$

the corresponding value for the 1s orbital is (negative and) of even larger magnitude. The SCF average coulomb interaction between the two 2s orbitals of  $1s^2 2s^2$  Be is:

$$\int |2s(\mathbf{r})|^2 |2s(\mathbf{r}')|^2 e^2/|\mathbf{r}-\mathbf{r}'| \mathbf{d}\mathbf{r} \mathbf{d}\mathbf{r}' = 5.95 \text{ eV}.$$

This data clearly shows that corrections to the SCF model (see the above table) represent significant fractions of the inter-electron interaction energies (e.g., 1.234 eV compared to  $5.95 - 1.234 = 4.72$  eV for the two 2s electrons of Be), and that the inter-electron interaction energies, in turn, constitute significant fractions of the total energy of each orbital (e.g.,  $5.95 - 1.234$  eV = 4.72 eV out of -15.4 eV for a 2s orbital of Be).

The task of describing the electronic states of atoms and molecules from first principles and in a chemically accurate manner ( $\pm 1$  kcal/mole) is clearly quite formidable. The orbital picture and its accompanying SCF potential take care of "most" of the interactions among the N electrons (which interact via long-range coulomb forces and whose dynamics requires the application of quantum physics and permutational symmetry). However, the residual fluctuation potential, although of shorter range than the bare coulomb potential, is large enough to cause significant corrections to the mean-field picture. This, in turn, necessitates the use of more sophisticated and computationally taxing techniques (e.g., high order perturbation theory or large variational expansion spaces) to reach the desired chemical accuracy.

Mean-field models are obviously approximations whose accuracy must be determined so scientists can know to what degree they can be "trusted". For electronic structures of atoms and molecules, they require quite substantial corrections to bring them into line with experimental fact. Electrons in atoms and molecules undergo dynamical motions in which their coulomb repulsions cause them to "avoid" one another at every instant of time, not only in the average-repulsion manner that the mean-field models embody. The inclusion of instantaneous spatial correlations among electrons is necessary to achieve a more accurate description of atomic and molecular electronic structure.

#### IV. Configuration Interaction (CI) Describes the Correct Electronic States

The most commonly employed tool for introducing such spatial correlations into electronic wavefunctions is called configuration interaction (CI); this approach is described briefly later in this Section and in considerable detail in **Section 6**.

Briefly, one employs the (in principle, complete as shown by P. O. Löwdin, Rev. Mod. Phys. 32, 328 (1960)) set of N-electron configurations that (i) can be formed by placing the N electrons into orbitals of the atom or molecule under study, and that (ii) possess the spatial, spin, and angular momentum symmetry of the electronic state of interest. This set of functions is then used, in a linear variational function, to achieve, via the CI technique, a more accurate and dynamically correct description of the electronic structure of that state. For example, to describe the ground  $^1S$  state of the Be atom, the  $1s^2 2s^2$  configuration (which yields the mean-field description) is augmented by including other configurations such as  $1s^2 3s^2$ ,  $1s^2 2p^2$ ,  $1s^2 3p^2$ ,  $1s^2 2s 3s$ ,  $3s^2 2s^2$ ,  $2p^2 2s^2$ , etc., all of which have overall  $^1S$  spin and angular momentum symmetry. The excited  $^1S$  states are also combinations of all such configurations. Of course, the ground-state wavefunction is dominated by the  $|1s^2 2s^2\rangle$  and excited states contain dominant contributions from  $|1s^2 2s 3s\rangle$ , etc. configurations. The resultant CI wavefunctions are formed as shown in **Section 6** as linear combinations of all such configurations.

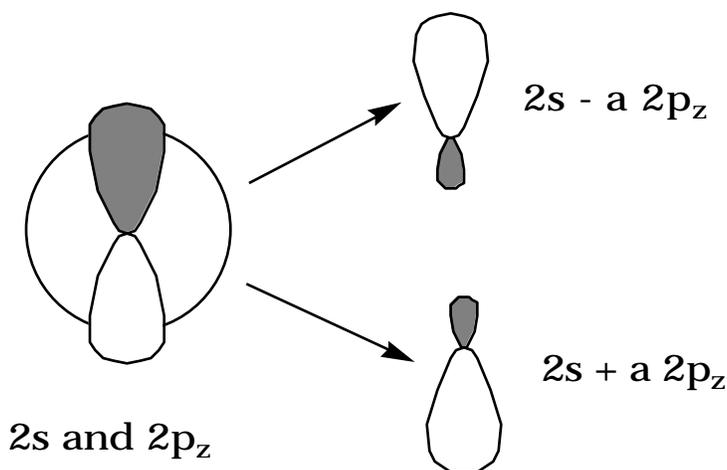
To clarify the physical significance of mixing such configurations, it is useful to consider what are found to be the two most important such configurations for the ground  $^1S$  state of the Be atom:

$$C_1 |1s^2 2s^2\rangle - C_2 [|1s^2 2p_x^2\rangle + |1s^2 2p_y^2\rangle + |1s^2 2p_z^2\rangle].$$

As proven in Chapter 13.III, this two-configuration description of Be's electronic structure is equivalent to a description in which two electrons reside in the 1s orbital (with opposite, and spins) while the other pair reside in 2s-2p hybrid orbitals (more correctly, polarized orbitals) in a manner that instantaneously correlates their motions:

$$\begin{aligned} & 1/6 C_1 |1s^2\{[(2s-a2p_x) (2s+a2p_x) - (2s-a2p_x) (2s+a2p_x) ] \\ & \quad + [(2s-a2p_y) (2s+a2p_y) - (2s-a2p_y) (2s+a2p_y) ] \\ & \quad + [(2s-a2p_z) (2s+a2p_z) - (2s-a2p_z) (2s+a2p_z) ]\}|, \end{aligned}$$

where  $a = \sqrt{3C_2/C_1}$ . The so-called polarized orbital pairs ( $2s \pm a 2p_{x,y, \text{ or } z}$ ) are formed by mixing into the 2s orbital an amount of the  $2p_{x,y, \text{ or } z}$  orbital, with the mixing amplitude determined by the ratio of  $C_2$  to  $C_1$ . As will be detailed in Section 6, this ratio is proportional to the magnitude of the coupling  $\langle 1s^2 2s^2 | H | 1s^2 2p^2 \rangle$  between the two configurations and inversely proportional to the energy difference  $[\langle 1s^2 2s^2 | H | 1s^2 2s^2 \rangle - \langle 1s^2 2p^2 | H | 1s^2 2p^2 \rangle]$  for these configurations. So, in general, configurations that have similar energies (Hamiltonian expectation values) and couple strongly give rise to strongly mixed polarized orbital pairs. The result of forming such polarized orbital pairs are described pictorially below.



## Polarized Orbital $2s$ and $2p_z$ Pairs

In each of the three equivalent terms in this wavefunction, one of the valence electrons moves in a  $2s+a2p$  orbital polarized in one direction while the other valence electron moves in the  $2s-a2p$  orbital polarized in the opposite direction. For example, the first term  $[(2s-a2p_x)(2s+a2p_x) - (2s-a2p_x)(2s+a2p_x)]$  describes one electron occupying a  $2s-a2p_x$  polarized orbital while the other electron occupies the  $2s+a2p_x$  orbital. In this picture, the electrons reduce their mutual coulomb repulsion by occupying different regions of space; in the SCF mean-field picture, both electrons reside in the same  $2s$  region of space. In this particular example, the electrons undergo angular correlation to "avoid" one another. The fact that equal amounts of  $x$ ,  $y$ , and  $z$  orbital polarization appear in  $^1S$  is what preserves the  $^1S$  symmetry of the wavefunction.

The fact that the CI wavefunction

$$C_1 |1s^2 2s^2| - C_2 [|1s^2 2p_x^2| + |1s^2 2p_y^2| + |1s^2 2p_z^2|]$$

mixes its two configurations with opposite sign is of significance. As will be seen later in Section 6, solution of the Schrödinger equation using the CI method in which two configurations (e.g.,  $|1s^2 2s^2|$  and  $|1s^2 2p^2|$ ) are employed gives rise to two solutions. One approximates the ground state wave function; the other approximates an excited state. The former is the one that mixes the two configurations with opposite sign.

To understand why the latter is of higher energy, it suffices to analyze a function of the form

$$C_1 |1s^2 2s^2| + C_2 [|1s^2 2p_x^2| + |1s^2 2p_y^2| + |1s^2 2p_z^2|]$$

in a manner analogous to above. In this case, it can be shown that

$$\begin{aligned} & \frac{1}{6} C_1 |1s^2 \{ [(2s - ia2p_x) (2s + ia2p_x) - (2s - ia2p_x) (2s + ia2p_x) ] \\ & \quad + [(2s - ia2p_y) (2s + ia2p_y) - (2s - ia2p_y) (2s + ia2p_y) ] \\ & \quad + [(2s - ia2p_z) (2s + ia2p_z) - (2s - ia2p_z) (2s + ia2p_z) ] \} | \end{aligned}$$

There is a fundamental difference, however, between the polarized orbital pairs introduced earlier  $\pm = (2s \pm a2p_{x,y,or z})$  and the corresponding functions  $\pm = (2s \pm ia2p_{x,y,or z})$  appearing here. The probability densities embodied in the former

$$|\pm|^2 = |2s|^2 + a^2 |2p_{x,y,or z}|^2 \pm 2a(2s \ 2p_{x,y,or z})$$

describe constructive (for the + case) and destructive (for the - case) superposition of the probabilities of the 2s and 2p orbitals. The probability densities of  $\pm$  are

$$\begin{aligned} |\pm|^2 &= (2s \pm ia2p_{x,y,or z})^* (2s \pm ia2p_{x,y,or z}) \\ &= |2s|^2 + a^2 |2p_{x,y,or z}|^2. \end{aligned}$$

These densities are identical to one another and do not describe polarized orbital densities. Therefore, the CI wavefunction which mixes the two configurations with like sign, when analyzed in terms of orbital pairs, places the electrons into orbitals  $\psi_{\pm} = (2s \pm ia2p_{x,y,or z})$  whose densities do not permit the electrons to avoid one another. Rather, both orbitals have the same spatial density  $|2s|^2 + a^2 |2p_{x,y,or z}|^2$ , which gives rise to higher coulombic interaction energy for this state.

## V. Summary

In summary, the dynamical interactions among electrons give rise to instantaneous spatial correlations that must be handled to arrive at an accurate picture of atomic and molecular structure. The simple, single-configuration picture provided by the mean-field model is a useful starting point, but improvements are often needed.

In Section 6, methods for treating electron correlation will be discussed in greater detail.

For the remainder of this Section, the primary focus is placed on forming proper N-electron wavefunctions by occupying the orbitals available to the system in a manner that guarantees that the resultant N-electron function is an eigenfunction of those operators that commute with the N-electron Hamiltonian.

For polyatomic molecules, these operators include point-group symmetry operators (which act on all N electrons) and the spin angular momentum ( $S^2$  and  $S_z$ ) of all of the electrons taken as a whole (this is true in the absence of spin-orbit coupling which is treated later as a perturbation). For linear molecules, the point group symmetry operations involve rotations  $R_z$  of all N electrons about the principal axis, as a result of which the total angular momentum  $L_z$  of the N electrons (taken as a whole) about this axis commutes with the Hamiltonian, H. Rotation of all N electrons about the x and y axes does not leave the total coulombic potential energy unchanged, so  $L_x$  and  $L_y$  do not commute with H. Hence for a linear molecule,  $L_z$ ,  $S^2$ , and  $S_z$  are the operators that commute with H. For atoms, the corresponding operators are  $L^2$ ,  $L_z$ ,  $S^2$ , and  $S_z$  (again, in the absence of spin-orbit coupling) where each operator pertains to the total orbital or spin angular momentum of the N electrons.

To construct N-electron functions that are eigenfunctions of the spatial symmetry or orbital angular momentum operators as well as the spin angular momentum operators, one has to "couple" the symmetry or angular momentum properties of the individual spin-orbitals used to construct the N-electrons functions. This coupling involves forming direct product symmetries in the case of polyatomic molecules that belong to finite point groups,

it involves vector coupling orbital and spin angular momenta in the case of atoms, and it involves vector coupling spin angular momenta and axis coupling orbital angular momenta when treating linear molecules. Much of this Section is devoted to developing the tools needed to carry out these couplings.

## Chapter 9

### *Electronic Wavefunctions Must be Constructed to Have Permutational Antisymmetry Because the $N$ Electrons are Indistinguishable Fermions*

#### I. Electronic Configurations

Atoms, linear molecules, and non-linear molecules have orbitals which can be labeled either according to the symmetry appropriate for that isolated species or for the species in an environment which produces lower symmetry. These orbitals should be viewed as regions of space in which electrons can move, with, of course, at most two electrons (of opposite spin) in each orbital. Specification of a particular occupancy of the set of orbitals available to the system gives an electronic configuration. For example,  $1s^2 2s^2 2p^4$  is an electronic configuration for the Oxygen atom (and for the  $F^{+1}$  ion and the  $N^{-1}$  ion);  $1s^2 2s^2 2p^3 3p^1$  is another configuration for O,  $F^{+1}$ , or  $N^{-1}$ . These configurations represent situations in which the electrons occupy low-energy orbitals of the system and, as such, are likely to contribute strongly to the true ground and low-lying excited states and to the low-energy states of molecules formed from these atoms or ions.

Specification of an electronic configuration does not, however, specify a particular electronic state of the system. In the above  $1s^2 2s^2 2p^4$  example, there are many ways (fifteen, to be precise) in which the 2p orbitals can be occupied by the four electrons. As a result, there are a total of fifteen states which cluster into three energetically distinct levels, lying within this single configuration. The  $1s^2 2s^2 2p^3 3p^1$  configuration contains thirty-six states which group into six distinct energy levels (the word level is used to denote one or more state with the same energy). Not all states which arise from a given electronic configuration have the same energy because various states occupy the degenerate (e.g., 2p and 3p in the above examples) orbitals differently. That is, some states have orbital occupancies of the form  $2p^2_1 2p^1_0 2p^1_{-1}$  while others have  $2p^2_1 2p^2_0 2p^0_{-1}$ ; as a result, the states can have quite different coulombic repulsions among the electrons (the state with two doubly occupied orbitals would lie higher in energy than that with two singly occupied orbitals). Later in this Section and in Appendix G techniques for constructing wavefunctions for each state contained within a particular configuration are given in detail. Mastering these tools is an important aspect of learning the material in this text.

In summary, an atom or molecule has many orbitals (core, bonding, non-bonding, Rydberg, and antibonding) available to it; occupancy of these orbitals in a particular manner gives rise to a configuration. If some orbitals are partially occupied in this configuration,

more than one state will arise; these states can differ in energy due to differences in how the orbitals are occupied. In particular, if degenerate orbitals are partially occupied, many states can arise and have energies which differ substantially because of differences in electron repulsions arising in these states. Systematic procedures for extracting all states from a given configuration, for labeling the states according to the appropriate symmetry group, for writing the wavefunctions corresponding to each state and for evaluating the energies corresponding to these wavefunctions are needed. Much of Chapters 10 and 11 are devoted to developing and illustrating these tools.

## II. Antisymmetric Wavefunctions

### A. General Concepts

The total electronic Hamiltonian

$$H = \sum_i \left( -\frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_a Z_a \frac{e^2}{r_{ia}} \right) + \sum_{i>j} \frac{e^2}{r_{ij}} + \sum_{a>b} Z_a Z_b \frac{e^2}{r_{ab}},$$

where  $i$  and  $j$  label electrons and  $a$  and  $b$  label the nuclei (whose charges are denoted  $Z_a$ ), commutes with the operators  $P_{ij}$  which permute the names of the electrons  $i$  and  $j$ . This, in turn, requires eigenfunctions of  $H$  to be eigenfunctions of  $P_{ij}$ . In fact, the set of such permutation operators form a group called the symmetric group (a good reference to this subject is contained in Chapter 7 of Group Theory, M. Hamermesh, Addison-Wesley, Reading, Mass. (1962)). In the present text, we will not exploit the full group theoretical nature of these operators; we will focus on the simple fact that all wavefunctions must be eigenfunctions of the  $P_{ij}$  (additional material on this subject is contained in Chapter XIV of Kemble).

Because  $P_{ij}$  obeys  $P_{ij} * P_{ij} = 1$ , the eigenvalues of the  $P_{ij}$  operators must be  $+1$  or  $-1$ . Electrons are Fermions (i.e., they have half-integral spin), and they have wavefunctions which are odd under permutation of any pair:  $P_{ij} \psi = -\psi$ . Bosons such as photons or deuterium nuclei (i.e., species with integral spin quantum numbers) have wavefunctions which obey  $P_{ij} \psi = +\psi$ .

These permutational symmetries are not only characteristics of the exact eigenfunctions of  $H$  belonging to any atom or molecule containing more than a single electron but they are also conditions which must be placed on any acceptable model or trial wavefunction (e.g., in a variational sense) which one constructs.

In particular, within the orbital model of electronic structure (which is developed more systematically in Section 6), one can not construct trial wavefunctions which are simple spin-orbital products (i.e., an orbital multiplied by an  $\alpha$  or  $\beta$  spin function for each electron) such as  $1s \alpha 1s \beta 2s \alpha 2s \beta p_1 \alpha p_0 \beta$ . Such spin-orbital product functions must be made permutationally antisymmetric if the N-electron trial function is to be properly antisymmetric. This can be accomplished for any such product wavefunction by applying the following antisymmetrizer operator:

$$A = (1/N!) \sum_P s_P P,$$

where N is the number of electrons, P runs over all N! permutations, and  $s_P$  is +1 or -1 depending on whether the permutation P contains an even or odd number of pairwise permutations (e.g., 231 can be reached from 123 by two pairwise permutations- $123 \Rightarrow 213 \Rightarrow 231$ , so 231 would have  $s_P = 1$ ). The permutation operator P in A acts on a product wavefunction and permutes the ordering of the spin-orbitals. For example,  $A \psi_{123} = (1/6) [\psi_{123} - \psi_{132} - \psi_{213} + \psi_{231} + \psi_{312} - \psi_{321}]$ , where the convention is that electronic coordinates  $r_1$ ,  $r_2$ , and  $r_3$  correspond to the orbitals as they appear in the product (e.g., the term  $\psi_{321}$  represents  $\psi_{\alpha}(r_1) \psi_{\beta}(r_2) \psi_{\alpha}(r_3)$ ).

It turns out that the permutations P can be allowed either to act on the "names" or labels of the electrons, keeping the order of the spin-orbitals fixed, or to act on the spin-orbitals, keeping the order and identity of the electrons' labels fixed. The resultant wavefunction, which contains N! terms, is exactly the same regardless of how one allows the permutations to act. Because we wish to use the above convention in which the order of the electronic labels remains fixed as 1, 2, 3, ... N, we choose to think of the permutations acting on the names of the spin-orbitals.

It should be noted that the effect of A on any spin-orbital product is to produce a function that is a sum of N! terms. In each of these terms the same spin-orbitals appear, but the order in which they appear differs from term to term. Thus antisymmetrization does not alter the overall orbital occupancy; it simply "scrambles" any knowledge of which electron is in which spin-orbital.

The antisymmetrized orbital product  $A \psi_{123}$  is represented by the short hand  $|\psi_{123}\rangle$  and is referred to as a Slater determinant. The origin of this notation can be made clear by noting that  $(1/N!)$  times the determinant of a matrix whose rows are labeled by the index i of the spin-orbital  $\psi_i$  and whose columns are labeled by the index j of the electron at  $\mathbf{r}_j$  is equal to the above function:  $A \psi_{123} = (1/3!) \det(\psi_i(\mathbf{r}_j))$ . The general structure of such Slater determinants is illustrated below:

$$(1/N!)^{1/2} \det\{ \psi_j(\mathbf{r}_i) \} = (1/N!)^{1/2} \begin{bmatrix} \psi_1(1) & \psi_2(1) & \psi_3(1) & \dots & \psi_k(1) & \dots & \psi_N(1) \\ \psi_1(2) & \psi_2(2) & \psi_3(2) & \dots & \psi_k(2) & \dots & \psi_N(2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \psi_1(i) & \psi_2(i) & \psi_3(i) & \dots & \psi_k(i) & \dots & \psi_N(i) \end{bmatrix}$$

The antisymmetry of many-electron spin-orbital products places constraints on any acceptable model wavefunction, which give rise to important physical consequences. For example, it is antisymmetry that makes a function of the form  $|\psi_1 \psi_1\rangle$  vanish (thereby enforcing the Pauli exclusion principle) while  $|\psi_1 \psi_2\rangle$  does not vanish, except at points  $\mathbf{r}_1$  and  $\mathbf{r}_2$  where  $\psi_1(\mathbf{r}_1) = \psi_2(\mathbf{r}_2)$ , and hence is acceptable. The Pauli principle is embodied in the fact that if any two or more columns (or rows) of a determinant are identical, the determinant vanishes. Antisymmetry also enforces indistinguishability of the electrons in that  $|\psi_1 \psi_1 \psi_2 \psi_2\rangle = -|\psi_1 \psi_2 \psi_1 \psi_2\rangle$ . That is, two wavefunctions which differ simply by the ordering of their spin-orbitals are equal to within a sign (+/- 1); such an overall sign difference in a wavefunction has no physical consequence because all physical properties depend on the product  $\psi^* \psi$ , which appears in any expectation value expression.

**B. Physical Consequences of Antisymmetry**

Once the rules for evaluating energies of determinantal wavefunctions and for forming functions which have proper spin and spatial symmetries have been put forth (in Chapter 11), it will be clear that antisymmetry and electron spin considerations, in addition to orbital occupancies, play substantial roles in determining energies and that it is precisely these aspects that are responsible for energy splittings among states arising from one configuration. A single example may help illustrate this point. Consider the  $1s^2 2s^2$  configuration of ethylene (ignore the other orbitals and focus on the properties of these two). As will be shown below when spin angular momentum is treated in full, the triplet spin states of this configuration are:

$$|S=1, M_S=1\rangle = |\psi_1^* \psi_2\rangle,$$

$$|S=1, M_S=-1\rangle = |\psi_2^* \psi_1\rangle,$$

and

$$|S=1, M_S=0\rangle = 2^{-1/2} [ | \uparrow \downarrow \rangle + | \downarrow \uparrow \rangle ].$$

The singlet spin state is:

$$|S=0, M_S=0\rangle = 2^{-1/2} [ | \uparrow \downarrow \rangle - | \downarrow \uparrow \rangle ].$$

To understand how the three triplet states have the same energy and why the singlet state has a different energy, and an energy different than the  $M_S=0$  triplet even though these two states are composed of the same two determinants, we proceed as follows:

1. We express the bonding and antibonding  $\sigma$  orbitals in terms of the atomic p-orbitals from which they are formed:  $\sigma = 2^{-1/2} [ L + R ]$  and  $\sigma^* = 2^{-1/2} [ L - R ]$ , where R and L denote the p-orbitals on the left and right carbon atoms, respectively.
2. We substitute these expressions into the Slater determinants that form the singlet and triplet states and collect terms and throw out terms for which the determinants vanish.
3. This then gives the singlet and triplet states in terms of atomic-orbital occupancies where it is easier to see the energy equivalences and differences.

Let us begin with the triplet states:

$$| \uparrow \downarrow \rangle = 1/2 [ |L \uparrow L \downarrow \rangle - |R \uparrow R \downarrow \rangle + |R \uparrow L \downarrow \rangle - |L \uparrow R \downarrow \rangle ]$$

$$= |R \uparrow L \downarrow \rangle;$$

$$2^{-1/2} [ | \uparrow \downarrow \rangle + | \downarrow \uparrow \rangle ] = 2^{-1/2} 1/2 [ |L \uparrow L \downarrow \rangle - |R \uparrow R \downarrow \rangle + |R \uparrow L \downarrow \rangle -$$

$$|L \uparrow R \downarrow \rangle + |L \downarrow L \uparrow \rangle - |R \downarrow R \uparrow \rangle + |R \downarrow L \uparrow \rangle - |L \downarrow R \uparrow \rangle ]$$

$$= 2^{-1/2} [ |R \uparrow L \downarrow \rangle + |R \downarrow L \uparrow \rangle ];$$

$$| \downarrow \uparrow \rangle = 1/2 [ |L \downarrow L \uparrow \rangle - |R \downarrow R \uparrow \rangle + |R \downarrow L \uparrow \rangle - |L \downarrow R \uparrow \rangle ]$$

$$= |\mathbf{R} \mathbf{L} \rangle.$$

The singlet state can be reduced in like fashion:

$$2^{-1/2} [ |\uparrow \uparrow \rangle - |\downarrow \downarrow \rangle ] = 2^{-1/2} 1/2 [ |\mathbf{L} \mathbf{L} \rangle - |\mathbf{R} \mathbf{R} \rangle + |\mathbf{R} \mathbf{L} \rangle - |\mathbf{L} \mathbf{R} \rangle - |\mathbf{L} \mathbf{L} \rangle + |\mathbf{R} \mathbf{R} \rangle - |\mathbf{R} \mathbf{L} \rangle + |\mathbf{L} \mathbf{R} \rangle ]$$

$$= 2^{-1/2} [ |\mathbf{L} \mathbf{L} \rangle - |\mathbf{R} \mathbf{R} \rangle ].$$

Notice that all three triplet states involve atomic orbital occupancy in which one electron is on one atom while the other is on the second carbon atom. In contrast, the singlet state places both electrons on one carbon (it contains two terms; one with the two electrons on the left carbon and the other with both electrons on the right carbon).

In a "valence bond" analysis of the physical content of the singlet and triplet  $^1 \ ^1$  states, it is clear that the energy of the triplet states will lie below that of the singlet because the singlet contains "zwitterion" components that can be denoted  $\text{C}^+\text{C}^-$  and  $\text{C}^-\text{C}^+$ , while the three triplet states are purely "covalent". This case provides an excellent example of how the spin and permutational symmetries of a state "conspire" to qualitatively affect its energy and even electronic character as represented in its atomic orbital occupancies.

Understanding this should provide ample motivation for learning how to form proper antisymmetric spin (and orbital) angular momentum eigenfunctions for atoms and molecules.

## Chapter 10

*Electronic Wavefunctions Must Also Possess Proper Symmetry. These Include Angular Momentum and Point Group Symmetries*

### I. Angular Momentum Symmetry and Strategies for Angular Momentum Coupling

Because the total Hamiltonian of a many-electron atom or molecule forms a mutually commutative set of operators with  $S^2$ ,  $S_z$ , and  $A = (1/N!) \prod P_{ij}$ , the exact eigenfunctions of  $H$  must be eigenfunctions of these operators. Being an eigenfunction of  $A$  forces the eigenstates to be odd under all  $P_{ij}$ . Any acceptable model or trial wavefunction should be constrained to also be an eigenfunction of these symmetry operators.

If the atom or molecule has additional symmetries (e.g., full rotation symmetry for atoms, axial rotation symmetry for linear molecules and point group symmetry for non-linear polyatomics), the trial wavefunctions should also conform to these spatial symmetries. This Chapter addresses those operators that commute with  $H$ ,  $P_{ij}$ ,  $S^2$ , and  $S_z$  and among one another for atoms, linear, and non-linear molecules.

As treated in detail in Appendix G, the full non-relativistic  $N$ -electron Hamiltonian of an atom or molecule

$$H = \sum_j \left( -\frac{\hbar^2}{2m} \nabla_j^2 - \sum_a Z_a \frac{e^2}{r_{j,a}} \right) + \sum_{j < k} \frac{e^2}{r_{j,k}}$$

commutes with the following operators:

- i. The inversion operator  $i$  and the three components of the total orbital angular momentum  $L_z = \sum_j L_z(j)$ ,  $L_y$ ,  $L_x$ , as well as the components of the total spin angular momentum  $S_z$ ,  $S_x$ , and  $S_y$  **for atoms** (but not the individual electrons'  $L_z(j)$ ,  $S_z(j)$ , etc). Hence,  $L^2$ ,  $L_z$ ,  $S^2$ ,  $S_z$  are the operators we need to form eigenfunctions of, and  $L$ ,  $M_L$ ,  $S$ , and  $M_S$  are the "good" quantum numbers.
- ii.  $L_z = \sum_j L_z(j)$ , as well as the  $N$ -electron  $S_x$ ,  $S_y$ , and  $S_z$  **for linear molecules** (also  $i$ , if the molecule has a center of symmetry). Hence,  $L_z$ ,  $S^2$ , and  $S_z$  are the operators we need to form eigenfunctions of, and  $M_L$ ,  $S$ , and  $M_S$  are the "good" quantum numbers;  $L$  no longer is!
- iii.  $S_x$ ,  $S_y$ , and  $S_z$  as well as all point group operations **for non-linear polyatomic molecules**. Hence  $S^2$ ,  $S_z$ , and the point group operations are used to characterize the

functions we need to form. When we include spin-orbit coupling into H (this adds another term to the potential that involves the spin and orbital angular momenta of the electrons),  $L^2$ ,  $L_z$ ,  $S^2$ ,  $S_z$  no longer commute with H. However,  $J_z = S_z + L_z$  and  $J^2 = (\mathbf{L} + \mathbf{S})^2$  now do commute with H.

### A. Electron Spin Angular Momentum

Individual electrons possess intrinsic spin characterized by angular momentum quantum numbers  $s$  and  $m_s$ ; for electrons,  $s = 1/2$  and  $m_s = 1/2$ , or  $-1/2$ . The  $m_s = 1/2$  spin state of the electron is represented by the symbol  $\alpha$  and the  $m_s = -1/2$  state is represented by  $\beta$ . These spin functions obey:  $S^2 \alpha = 1/2(1/2 + 1)\hbar^2 \alpha$ ,  $S_z \alpha = 1/2\hbar \alpha$ ,  $S^2 \beta = 1/2(1/2 + 1)\hbar^2 \beta$ , and  $S_z \beta = -1/2\hbar \beta$ . The  $\alpha$  and  $\beta$  spin functions are connected via lowering  $S_-$  and raising  $S_+$  operators, which are defined in terms of the x and y components of  $\mathbf{S}$  as follows:  $S_+ = S_x + iS_y$ , and  $S_- = S_x - iS_y$ . In particular  $S_+ \alpha = \hbar \beta$ ,  $S_+ \beta = 0$ ,  $S_- \alpha = \hbar \beta$ , and  $S_- \beta = 0$ . These expressions are examples of the more general relations (these relations are developed in detail in Appendix G) which all angular momentum operators and their eigenstates obey:

$$J^2 |j, m\rangle = j(j+1)\hbar^2 |j, m\rangle,$$

$$J_z |j, m\rangle = m\hbar |j, m\rangle,$$

$$J_+ |j, m\rangle = \hbar \{j(j+1) - m(m+1)\}^{1/2} |j, m+1\rangle, \text{ and}$$

$$J_- |j, m\rangle = \hbar \{j(j+1) - m(m-1)\}^{1/2} |j, m-1\rangle.$$

In a many-electron system, one must combine the spin functions of the individual electrons to generate eigenfunctions of the total  $S_z = \sum_i S_z(i)$  (expressions for  $S_x = \sum_i S_x(i)$  and  $S_y = \sum_i S_y(i)$  also follow from the fact that the total angular momentum of a collection of particles is the sum of the angular momenta, component-by-component, of the individual angular momenta) and total  $S^2$  operators because only these operators commute with the full Hamiltonian, H, and with the permutation operators  $P_{ij}$ . No longer are the individual  $S^2(i)$  and  $S_z(i)$  good quantum numbers; these operators do not commute with  $P_{ij}$ .

Spin states which are eigenfunctions of the total  $S^2$  and  $S_z$  can be formed by using angular momentum coupling methods or the explicit construction methods detailed in

Appendix (G). In the latter approach, one forms, consistent with the given electronic configuration, the spin state having maximum  $S_z$  eigenvalue (which is easy to identify as shown below and which corresponds to a state with  $S$  equal to this maximum  $S_z$  eigenvalue) and then generating states of lower  $S_z$  values and lower  $S$  values using the angular momentum raising and lowering operators ( $S_- = \sum_i S_{-i}$ ) and  $S_+ = \sum_i S_{+i}$ ).

To illustrate, consider a three-electron example with the configuration  $1s2s3s$ . Starting with the determinant  $|1s \uparrow 2s \uparrow 3s \uparrow|$ , which has the maximum  $M_s = 3/2$  and hence has  $S = 3/2$  (this function is denoted  $|3/2, 3/2\rangle$ ), apply  $S_-$  in the additive form  $S_- = \sum_i S_{-i}$  to generate the following combination of three determinants:

$$\hbar[|1s \uparrow 2s \uparrow 3s \downarrow| + |1s \uparrow 2s \downarrow 3s \uparrow| + |1s \downarrow 2s \uparrow 3s \uparrow|],$$

which, according to the above identities, must equal

$$\hbar\sqrt{3/2(3/2+1)-3/2(3/2-1)}|3/2, 1/2\rangle.$$

So the state  $|3/2, 1/2\rangle$  with  $S = 3/2$  and  $M_s = 1/2$  can be solved for in terms of the three determinants to give

$$|3/2, 1/2\rangle = 1/\sqrt{3}[|1s \uparrow 2s \uparrow 3s \downarrow| + |1s \uparrow 2s \downarrow 3s \uparrow| + |1s \downarrow 2s \uparrow 3s \uparrow|].$$

The states with  $S = 3/2$  and  $M_s = -1/2$  and  $-3/2$  can be obtained by further application of  $S_-$  to  $|3/2, 1/2\rangle$  (actually, the  $M_s = -3/2$  can be identified as the "spin flipped" image of the state with  $M_s = 3/2$  and the one with  $M_s = -1/2$  can be formed by interchanging all  $\uparrow$ 's and  $\downarrow$ 's in the  $M_s = 1/2$  state).

Of the eight total spin states (each electron can take on either  $\uparrow$  or  $\downarrow$  spin and there are three electrons, so the number of states is  $2^3$ ), the above process has identified proper combinations which yield the four states with  $S = 3/2$ . Doing so consumed the determinants with  $M_s = 3/2$  and  $-3/2$ , one combination of the three determinants with  $M_s = 1/2$ , and one combination of the three determinants with  $M_s = -1/2$ . There still remain two combinations of the  $M_s = 1/2$  and two combinations of the  $M_s = -1/2$  determinants to deal with. These functions correspond to two sets of  $S = 1/2$  eigenfunctions having  $M_s = 1/2$  and  $-1/2$ . Combinations of the determinants must be used in forming the  $S = 1/2$  functions to keep the  $S = 1/2$  eigenfunctions orthogonal to the above  $S = 3/2$  functions (which is required because  $S^2$  is a hermitian operator whose eigenfunctions belonging to different eigenvalues must be orthogonal). The two independent  $S = 1/2$ ,  $M_s = 1/2$  states

can be formed by simply constructing combinations of the above three determinants with  $M_s = 1/2$  which are orthogonal to the  $S = 3/2$  combination given above and orthogonal to each other. For example,

$$|1/2, 1/2\rangle = 1/\sqrt{2} [ |1s\ 2s\ 3s\rangle - |1s\ 2s\ 3s\rangle + 0x |1s\ 2s\ 3s\rangle ],$$

$$|1/2, 1/2\rangle = 1/\sqrt{6} [ |1s\ 2s\ 3s\rangle + |1s\ 2s\ 3s\rangle - 2x |1s\ 2s\ 3s\rangle ]$$

are acceptable (as is any combination of these two functions generated by a unitary transformation). A pair of independent orthonormal states with  $S = 1/2$  and  $M_s = -1/2$  can be generated by applying  $S_-$  to each of these two functions (or by constructing a pair of orthonormal functions which are combinations of the three determinants with  $M_s = -1/2$  and which are orthogonal to the  $S = 3/2$ ,  $M_s = -1/2$  function obtained as detailed above).

The above treatment of a three-electron case shows how to generate quartet (spin states are named in terms of their spin degeneracies  $2S+1$ ) and doublet states for a configuration of the form

$1s^2 2s 3s$ . Not all three-electron configurations have both quartet and doublet states; for example, the  $1s^2 2s$  configuration only supports one doublet state. The methods used above to generate  $S = 3/2$  and

$S = 1/2$  states are valid for any three-electron situation; however, some of the determinantal functions vanish if doubly occupied orbitals occur as for  $1s^2 2s$ . In particular, the  $|1s\ 1s\ 2s\rangle$  and

$$|1s\ 1s\ 2s\rangle \text{ with } M_s = 3/2, -3/2 \text{ and } |1s\ 1s\ 2s\rangle \text{ and } |1s\ 1s\ 2s\rangle \text{ with } M_s = 1/2, -1/2$$

determinants vanish because they violate the Pauli principle; only  $|1s\ 1s\ 2s\rangle$  and  $|1s\ 1s\ 2s\rangle$

do not vanish. These two remaining determinants form the  $S = 1/2$ ,  $M_s = 1/2$ ,  $-1/2$  doublet spin functions which pertain to the  $1s^2 2s$  configuration. It should be noted that all closed-shell components of a configuration (e.g., the  $1s^2$  part of  $1s^2 2s$  or the  $1s^2 2s^2 2p^6$  part of  $1s^2 2s^2 2p^6 3s^1 3p^1$ ) must involve  $\uparrow\downarrow$  spin functions for each doubly occupied orbital and, as such, can contribute nothing to the total  $M_s$  value; only the open-shell components need to be treated with the angular momentum operator tools to arrive at proper total-spin eigenstates.

In summary, proper spin eigenfunctions must be constructed from antisymmetric (i.e., determinantal) wavefunctions as demonstrated above because the total  $S^2$  and total  $S_z$  remain valid symmetry operators for many-electron systems. Doing so results in the spin-adapted wavefunctions being expressed as combinations of determinants with coefficients determined via spin angular momentum techniques as demonstrated above. In

configurations with closed-shell components, not all spin functions are possible because of the antisymmetry of the wavefunction; in particular, any closed-shell parts must involve spin pairings for each of the doubly occupied orbitals, and, as such, contribute zero to the total  $M_S$ .

## B. Vector Coupling of Angular Momenta

Given two angular momenta (of any kind)  $\mathbf{L}_1$  and  $\mathbf{L}_2$ , when one generates states that are eigenstates of their vector sum  $\mathbf{L} = \mathbf{L}_1 + \mathbf{L}_2$ , one can obtain L values of  $L_1 + L_2, L_1 + L_2 - 1, \dots, |L_1 - L_2|$ . This can apply to two electrons for which the total spin S can be 1 or 0 as illustrated in detail above, or to a p and a d orbital for which the total orbital angular momentum L can be 3, 2, or 1. Thus for a  $p^1d^1$  electronic configuration,  $^3F, ^1F, ^3D, ^1D, ^3P,$  and  $^1P$  energy levels (and corresponding wavefunctions) arise. Here the term symbols are specified as the spin degeneracy ( $2S+1$ ) and the letter that is associated with the L-value. If spin-orbit coupling is present, the  $^3F$  level further splits into  $J=4, 3,$  and  $2$  levels which are denoted  $^3F_4, ^3F_3,$  and  $^3F_2$ .

This simple "vector coupling" method applies to any angular momenta. However, if the angular momenta are "equivalent" in the sense that they involve indistinguishable particles that occupy the same orbital shell (e.g.,  $2p^3$  involves 3 equivalent electrons;  $2p^13p^14p^1$  involves 3 non-equivalent electrons;  $2p^23p^1$  involves 2 equivalent electrons and one non-equivalent electron), the Pauli principle eliminates some of the expected term symbols (i.e., when the corresponding wavefunctions are formed, some vanish because their Slater determinants vanish). Later in this section, techniques for dealing with the equivalent-angular momenta case are introduced. These techniques involve using the above tools to obtain a list of candidate term symbols after which Pauli-violating term symbols are eliminated.

## C. Non-Vector Coupling of Angular Momenta

For linear molecules, one does not vector couple the orbital angular momenta of the individual electrons (because only  $L_z$  not  $L^2$  commutes with H), but one does vector couple the electrons' spin angular momenta. Coupling of the electrons' orbital angular momenta involves simply considering the various  $L_z$  eigenvalues that can arise from adding the  $L_z$  values of the individual electrons. For example, coupling two orbitals (each of which can have  $m = \pm 1$ ) can give  $M_L = 1+1, 1-1, -1+1,$  and  $-1-1,$  or  $2, 0, 0,$  and  $-2$ . The level with  $M_L = \pm 2$  is called a state (much like an orbital with  $m = \pm 2$  is called a orbital), and the

two states with  $M_L = 0$  are called  $\pi$  states. States with  $L_z$  eigenvalues of  $M_L$  and  $-M_L$  are degenerate because the total energy is independent of which direction the electrons are moving about the linear molecule's axis (just a  $+1$  and  $-1$  orbitals are degenerate).

Again, if the two electrons are non-equivalent, all possible couplings arise (e.g., a  $1^1 1^1$  configuration yields  $^3\Sigma$ ,  $^3\Pi$ ,  $^3\Delta$ ,  $^1\Sigma$ ,  $^1\Pi$ , and  $^1\Delta$  states). In contrast, if the two electrons are equivalent, certain of the term symbols are Pauli forbidden. Again, techniques for dealing with such cases are treated later in this Chapter.

#### D. Direct Products for Non-Linear Molecules

For non-linear polyatomic molecules, one vector couples the electrons' spin angular momenta but their orbital angular momenta are not even considered. Instead, their point group symmetries must be combined, by forming direct products, to determine the symmetries of the resultant spin-orbital product states. For example, the  $b_1^1 b_2^1$  configuration in  $C_{2v}$  symmetry gives rise to  $^3A_2$  and  $^1A_2$  term symbols. The  $e^1 e^1$  configuration in  $C_{3v}$  symmetry gives  $^3E$ ,  $^3A_2$ ,  $^3A_1$ ,  $^1E$ ,  $^1A_2$ , and  $^1A_1$  term symbols. For two equivalent electrons such as in the  $e^2$  configuration, certain of the  $^3E$ ,  $^3A_2$ ,  $^3A_1$ ,  $^1E$ ,  $^1A_2$ , and  $^1A_1$  term symbols are Pauli forbidden. Once again, the methods needed to identify which term symbols arise in the equivalent-electron case are treated later.

One needs to learn how to tell which term symbols will be Pauli excluded, and to learn how to write the spin-orbit product wavefunctions corresponding to each term symbol and to evaluate the corresponding term symbols' energies.

## II. Atomic Term Symbols and Wavefunctions

### A. Non-Equivalent Orbital Term Symbols

When coupling non-equivalent angular momenta (e.g., a spin and an orbital angular momenta or two orbital angular momenta of non-equivalent electrons), one vector couples using the fact that the coupled angular momenta range from the sum of the two individual angular momenta to the absolute value of their difference. For example, when coupling the spins of two electrons, the total spin  $S$  can be 1 or 0; when coupling a  $p$  and a  $d$  orbital, the total orbital angular momentum can be 3, 2, or 1. Thus for a  $p^1 d^1$  electronic configuration,  $^3F$ ,  $^1F$ ,  $^3D$ ,  $^1D$ ,  $^3P$ , and  $^1P$  energy levels (and corresponding wavefunctions) arise. The energy differences among these levels has to do with the different electron-electron repulsions that occur in these levels; that is, their wavefunctions involve different

occupancy of the p and d orbitals and hence different repulsion energies. If spin-orbit coupling is present, the L and S angular momenta are further vector coupled. For example, the  ${}^3F$  level splits into J= 4, 3, and 2 levels which are denoted  ${}^3F_4$ ,  ${}^3F_3$ , and  ${}^3F_2$ . The energy differences among these J-levels are caused by spin-orbit interactions.

## B. Equivalent Orbital Term Symbols

If equivalent angular momenta are coupled (e.g., to couple the orbital angular momenta of a  $p^2$  or  $d^3$  configuration), one must use the "box" method to determine which of the term symbols, that would be expected to arise if the angular momenta were non-equivalent, violate the Pauli principle. To carry out this step, one forms all possible unique (determinantal) product states with non-negative  $M_L$  and  $M_S$  values and arranges them into groups according to their  $M_L$  and  $M_S$  values. For example, the boxes appropriate to the  $p^2$  orbital occupancy are shown below:

$M_L$	2	1	0
$M_S$	1	$ p_1 p_0 \rangle$	$ p_1 p_{-1} \rangle$
0	$ p_1 p_1 \rangle$	$ p_1 p_0 \rangle,  p_0 p_1 \rangle$	$ p_1 p_{-1} \rangle,$ $ p_{-1} p_1 \rangle,$ $ p_0 p_0 \rangle$

There is no need to form the corresponding states with negative  $M_L$  or negative  $M_S$  values because they are simply "mirror images" of those listed above. For example, the state with  $M_L = -1$  and  $M_S = -1$  is  $|p_{-1} p_0 \rangle$ , which can be obtained from the  $M_L = 1, M_S = 1$  state  $|p_1 p_0 \rangle$  by replacing  $p_1$  by  $p_{-1}$ .

Given the box entries, one can identify those term symbols that arise by applying the following procedure over and over until all entries have been accounted for:

1. One identifies the highest  $M_S$  value (this gives a value of the total spin quantum number that arises,  $S$ ) in the box. For the above example, the answer is  $S = 1$ .
2. For all product states of this  $M_S$  value, one identifies the highest  $M_L$  value (this gives a value of the total orbital angular momentum,  $L$ , that can arise for this  $S$ ). For the above example, the highest  $M_L$  within the  $M_S = 1$  states is  $M_L = 1$  (not  $M_L = 2$ ), hence  $L = 1$ .
3. Knowing an  $S, L$  combination, one knows the first term symbol that arises from this configuration. In the  $p^2$  example, this is  $^3P$ .
4. Because the level with this  $L$  and  $S$  quantum numbers contains  $(2L+1)(2S+1)$  states with  $M_L$  and  $M_S$  quantum numbers running from  $-L$  to  $L$  and from  $-S$  to  $S$ , respectively, one must remove from the original box this number of product states. To do so, one simply erases from the box one entry with each such  $M_L$  and  $M_S$  value. Actually, since the box need only show those entries with non-negative  $M_L$  and  $M_S$  values, only these entries need be explicitly deleted. In the  $^3P$  example, this amounts to deleting nine product states with  $M_L, M_S$  values of 1,1; 1,0; 1,-1; 0,1; 0,0; 0,-1; -1,1; -1,0; -1,-1.
5. After deleting these entries, one returns to step 1 and carries out the process again. For the  $p^2$  example, the box after deleting the first nine product states looks as follows (those that appear in italics should be viewed as already cancelled in counting all of the  $^3P$  states):

$M_L$	2	1	0

$$\begin{array}{r}
M_S = 1 \\
\\
0
\end{array}
\begin{array}{l}
|p_1 p_0 \rangle \\
|p_1 p_1 \rangle \\
|p_1 p_0 \rangle, |p_0 p_1 \rangle \\
|p_1 p_{-1} \rangle, \\
|p_{-1} p_1 \rangle, \\
|p_0 p_0 \rangle
\end{array}$$

It should be emphasized that the process of deleting or crossing off entries in various  $M_L$ ,  $M_S$  boxes involves only counting how many states there are; by no means do we identify the particular  $L, S, M_L, M_S$  wavefunctions when we cross out any particular entry in a box. For example, when the  $|p_1 p_0 \rangle$  product is deleted from the  $M_L = 1, M_S = 0$  box in accounting for the states in the  $^3P$  level, we do not claim that  $|p_1 p_0 \rangle$  itself is a member of the  $^3P$  level; the  $|p_0 p_1 \rangle$  product state could just as well been eliminated when accounting for the  $^3P$  states. As will be shown later, the  $^3P$  state with  $M_L = 1, M_S = 0$  will be a combination of  $|p_1 p_0 \rangle$  and  $|p_0 p_1 \rangle$ .

Returning to the  $p^2$  example at hand, after the  $^3P$  term symbol's states have been accounted for, the highest  $M_S$  value is 0 (hence there is an  $S=0$  state), and within this  $M_S$  value, the highest  $M_L$  value is 2 (hence there is an  $L=2$  state). This means there is a  $^1D$  level with five states having  $M_L = 2, 1, 0, -1, -2$ . Deleting five appropriate entries from the above box (again denoting deletions by italics) leaves the following box:

$M_L$	2	1	0
$M_S$	1	$ p_1 p_0 \rangle$	$ p_1 p_{-1} \rangle$
0	$ p_1 p_1 \rangle$	$ p_1 p_0 \rangle,  p_0 p_1 \rangle$	$ p_1 p_{-1} \rangle,$ $ p_{-1} p_1 \rangle,$ $ p_0 p_0 \rangle$

The only remaining entry, which thus has the highest  $M_S$  and  $M_L$  values, has  $M_S = 0$  and  $M_L = 0$ . Thus there is also a  $^1S$  level in the  $p^2$  configuration.

Thus, unlike the non-equivalent  $2p^1 3p^1$  case, in which  $^3P$ ,  $^1P$ ,  $^3D$ ,  $^1D$ ,  $^3S$ , and  $^1S$  levels arise, only the  $^3P$ ,  $^1D$ , and  $^1S$  arise in the  $p^2$  situation. This "box method" is necessary to carry out whenever one is dealing with equivalent angular momenta.

If one has mixed equivalent and non-equivalent angular momenta, one can determine all possible couplings of the equivalent angular momenta using this method and then use the simpler vector coupling method to add the non-equivalent angular momenta to each of these coupled angular momenta. For example, the  $p^2 d^1$  configuration can be handled by vector coupling (using the straightforward non-equivalent procedure)  $L=2$  (the  $d$  orbital) and  $S=1/2$  (the third electron's spin) to each of  $^3P$ ,  $^1D$ , and  $^1S$ . The result is  $^4F$ ,  $^4D$ ,  $^4P$ ,  $^2F$ ,  $^2D$ ,  $^2P$ ,  $^2G$ ,  $^2F$ ,  $^2D$ ,  $^2P$ ,  $^2S$ , and  $^2D$ .

### C. Atomic Configuration Wavefunctions

To express, in terms of Slater determinants, the wavefunctions corresponding to each of the states in each of the levels, one proceeds as follows:

1. For each  $M_S$ ,  $M_L$  combination for which one can write down only one product function (i.e., in the non-equivalent angular momentum situation, for each case where only one product function sits at a given box row and column point), that product function itself is one of the desired states. For the  $p^2$  example, the  $|p_1 p_0 \rangle$  and  $|p_1 p_{-1} \rangle$  (as well as their four other  $M_L$  and  $M_S$  "mirror images") are members of the  $^3P$  level (since they have  $M_S = \pm 1$ ) and  $|p_1 p_1 \rangle$  and its  $M_L$  mirror image are members of the  $^1D$  level (since they have  $M_L = \pm 2$ ).

2. After identifying as many such states as possible by inspection, one uses  $L_{\pm}$  and  $S_{\pm}$  to generate states that belong to the same term symbols as those already identified but which have higher or lower  $M_L$  and/or  $M_S$  values.

3. If, after applying the above process, there are term symbols for which states have not yet been formed, one may have to construct such states by forming linear combinations that are orthogonal to all those states that have thus far been found.

To illustrate the use of raising and lowering operators to find the states that can not be identified by inspection, let us again focus on the  $p^2$  case. Beginning with three of the  $^3P$  states that are easy to recognize,  $|p_1 p_0\rangle$ ,  $|p_1 p_{-1}\rangle$ , and  $|p_{-1} p_0\rangle$ , we apply  $S_-$  to obtain the  $M_S=0$  functions:

$$\begin{aligned} S_- ^3P(M_L=1, M_S=1) &= [S_-(1) + S_-(2)] |p_1 p_0\rangle \\ &= \hbar(1(2)-1(0))^{1/2} ^3P(M_L=1, M_S=0) \\ &= \hbar(1/2(3/2)-1/2(-1/2))^{1/2} |p_1 p_0\rangle + \hbar(1)^{1/2} |p_1 p_{-1}\rangle, \end{aligned}$$

so,

$$^3P(M_L=1, M_S=0) = 2^{-1/2} [|p_1 p_0\rangle + |p_1 p_{-1}\rangle].$$

The same process applied to  $|p_1 p_{-1}\rangle$  and  $|p_{-1} p_0\rangle$  gives

$$1/\sqrt{2} [|p_1 p_{-1}\rangle + |p_1 p_{-1}\rangle] \text{ and } 1/\sqrt{2} [|p_{-1} p_0\rangle + |p_{-1} p_0\rangle],$$

respectively.

The  $^3P(M_L=1, M_S=0) = 2^{-1/2} [|p_1 p_0\rangle + |p_1 p_{-1}\rangle]$  function can be acted on with  $L_-$  to generate  $^3P(M_L=0, M_S=0)$ :

$$\begin{aligned} L_- ^3P(M_L=1, M_S=0) &= [L_-(1) + L_-(2)] 2^{-1/2} [|p_1 p_0\rangle + |p_1 p_{-1}\rangle] \\ &= \hbar(1(2)-1(0))^{1/2} ^3P(M_L=0, M_S=0) \\ &= \hbar(1(2)-1(0))^{1/2} 2^{-1/2} [|p_0 p_0\rangle + |p_0 p_0\rangle] \\ &\quad + \hbar(1(2)-0(-1))^{1/2} 2^{-1/2} [|p_1 p_{-1}\rangle + |p_1 p_{-1}\rangle], \end{aligned}$$

so,

$$^3P(M_L=0, M_S=0) = 2^{-1/2} [|p_1 p_{-1}\rangle + |p_1 p_{-1}\rangle].$$

The  $^1D$  term symbol is handled in like fashion. Beginning with the  $M_L = 2$  state  $|p_1 p_1 \rangle$ , one applies  $L_-$  to generate the  $M_L = 1$  state:

$$\begin{aligned} L_- ^1D(M_L=2, M_S=0) &= [L_-(1) + L_-(2)] |p_1 p_1 \rangle \\ &= \hbar(2(3)-2(1))^{1/2} ^1D(M_L=1, M_S=0) \\ &= \hbar(1(2)-1(0))^{1/2} [|p_0 p_1 \rangle + |p_1 p_0 \rangle], \end{aligned}$$

so,

$$^1D(M_L=1, M_S=0) = 2^{-1/2} [|p_0 p_1 \rangle + |p_1 p_0 \rangle].$$

Applying  $L_-$  once more generates the  $^1D(M_L=0, M_S=0)$  state:

$$\begin{aligned} L_- ^1D(M_L=1, M_S=0) &= [L_-(1) + L_-(2)] 2^{-1/2} [|p_0 p_1 \rangle + |p_1 p_0 \rangle] \\ &= \hbar(2(3)-1(0))^{1/2} ^1D(M_L=0, M_S=0) \\ &= \hbar(1(2)-0(-1))^{1/2} 2^{-1/2} [|p_{-1} p_1 \rangle + |p_1 p_{-1} \rangle] \\ &\quad + \hbar(1(2)-1(0))^{1/2} 2^{-1/2} [|p_0 p_0 \rangle + |p_0 p_0 \rangle], \end{aligned}$$

so,

$$^1D(M_L=0, M_S=0) = 6^{-1/2} [2|p_0 p_0 \rangle + |p_{-1} p_1 \rangle + |p_1 p_{-1} \rangle].$$

Notice that the  $M_L=0, M_S=0$  states of  $^3P$  and of  $^1D$  are given in terms of the three determinants that appear in the "center" of the  $p^2$  box diagram:

$$^1D(M_L=0, M_S=0) = 6^{-1/2} [2|p_0 p_0 \rangle + |p_{-1} p_1 \rangle + |p_1 p_{-1} \rangle],$$

$$^3P(M_L=0, M_S=0) = 2^{-1/2} [|p_1 p_{-1} \rangle + |p_{-1} p_1 \rangle]$$

$$= 2^{-1/2} [-|p_{-1} p_1 \rangle + |p_1 p_{-1} \rangle].$$

The only state that has eluded us thus far is the  $^1S$  state, which also has  $M_L=0$  and  $M_S=0$ . To construct this state, which must also be some combination of the three determinants with  $M_L=0$  and  $M_S=0$ , we use the fact that the  $^1S$  wavefunction must be orthogonal to the

$^3P$  and  $^1D$  functions because  $^1S$ ,  $^3P$ , and  $^1D$  are eigenfunctions of the hermitian operator  $L^2$  having different eigenvalues. The state that is normalized and is a combination of  $|p_0 p_0\rangle$ ,  $|p_{-1} p_1\rangle$ , and  $|p_1 p_{-1}\rangle$  is given as follows:

$$^1S = 3^{-1/2} [ |p_0 p_0\rangle - |p_{-1} p_1\rangle - |p_1 p_{-1}\rangle ].$$

The procedure used here to form the  $^1S$  state illustrates point 3 in the above prescription for determining wavefunctions. Additional examples for constructing wavefunctions for atoms are provided later in this chapter and in Appendix G.

#### D. Inversion Symmetry

One more quantum number, that relating to the inversion ( $i$ ) symmetry operator can be used in atomic cases because the total potential energy  $V$  is unchanged when all of the electrons have their position vectors subjected to inversion ( $i \mathbf{r} = -\mathbf{r}$ ). This quantum number is straightforward to determine. Because each  $L$ ,  $S$ ,  $M_L$ ,  $M_S$ ,  $H$  state discussed above consist of a few (or, in the case of configuration interaction several) symmetry adapted combinations of Slater determinant functions, the effect of the inversion operator on such a wavefunction can be determined by:

(i) applying  $i$  to each orbital occupied in thereby generating a  $\pm 1$  factor for each orbital (+1 for  $s$ ,  $d$ ,  $g$ ,  $i$ , etc orbitals; -1 for  $p$ ,  $f$ ,  $h$ ,  $j$ , etc orbitals),

(ii) multiplying these  $\pm 1$  factors to produce an overall sign for the character of under  $i$ .

When this overall sign is positive, the function is termed "even" and its term symbol is appended with an "e" superscript (e.g., the  $^3P$  level of the O atom, which has  $1s^2 2s^2 2p^4$  occupancy is labeled  $^3P^e$ ); if the sign is negative is called "odd" and the term symbol is so amended (e.g., the  $^3P$  level of  $1s^2 2s^1 2p^1$   $B^+$  ion is labeled  $^3P^o$ ).

#### E. Review of Atomic Cases

The orbitals of an atom are labeled by  $l$  and  $m$  quantum numbers; the orbitals belonging to a given energy and  $l$  value are  $2l+1$ - fold degenerate. The many-electron Hamiltonian,  $H$ , of an atom and the antisymmetrizer operator  $A = (1/N!) \sum_p s_p P$  commute with total  $L_z = \sum_i L_z(i)$ , as in the linear-molecule case. The additional symmetry present in the spherical atom reflects itself in the fact that  $L_x$ , and  $L_y$  now also commute with  $H$  and  $A$ . However, since  $L_z$  does not commute with  $L_x$  or  $L_y$ , new quantum

numbers can not be introduced as symmetry labels for these other components of  $\mathbf{L}$ . A new symmetry label does arise when  $L^2 = L_z^2 + L_x^2 + L_y^2$  is introduced;  $L^2$  commutes with  $H$ ,  $A$ , and  $L_z$ , so proper eigenstates (and trial wavefunctions) can be labeled with  $L$ ,  $M_L$ ,  $S$ ,  $M_S$ , and  $H$  quantum numbers.

To identify the states which arise from a given atomic configuration and to construct properly symmetry-adapted determinantal wave functions corresponding to these symmetries, one must employ  $L$  and  $M_L$  and  $S$  and  $M_S$  angular momentum tools. One first identifies those determinants with maximum  $M_S$  (this then defines the maximum  $S$  value that occurs); within that set of determinants, one then identifies the determinant(s) with maximum  $M_L$  (this identifies the highest  $L$  value). This determinant has  $S$  and  $L$  equal to its  $M_S$  and  $M_L$  values (this can be verified, for example for  $L$ , by acting on this determinant with  $L^2$  in the form

$$L^2 = L \cdot L + L_z^2 + \hbar L_z$$

and realizing that  $L_+$  acting on the state must vanish); other members of this  $L, S$  energy level can be constructed by sequential application of  $S_-$  and  $L_- = \sum_i L_-(i)$ . Having exhausted a set of  $(2L+1)(2S+1)$  combinations of the determinants belonging to the given configuration, one proceeds to apply the same procedure to the remaining determinants (or combinations thereof). One identifies the maximum  $M_S$  and, within it, the maximum  $M_L$  which thereby specifies another  $S, L$  label and a new "maximum" state. The determinantal functions corresponding to these  $L, S$  (and various  $M_L, M_S$ ) values can be constructed by applying  $S_-$  and  $L_-$  to this "maximum" state. This process is continued until all of the states and their determinantal wave functions are obtained.

As illustrated above, any  $p^2$  configuration gives rise to  $^3P^e$ ,  $^1D^e$ , and  $^1S^e$  levels which contain nine, five, and one state respectively. The use of  $L$  and  $S$  angular momentum algebra tools allows one to identify the wavefunctions corresponding to these states. As shown in detail in Appendix G, in the event that spin-orbit coupling causes the Hamiltonian,  $\mathbf{H}$ , not to commute with  $\mathbf{L}$  or with  $\mathbf{S}$  but only with their vector sum  $\mathbf{J} = \mathbf{L} + \mathbf{S}$ , then these  $L^2 S^2 L_z S_z$  eigenfunctions must be coupled (i.e., recombined) to generate  $J^2 J_z$  eigenstates. The steps needed to effect this coupling are developed and illustrated for the above  $p^2$  configuration case in Appendix G.

In the case of a pair of non-equivalent  $p$  orbitals (e.g., in a  $2p^1 3p^1$  configuration), even more states would arise. They can also be found using the tools provided above. Their symmetry labels can be obtained by vector coupling (see Appendix G) the spin and orbital angular momenta of the two subsystems. The orbital angular momentum coupling

with  $l = 1$  and  $l = 1$  gives  $L = 2, 1,$  and  $0$  or D, P, and S states. The spin angular momentum coupling with  $s = 1/2$  and  $s = 1/2$  gives  $S = 1$  and  $0$ , or triplet and singlet states. So, vector coupling leads to the prediction that  $^3D^e, ^1D^e, ^3P^e, ^1P^e, ^3S^e,$  and  $^1S^e$  states can be formed from a pair of non-equivalent p orbitals. It is seen that more states arise when non-equivalent orbitals are involved; for equivalent orbitals, some determinants vanish, thereby decreasing the total number of states that arise.

### III. Linear Molecule Term Symbols and Wavefunctions

#### A. Non-Equivalent Orbital Term Symbols

Equivalent angular momenta arising in linear molecules also require use of specialized angular momentum coupling. Their spin angular momenta are coupled exactly as in the atomic case because both for atoms and linear molecules,  $S^2$  and  $S_z$  commute with H. However, unlike atoms, linear molecules no longer permit  $L^2$  to be used as an operator that commutes with H;  $L_z$  still does, but  $L^2$  does not. As a result, when coupling non-equivalent linear molecule angular momenta, one vector couples the electron spins as before. However, in place of vector coupling the individual orbital angular momenta, one adds the individual  $L_z$  values to obtain the  $L_z$  values of the coupled system. For example, the  $\pi^1 \pi^1$  configuration gives rise to  $S=1$  and  $S=0$  spin states. The individual  $m_l$  values of the two pi-orbitals can be added to give  $M_L = 1+1, 1-1, -1+1,$  and  $-1-1,$  or  $2, 0, 0,$  and  $-2$ . The  $M_L = 2$  and  $-2$  cases are degenerate (just as the  $m_l = 2$  and  $-2$  orbitals are and the  $m_l = 1$  and  $-1$  orbitals are) and are denoted by the term symbol  $^3\Delta_g$ ; there are two distinct  $M_L = 0$  states that are denoted  $^3\Sigma_g^+$  and  $^3\Sigma_g^-$ . Hence, the  $\pi^1 \pi^1$  configuration yields  $^3\Delta_g, ^3\Sigma_g^+, ^3\Sigma_g^-, ^1\Delta_g, ^1\Sigma_g^+,$  and  $^1\Sigma_g^-$  term symbols.

#### B. Equivalent-Orbital Term Symbols

To treat the equivalent-orbital case  $\pi^2$ , one forms a box diagram as in the atom case:

	$M_L = 2$	$1$	$0$
$M_S = 1$			1 -1
$M_S = 0$	1 1		1 -1  ,

$$| -1 \quad 1 |$$

The process is very similar to that used for atoms. One first identifies the highest  $M_S$  value (and hence an  $S$  value that occurs) and within that  $M_S$ , the highest  $M_L$ . However, the highest  $M_L$  does not specify an  $L$ -value, because  $L$  is no longer a "good quantum number" because  $L^2$  no longer commutes with  $H$ . Instead, we simply take the highest  $M_L$  value (and minus this value) as specifying a  $l$ ,  $l$ ,  $l$ ,  $l$ , etc. term symbol. In the above example, the highest  $M_S$  value is  $M_S = 1$ , so there is an  $S = 1$  level. Within  $M_S = 1$ , the highest  $M_L = 0$ ; hence, there is a  $^3$  level.

After deleting from the box diagram entries corresponding to  $M_S$  values ranging from  $-S$  to  $S$  and  $M_L$  values of  $M_L$  and  $-M_L$ , one has (again using italics to denote the deleted entries):

$M_L$	2	1	0
-----			
$M_S$	1		/ <i>1</i> <i>-1</i> /
	0	1 1	/ <i>1</i> <i>-1</i> /   -1 1

Among the remaining entries, the highest  $M_S$  value is  $M_S = 0$ , and within this  $M_S$  the highest  $M_L$  is  $M_L = 2$ . Thus, there is a  $^1$  state. Deleting entries with  $M_S = 0$  and  $M_L = 2$  and  $-2$ , one has left the following box diagram:

$M_L$	2	1	0
-----			
$M_S$	1		/ <i>1</i> <i>-1</i> /
	0	/ <i>1</i> <i>1</i> /	/ <i>1</i> <i>-1</i> /   -1 1

There still remains an entry with  $M_S = 0$  and  $M_L = 0$ ; hence, there is also a  $^1$  level.

Recall that the non-equivalent  $1^1 1^1$  case yielded  $^3$ ,  $^3$ ,  $^3$ ,  $^1$ ,  $^1$ , and  $^1$  term symbols. The equivalent  $2^2$  case yields only  $^3$ ,  $^1$ , and  $^1$  term symbols. Again,

whenever one is faced with equivalent angular momenta in a linear-molecule case, one must use the box method to determine the allowed term symbols. If one has a mixture of equivalent and non-equivalent angular momenta, it is possible to treat the equivalent angular momenta using boxes and to then add in the non-equivalent angular momenta using the more straightforward technique. For example, the  $2^2 1$  configuration can be treated by coupling the  $2^2$  as above to generate  $3^1$ ,  $1^1$ , and  $1^1$  and then vector coupling the spin of the third electron and additively coupling the  $m_l = 2$  and  $-2$  of the third orbital. The resulting term symbols are  $4^2$ ,  $2^2$ ,  $2^2$ ,  $2^2$ , and  $2^2$  (e.g., for the  $1^1$  intermediate state, adding the orbital's  $m_l$  values gives total  $M_L$  values of  $M_L = 2+2, 2-2, -2+2,$  and  $-2-2,$  or  $4, 0, 0,$  and  $-4$ ).

### C. Linear-Molecule Configuration Wavefunctions

Procedures analogous to those used for atoms can be applied to linear molecules. However, in this case only  $S_{\pm}$  can be used;  $L_{\pm}$  no longer applies because  $L$  is no longer a good quantum number. One begins as in the atom case by identifying determinantal functions for which  $M_L$  and  $M_S$  are unique. In the  $2^2$  example considered above, these states include  $|1 \uparrow 1 \downarrow\rangle$ ,  $|1 \uparrow 1 \uparrow\rangle$ , and their mirror images. These states are members of the  $3^1$  and  $1^1$  levels, respectively, because the first has  $M_S=1$  and because the latter has  $M_L = 2$ .

Applying  $S_-$  to this  $3^1$  state with  $M_S=1$  produces the  $3^0$  state with  $M_S = 0$ :

$$S_- 3^1 (M_L=0, M_S=1) = [S_-(1) + S_-(2)] |1 \uparrow 1 \downarrow\rangle$$

$$= \hbar(1(2)-1(0))^{1/2} 3^0 (M_L=0, M_S=0)$$

$$= \hbar (1)^{1/2} [|1 \uparrow 1 \downarrow\rangle + |1 \downarrow 1 \uparrow\rangle],$$

so,

$$3^0 (M_L=0, M_S=0) = 2^{-1/2} [|1 \uparrow 1 \downarrow\rangle + |1 \downarrow 1 \uparrow\rangle].$$

The only other state that can have  $M_L=0$  and  $M_S=0$  is the  $1^0$  state, which must itself be a combination of the two determinants,  $|1 \uparrow 1 \downarrow\rangle$  and  $|1 \downarrow 1 \uparrow\rangle$ , with  $M_L=0$  and  $M_S=0$ .

Because the  $1^0$  state has to be orthogonal to the  $3^0$  state, the combination must be

$$1^0 = 2^{-1/2} [|1 \uparrow 1 \downarrow\rangle - |1 \downarrow 1 \uparrow\rangle].$$

As with the atomic systems, additional examples are provided later in this chapter and in Appendix G.

#### D. Inversion Symmetry and $\sigma_v$ Reflection Symmetry

For homonuclear molecules (e.g., O<sub>2</sub>, N<sub>2</sub>, etc.) the inversion operator  $i$  (where inversion of all electrons now takes place through the center of mass of the nuclei rather than through an individual nucleus as in the atomic case) is also a valid symmetry, so wavefunctions may also be labeled as even or odd. The former functions are referred to as **gerade** (g) and the latter as **ungerade** (u) (derived from the German words for even and odd). The g or u character of a term symbol is straightforward to determine. Again one

(i) applies  $i$  to each orbital occupied in  $\Psi$  thereby generating a  $\pm 1$  factor for each orbital (+1 for  $\sigma$ ,  $\pi^*$ ,  $\delta$ ,  $\delta^*$ , etc orbitals; -1 for  $\pi^*$ ,  $\sigma$ ,  $\pi$ ,  $\delta$ , etc orbitals),

(ii) multiplying these  $\pm 1$  factors to produce an overall sign for the character of  $\Psi$  under  $i$ .

When this overall sign is positive, the function  $\Psi$  is gerade and its term symbol is appended with a "g" subscript (e.g., the <sup>3</sup> level of the O<sub>2</sub> molecule, which has  $\sigma_u^4 \pi_g^{*2}$  occupancy is labeled <sup>3</sup> g); if the sign is negative,  $\Psi$  is ungerade and the term symbol is so amended (e.g., the <sup>3</sup> level of the  $1 \sigma_g^2 1 \pi_u^2 1 \sigma_g^1 1 \pi_u^1$  configuration of the Li<sub>2</sub> molecule is labeled <sup>3</sup> u).

Finally, for linear molecules in  $\Sigma$  states, the wavefunctions can be labeled by one additional quantum number that relates to their symmetry under reflection of all electrons through a  $\sigma_v$  plane passing through the molecule's C axis. If  $\Sigma$  is even, a + sign is appended as a superscript to the term symbol; if  $\Sigma$  is odd, a - sign is added.

To determine the  $\sigma_v$  symmetry of  $\Psi$ , one first applies  $\sigma_v$  to each orbital in  $\Psi$ . Doing so replaces the azimuthal angle  $\phi$  of the electron in that orbital by  $2\pi - \phi$ ; because orbitals of linear molecules depend on  $\phi$  as  $\exp(im\phi)$ , this changes the orbital into  $\exp(im(-\phi)) \exp(2i\pi) = \exp(-im\phi)$ . In effect,  $\sigma_v$  applied to  $\Psi$  changes the signs of all of the  $m$  values of the orbitals in  $\Psi$ . One then determines whether the resultant  $\sigma_v \Psi$  is equal to or opposite in sign from the original  $\Psi$  by inspection. For example, the <sup>3</sup> g ground state of O<sub>2</sub>, which has a Slater determinant function

$$|S=1, M_S=1\rangle = \frac{1}{\sqrt{2}} \left[ \psi_{1,1}^* \psi_{1,-1}^* - \psi_{1,-1}^* \psi_{1,1}^* \right]$$

$$= 2^{-1/2} \left[ \psi_{1,1}^*(\mathbf{r}_1) \psi_{1,-1}^*(\mathbf{r}_2) - \psi_{1,-1}^*(\mathbf{r}_2) \psi_{1,1}^*(\mathbf{r}_1) \right].$$

Recognizing that  $\psi_{\uparrow}^* = \psi_{\downarrow}^*$  and  $\psi_{\downarrow}^* = \psi_{\uparrow}^*$ , then gives

$$\begin{aligned} \psi_{\uparrow} |S=1, M_S=1\rangle &= | \psi_{\uparrow}^* \psi_{\downarrow}^* | \\ &= 2^{-1/2} [ \psi_{\uparrow}^*(\mathbf{r}_1) \psi_{\downarrow}^*(\mathbf{r}_2) - \psi_{\downarrow}^*(\mathbf{r}_2) \psi_{\uparrow}^*(\mathbf{r}_1) ] \\ &= (-1)^{2-1/2} [ \psi_{\uparrow}^*(\mathbf{r}_1) \psi_{\downarrow}^*(\mathbf{r}_2) - \psi_{\downarrow}^*(\mathbf{r}_2) \psi_{\uparrow}^*(\mathbf{r}_1) ], \end{aligned}$$

so this wavefunction is odd under  $\psi_{\uparrow}^*$  which is written as  $^3g^-$ .

### E. Review of Linear Molecule Cases

Molecules with axial symmetry have orbitals of  $\sigma$ ,  $\pi$ ,  $\delta$ , etc symmetry; these orbitals carry angular momentum about the z-axis in amounts (in units of  $\hbar$ ) 0, +1 and -1, +2 and -2, +3 and -3, etc. The axial point-group symmetries of configurations formed by occupying such orbitals can be obtained by adding, in all possible ways, the angular momenta contributed by each orbital to arrive at a set of possible total angular momenta. The eigenvalue of total  $L_z = \sum_i L_z(i)$  is a valid quantum number because total  $L_z$  commutes with the Hamiltonian and with  $P_{ij}$ ; one obtains the eigenvalues of total  $L_z$  by adding the individual spin-orbitals' m eigenvalues because of the additive form of the  $L_z$  operator.  $L^2$  no longer commutes with the Hamiltonian, so it is no longer appropriate to construct N-electron functions that are eigenfunctions of  $L^2$ . Spin symmetry is treated as usual via the spin angular momentum methods described in the preceding sections and in Appendix G. For molecules with centers of symmetry (e.g., for homonuclear diatomics or ABA linear triatomics), the many-electron spin-orbital product inversion symmetry, which is equal to the product of the individual spin-orbital inversion symmetries, provides another quantum number with which the states can be labeled. Finally the  $\psi_{\uparrow}^*$  symmetry of  $\psi_{\uparrow}^*$  states can be determined by changing the m values of all orbitals in  $\psi_{\uparrow}^*$  and then determining whether the resultant function is equal to  $\psi_{\uparrow}^*$  or to  $-\psi_{\uparrow}^*$ .

If, instead of a  $\pi^2$  configuration like that treated above, one had a  $\delta^2$  configuration, the above analysis would yield  $^1\delta$ ,  $^1\delta$  and  $^3\delta$  symmetries (because the two  $\delta$  orbitals' m values could be combined as  $2 + 2$ ,  $2 - 2$ ,  $-2 + 2$ , and  $-2 - 2$ ); the wavefunctions would be identical to those given above with the  $\pi_1$  orbitals replaced by  $\delta_2$  orbitals and  $\pi_{-1}$  replaced by  $-\delta_2$ . Likewise,  $\delta^2$  gives rise to  $^1\delta$ ,  $^1\delta$ , and  $^3\delta$  symmetries.

For a  $1^1 1^1$  configuration in which two non-equivalent orbitals (i.e., orbitals which are of  $\sigma$  symmetry but which are not both members of the same degenerate set; an example would be the  $\sigma$  and  $\sigma^*$  orbitals in the  $B_2$  molecule) are occupied, the above analysis must be expanded by including determinants of the form:  $|1^1 1^1|$ ,  $|1^1 1^1|$ ,  $|1^1 1^1|$ ,  $|1^1 1^1|$ . Such determinants were excluded in the  $2^2$  case because they violated the Pauli principle (i.e., they vanish identically when  $1^1 = 1^1$ ). Determinants of the form  $|1^1 1^1|$ ,  $|1^1 1^1|$ ,  $|1^1 1^1|$ ,  $|1^1 1^1|$ ,  $|1^1 1^1|$ , and  $|1^1 1^1|$  are now distinct and must be included as must the determinants  $|1^1 1^1|$ ,  $|1^1 1^1|$ ,  $|1^1 1^1|$ ,  $|1^1 1^1|$ , and  $|1^1 1^1|$ , which are analogous to those used above. The result is that there are more possible determinants in the case of non-equivalent orbitals. However, the techniques for identifying space-spin symmetries and creating proper determinantal wavefunctions are the same as in the equivalent-orbital case.

For any  $2^2$  configuration, one finds  $1^1$ ,  $1^1$ , and  $3^1$  wavefunctions as detailed earlier; for the  $1^1 1^1$  case, one finds  $3^1$ ,  $1^1$ ,  $3^1$ ,  $1^1$ ,  $3^1$ , and  $1^1$  wavefunctions by starting with the determinants with the maximum  $M_S$  value, identifying states by their  $|M_L|$  values, and using spin angular momentum algebra and orthogonality to generate states with lower  $M_S$  and, subsequently, lower  $S$  values. Because  $L^2$  is not an operator of relevance in such cases, raising and lowering operators relating to  $\mathbf{L}$  are not used to generate states with lower  $M_L$  values. States with specific  $M_L$  values are formed by occupying the orbitals in all possible manners and simply computing  $M_L$  as the absolute value of the sum of the individual orbitals'  $m$ -values.

If a center of symmetry is present, all of the states arising from  $2^2$  are gerade; however, the states arising from  $1^1 1^1$  can be gerade if  $1^1$  and  $1^1$  are both  $g$  or both  $u$  or ungerade if  $1^1$  and  $1^1$  are of opposite inversion symmetry.

The state symmetries appropriate to the non-equivalent  $1^1 1^1$  case can, alternatively, be identified by "coupling" the spin and  $L_z$  angular momenta of two "independent" subsystems—the  $1^1$  system which gives rise to  $2^2$  symmetry (with  $M_L = 1$  and  $-1$  and  $S = 1/2$ ) and the  $1^1$  system which also give  $2^2$  symmetry. The coupling gives rise to triplet and singlet spins (whenever two full vector angular momenta  $|j, m\rangle$  and  $|j', m'\rangle$  are coupled, one can obtain total angular momentum values of  $J = j+j'$ ,  $j+j'-1$ ,  $j+j'-2$ , ...  $|j-j'|$ ; see Appendix G for details) and to  $M_L$  values of  $1+1=2$ ,  $-1-1=-2$ ,  $1-1=0$  and  $-1+1=0$  (i.e., to  $2$ ,  $0$ , and  $0$  states). The  $L_z$  angular momentum coupling is not carried out in the full vector coupling scheme used for the electron spins because, unlike the spin case where one is forming eigenfunctions of total  $S^2$  and  $S_z$ , one is only forming  $L_z$  eigenstates (i.e.,  $L^2$  is not a valid quantum label). In the case of axial angular momentum coupling, the various possible  $M_L$  values of each subsystem are added to those of the other subsystem to

arrive at the total  $M_L$  value. This angular momentum coupling approach gives the same set of symmetry labels ( $^3A_1$ ,  $^1A_1$ ,  $^3A_2$ ,  $^1A_2$ ,  $^3E_1$ , and  $^1E_1$ ) as are obtained by considering all of the determinants of the composite system as treated above.

#### IV. Non-Linear Molecule Term Symbols and Wavefunctions

##### A. Term Symbols for Non-Degenerate Point Group Symmetries

The point group symmetry labels of the individual orbitals which are occupied in any determinantal wave function can be used to determine the overall spatial symmetry of the determinant. When a point group symmetry operation is applied to a determinant, it acts on all of the electrons in the determinant; for example,  $\nu | \psi_1 \psi_2 \psi_3 | = | \nu \psi_1 \nu \psi_2 \nu \psi_3 |$ . If each of the spin-orbitals  $\psi_i$  belong to non-degenerate representations of the point group,  $\nu \psi_i$  will yield the character  $\chi_i(\nu)$  appropriate to that spin-orbital multiplying  $\psi_i$ . As a result,  $\nu | \psi_1 \psi_2 \psi_3 |$  will equal the product of the three characters (one for each spin-orbital)  $\chi_i(\nu)$  times  $| \psi_1 \psi_2 \psi_3 |$ . This gives an example of how the symmetry of a spin-orbital product (or an antisymmetrized product) is given as the direct product of the symmetries of the individual spin-orbitals in the product; the point group symmetry operator, because of its product nature, passes through or commutes with the antisymmetrizer. It should be noted that any closed-shell parts of the determinant (e.g.,  $1a_1^2 2a_1^2 1b_2^2$  in the configuration  $1a_1^2 2a_1^2 1b_2^2 1b_1^1$ ) contribute unity to the direct product because the squares of the characters of any non-degenerate point group for any group operation equals unity. Therefore, only the open-shell parts need to be considered further in the symmetry analysis. For a brief introduction to point group symmetry and the use of direct products in this context, see Appendix E.

An example will help illustrate these ideas. Consider the formaldehyde molecule  $H_2CO$  in  $C_{2v}$  symmetry. The configuration which dominates the ground-state wavefunction has doubly occupied O and C 1s orbitals, two CH bonds, a CO  $\sigma$  bond, a CO  $\pi$  bond, and two O-centered lone pairs; this configuration is described in terms of symmetry adapted orbitals as follows:  $(1a_1^2 2a_1^2 3a_1^2 1b_2^2 4a_1^2 1b_1^2 5a_1^2 2b_2^2)$  and is of  $^1A_1$  symmetry because it is entirely closed-shell (note that lower case letters are used to denote the symmetries of orbitals and capital letters are used for many-electron functions' symmetries).

The lowest-lying  $n \Rightarrow ^*$  states correspond to a configuration (only those orbitals whose occupancies differ from those of the ground state are listed) of the form  $2b_2^1 2b_1^1$ , which gives rise to  $^1A_2$  and  $^3A_2$  wavefunctions (the direct product of the open-shell spin

orbitals is used to obtain the symmetry of the product wavefunction:  $A_2 = b_1 \times b_2$ ). The  $\Rightarrow$  \* excited configuration  $1b_1^2 2b_1^1$  gives  $^1A_1$  and  $^3A_1$  states because  $b_1 \times b_1 = A_1$ .

The only angular momentum coupling that occurs in non-linear molecules involves the electron spin angular momenta, which are treated in a vector coupling manner. For example, in the lowest-energy state of formaldehyde, the orbitals are occupied in the configuration  $1a_1^2 2a_1^2 3a_1^2 1b_2^2 4a_1^2 1b_1^2 5a_1^2 2b_2^2$ . This configuration has only a single entry in its "box". Its highest  $M_S$  value is  $M_S = 0$ , so there is a singlet  $S = 0$  state. The spatial symmetry of this singlet state is totally symmetric  $A_1$  because this is a closed-shell configuration.

The lowest-energy n \* excited configuration of formaldehyde has a  $1a_1^2 2a_1^2 3a_1^2 1b_2^2 4a_1^2 1b_1^2 5a_1^2 2b_2^1 2b_1^1$  configuration, which has a total of four entries in its "box" diagram:

$$\begin{array}{ll} M_S = 1 & |2b_2^1 \ 2b_1^1 \rangle, \\ M_S = 0 & |2b_2^1 \ 2b_1^1 \rangle, \\ M_S = 0 & |2b_2^1 \ 2b_1^1 \rangle, \\ M_S = -1 & |2b_2^1 \ 2b_1^1 \rangle. \end{array}$$

The highest  $M_S$  value is  $M_S = 1$ , so there is an  $S = 1$  state. After deleting one entry each with  $M_S = 1, 0$ , and  $-1$ , there is one entry left with  $M_S = 0$ . Thus, there is an  $S = 0$  state also.

As illustrated above, the spatial symmetries of these four  $S = 1$  and  $S = 0$  states are obtained by forming the direct product of the "open-shell" orbitals that appear in this configuration:  $b_2 \times b_1 = A_2$ .

All four states have this spatial symmetry. In summary, the above configuration yields  $^3A_2$  and  $^1A_2$  term symbols. The  $^1 * 1$  configuration  $1a_1^2 2a_1^2 3a_1^2 1b_2^2 4a_1^2 1b_1^1 5a_1^2 2b_2^2 2b_1^1$  produces  $^3A_1$  and  $^1A_1$  term symbols (because  $b_1 \times b_1 = A_1$ ).

## B. Wavefunctions for Non-Degenerate Non-Linear Point Molecules

The techniques used earlier for linear molecules extend easily to non-linear molecules. One begins with those states that can be straightforwardly identified as unique entries within the box diagram. For polyatomic molecules with no degenerate representations, the spatial symmetry of each box entry is identical and is given as the direct product of the open-shell orbitals. For the formaldehyde example considered earlier, the spatial symmetries of the n \* and \* states were  $A_2$  and  $A_1$ , respectively.

After the unique entries of the box have been identified, one uses  $S_{\pm}$  operations to find the other functions. For example, the wavefunctions of the  ${}^3A_2$  and  ${}^1A_2$  states of the  $n^* 1a_1^2 2a_1^2 3a_1^2 1b_2^2 4a_1^2 1b_1^2 5a_1^2 2b_2^1 2b_1^1$  configuration of formaldehyde are formed by first identifying the  $M_S = \pm 1$  components of the  $S = 1$  state as  $|2b_2 \ 2b_1 \ |$  and  $|2b_2 \ -2b_1 \ |$  (all of the closed-shell components of the determinants are not explicitly given). Then, applying  $S_{-}$  to the  $M_S = 1$  state, one obtains the  $M_S = 0$  component  $(1/2)^{1/2} [|2b_2 \ 2b_1 \ | + |2b_2 \ -2b_1 \ |]$ . The singlet state is then constructed as the combination of the two determinants appearing in the  $S = 1, M_S = 0$  state that is orthogonal to this triplet state. The result is  $(1/2)^{1/2} [|2b_2 \ 2b_1 \ | - |2b_2 \ -2b_1 \ |]$ .

The results of applying these rules to the  $n^*$  and  ${}^*$  states are as follows:

$${}^3A_2 (M_S = 1) = |1a_1 \ 1a_1 \ 2a_1 \ 2a_1 \ 3a_1 \ 3a_1 \ 1b_2 \ 1b_2 \ 4a_1 \ 4a_1 \ 1 \ 1 \ 1b_1$$

$$5a_1 \ 5a_1 \ 2b_2 \ 2b_1 \ |,$$

$${}^3A_2 (M_S = 0) = 1/\sqrt{2} [|2b_2 \ 2b_1 \ | + |2b_2 \ -2b_1 \ |],$$

$${}^3A_2 (M_S = -1) = |2b_2 \ -2b_1 \ |,$$

$${}^1A_2 = 1/\sqrt{2} [|2b_2 \ 2b_1 \ | - |2b_2 \ -2b_1 \ |].$$

The lowest  ${}^*$  states of triplet and singlet spin involve the following:

$${}^3A_1 (M_S = 1) = |1b_1 \ 2b_1 \ |,$$

$${}^1A_1 = 1/\sqrt{2} [|1b_1 \ 2b_1 \ | - |1b_1 \ -2b_1 \ |].$$

In summary, forming spatial- and spin- adapted determinantal functions for molecules whose point groups have no degenerate representations is straightforward. The direct product of all of the open-shell spin orbitals gives the point-group symmetry of the determinant. The spin symmetry is handled using the spin angular momentum methods introduced and illustrated earlier.

### C. Extension to Degenerate Representations for Non-Linear Molecules

Point groups in which degenerate orbital symmetries appear can be treated in like fashion but require more analysis because a symmetry operation  $R$  acting on a degenerate

orbital generally yields a linear combination of the degenerate orbitals rather than a multiple of the original orbital (i.e.,  $R \psi_i = \psi_i(R) \psi_i$  is no longer valid). For example, when a pair of degenerate orbitals (denoted  $e_1$  and  $e_2$ ) are involved, one has

$$R \psi_i = \sum_j R_{ij} \psi_j,$$

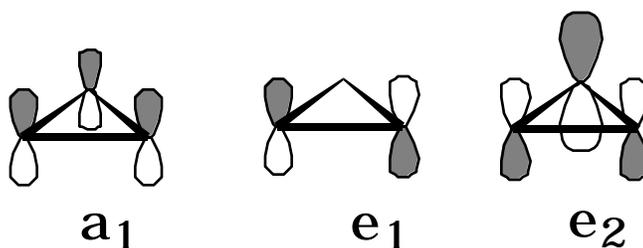
where  $R_{ij}$  is the 2x2 matrix representation of the effect of  $R$  on the two orbitals. The effect of  $R$  on a product of orbitals can be expressed as:

$$R \psi_i \psi_j = \sum_{k,l} R_{ik} R_{jl} \psi_k \psi_l.$$

The matrix  $R_{ij,kl} = R_{ik} R_{jl}$  represents the effect of  $R$  on the orbital products in the same way  $R_{ik}$  represents the effect of  $R$  on the orbitals. One says that the orbital products also form a basis for a representation of the point group. The character (i.e., the trace) of the representation matrix  $R_{ij,kl}$  appropriate to the orbital product basis is seen to equal the product of the characters of the matrix  $R_{ik}$  appropriate to the orbital basis:  $\chi^2(R) = \chi_e(R) \chi_e(R)$ , which is, of course, why the term "direct product" is used to describe this relationship.

For point groups which contain no degenerate representations, the direct product of one symmetry with another is equal to a unique symmetry; that is, the characters  $\chi_a(R) \chi_b(R)$  obtained as  $\chi_a(R) \chi_b(R)$  belong to a pure symmetry and can be immediately identified in a point-group character table. However, for point groups in which degenerate representations occur, such is not the case. The direct product characters will, in general, not correspond to the characters of a single representation; they will contain contributions from more than one representation and these contributions will have to be sorted out using the tools provided below.

A concrete example will help clarify these concepts. In  $C_{3v}$  symmetry, the orbitals of the cyclopropenyl anion transform according to  $a_1$  and  $e$  symmetries



and can be expressed as LCAO-MO's in terms of the individual  $p_i$  orbitals as follows:

$$a_1 = 1/\sqrt{3} [p_1 + p_2 + p_3], e_1 = 1/\sqrt{2} [p_1 - p_3],$$

and

$$e_2 = 1/\sqrt{6} [2p_2 - p_1 - p_3].$$

For the anion's lowest energy configuration, the orbital occupancy  $a_1^2 e^2$  must be considered, and hence the spatial and spin symmetries arising from the  $e^2$  configuration are of interest. The character table shown below

$C_{3v}$	E	$3 \sigma_v$	$2 C_3$
$a_1$	1	1	1
$a_2$	1	-1	1
e	2	0	-1

allows one to compute the characters appropriate to the direct product ( $e \times e$ ) as  $\chi(E) = 2 \times 2 = 4$ ,  $\chi(\sigma_v) = 0 \times 0 = 0$ ,  $\chi(C_3) = (-1) \times (-1) = 1$ .

This reducible representation (the occupancy of two e orbitals in the anion gives rise to more than one state, so the direct product  $e \times e$  contains more than one symmetry component) can be decomposed into pure symmetry components (labels are used to denote the irreducible symmetries) by using the decomposition formula given in Appendix E:

$$n(\Gamma) = 1/g \sum_R \chi(\Gamma) \chi(R) \chi(R).$$

Here  $g$  is the order of the group (the number of symmetry operations in the group- 6 in this case) and  $\chi(R)$  is the character for the particular symmetry whose component in the direct product is being calculated.

For the case given above, one finds  $n(a_1) = 1$ ,  $n(a_2) = 1$ , and  $n(e) = 1$ ; so within the configuration  $e^2$  there is one  $A_1$  wavefunction, one  $A_2$  wavefunction and a pair of  $E$  wavefunctions (where the symmetry labels now refer to the symmetries of the determinantal wavefunctions). This analysis tells one how many different wavefunctions of various spatial symmetries are contained in a configuration in which degenerate orbitals are fractionally occupied. Considerations of spin symmetry and the construction of proper determinantal wavefunctions, as developed earlier in this Section, still need to be applied to each spatial symmetry case.

To generate the proper  $A_1$ ,  $A_2$ , and  $E$  wavefunctions of singlet and triplet spin symmetry (thus far, it is not clear which spin can arise for each of the three above spatial symmetries; however, only singlet and triplet spin functions can arise for this two-electron example), one can apply the following (un-normalized) symmetry projection operators (see Appendix E where these projectors are introduced) to all determinantal wavefunctions arising from the  $e^2$  configuration:

$$P = \frac{1}{g} \sum_R \chi(R) R .$$

Here,  $\chi(R)$  is the character belonging to symmetry for the symmetry operation  $R$ . Applying this projector to a determinantal function of the form  $|i j|$  generates a sum of determinants with coefficients determined by the matrix representations  $R_{ik}$ :

$$P |i j| = \frac{1}{g} \sum_k \chi(R) R_{ik} R_{jl} |k l| .$$

For example, in the  $e^2$  case, one can apply the projector to the determinant with the maximum  $M_s$  value to obtain

$$\begin{aligned} P |e_1 e_2| &= \frac{1}{g} \sum_R \chi(R) [R_{11}R_{22} |e_1 e_2| + R_{12}R_{21} |e_2 e_1|] \\ &= \frac{1}{g} \sum_R \chi(R) [R_{11}R_{22} - R_{12}R_{21}] |e_1 e_2| , \end{aligned}$$

or to the other two members of this triplet manifold, thereby obtaining

$$P |e_1 e_2| = \frac{1}{g} \sum_R \chi(R) [R_{11}R_{22} - R_{12}R_{21}] |e_1 e_2|$$

and

$$P \frac{1}{2} [|e_1 e_2 | + |e_1 e_2 |] = R \quad (R) [R_{11}R_{22} - R_{12}R_{21}]$$

$$\frac{1}{2} [|e_1 e_2 | + |e_1 e_2 |].$$

The other (singlet) determinants can be symmetry analyzed in like fashion and result in the following:

$$P |e_1 e_1 | = R \quad (R) \{ R_{11}R_{11}|e_1 e_1 | + R_{12}R_{12}|e_2 e_2 | + R_{11}R_{12} [|e_1 e_2 | - |e_1 e_2 |] \},$$

$$P |e_2 e_2 | = R \quad (R) \{ R_{22}R_{22}|e_2 e_2 | + R_{21}R_{21}|e_1 e_1 | + R_{22}R_{21} [|e_2 e_1 | - |e_2 e_1 |] \},$$

and

$$P \frac{1}{2} [|e_1 e_2 | - |e_1 e_2 |] = R \quad (R) \{ 2 R_{11}R_{21}|e_1 e_1 | + 2 R_{22}R_{12}|e_2 e_2 | + (R_{11}R_{22} + R_{12}R_{21}) [|e_1 e_2 | - |e_1 e_2 |] \}.$$

To make further progress, one needs to evaluate the  $R_{ik}$  matrix elements for the particular orbitals given above and to then use these explicit values in the above equations. The matrix representations for the two e orbitals can easily be formed and are as follows:

$$\begin{matrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} & \begin{pmatrix} -1/2 & 3/2 \\ 3/2 & 1/2 \end{pmatrix} \\ E & v & 'v \end{matrix}$$

$$\begin{matrix} \begin{pmatrix} -1/2 & -3/2 \\ -3/2 & 1/2 \end{pmatrix} & \begin{pmatrix} -1/2 & 3/2 \\ -3/2 & -1/2 \end{pmatrix} & \begin{pmatrix} -1/2 & -3/2 \\ 3/2 & -1/2 \end{pmatrix} \\ ''v & C_3 & C'_3 \end{matrix}$$

Turning first to the three triplet functions, one notes that the effect of the symmetry projector acting on each of these three was the following multiple of the respective function:

$$R \quad (R) [R_{11}R_{22}$$

$-R_{12}R_{21}]$ . Evaluating this sum for each of the three symmetries  $= A_1, A_2,$  and  $E$ , one obtains values of 0, 2, and 0, respectively. That is, the projection of each of the original triplet determinants gives zero except for  $A_2$  symmetry. This allows one to conclude that there are no  $A_1$  or  $E$  triplet functions in this case; the triplet functions are of pure  ${}^3A_2$  symmetry.

Using the explicit values for  $R_{ik}$  matrix elements in the expressions given above for the projection of each of the singlet determinantal functions, one finds only the following non-vanishing contributions:

(i) For  $A_1$  symmetry-  $P|e_1 e_1| = 3[|e_1 e_1| + |e_2 e_2|] = P|e_2 e_2|,$

(ii) For  $A_2$  symmetry- all projections vanish,

(iii) For  $E$  symmetry-  $P|e_1 e_1| = 3/2 [|e_1 e_1| - |e_2 e_2|] = -P|e_2 e_2|$

and  $P1/2[|e_1 e_2| - |e_1 e_2|] = 3/2[|e_1 e_2| - |e_1 e_2|].$

Remembering that the projection process does not lead to a normalized function, although it does generate a function of pure symmetry, one can finally write down the normalized symmetry-adapted singlet functions as:

(i)  ${}^1A_1 = 1/2[|e_1 e_1| + |e_2 e_2|],$

(ii)  ${}^1E = \{ 1/2[|e_1 e_1| - |e_2 e_2|], \text{ and } 1/2[|e_1 e_2| - |e_1 e_2|] \}.$

The triplet functions given above are:

(iii)  ${}^3A_2 = \{ |e_1 e_2|, 1/2[|e_1 e_2| + |e_1 e_2|], \text{ and } |e_1 e_2| \}.$

In summary, whenever one has partially occupied degenerate orbitals, the characters corresponding to the direct product of the open-shell orbitals (as always, closed-shells contribute nothing to the symmetry analysis and can be ignored, although their presence must, of course, be specified when one finally writes down complete symmetry-adapted wavefunctions) must be reduced to identify the spatial symmetry components of the configuration. Given knowledge of the various spatial symmetries, one must then form determinantal wavefunctions of each possible space and spin symmetry. In doing so, one

starts with the maximum  $M_s$  function and uses spin angular momentum algebra and orthogonality to form proper spin eigenfunctions, and then employs point group projection operators (which require the formation of the  $R_{jk}$  representation matrices). Antisymmetry, as embodied in the determinants, causes some space-spin symmetry combinations to vanish (e.g.,  $^3A_1$  and  $^3E$  and  $^1A_2$  in the above example) thereby enforcing the Pauli principle. This procedure, although tedious, is guaranteed to generate all space- and spin-symmetry adapted determinants for any configuration involving degenerate orbitals. The results of certain such combined spin and spatial symmetry analyses have been tabulated. For example, in Appendix 11 of Atkins such information is given in the form of tables of direct products for several common point groups.

For cases in which one has a non-equivalent set of degenerate orbitals (e.g., for a configuration whose open-shell part is  $e^1e'^1$ ), the procedure is exactly the same as above except that the determination of the possible space-spin symmetries is more straightforward. In this case, singlet and triplet functions exist for all three space symmetries-  $A_1$ ,  $A_2$ , and  $E$ , because the Pauli principle does not exclude determinants of the form  $|e_1 e'_1|$  or  $|e_2 e'_2|$ , whereas the equivalent determinants ( $|e_1 e_1|$  or  $|e_2 e_2|$ ) vanish when the degenerate orbitals belong to the same set (in which case, one says that the orbitals are equivalent).

For all point, axial rotation, and full rotation group symmetries, this observation holds: if the orbitals are equivalent, certain space-spin symmetry combinations will vanish due to antisymmetry; if the orbitals are not equivalent, all space-spin symmetry combinations consistent with the content of the direct product analysis are possible. In either case, one must proceed through the construction of determinantal wavefunctions as outlined above.

## V. Summary

The ability to identify all term symbols and to construct all determinantal wavefunctions that arise from a given electronic configuration is important. This knowledge allows one to understand and predict the changes (i.e., physical couplings due to external fields or due to collisions with other species and chemical couplings due to interactions with orbitals and electrons of a 'ligand' or another species) that each state experiences when the atom or molecule is subjected to some interaction. Such understanding plays central roles in interpreting the results of experiments in spectroscopy and chemical reaction dynamics.

The essence of this analysis involves being able to write each wavefunction as a combination of determinants each of which involves occupancy of particular spin-orbitals. Because different spin-orbitals interact differently with, for example, a colliding molecule, the various determinants will interact differently. These differences thus give rise to different interaction potential energy surfaces.

For example, the Carbon-atom  ${}^3P(M_L=1, M_S=0) = 2^{-1/2} [|p_1 p_0 | + |p_1 p_0 |]$  and  ${}^3P(M_L=0, M_S=0) = 2^{-1/2} [|p_1 p_{-1} | + |p_1 p_{-1} |]$  states interact quite differently in a collision with a closed-shell Ne atom. The  $M_L = 1$  state's two determinants both have an electron in an orbital directed toward the Ne atom (the  $2p_0$  orbital) as well as an electron in an orbital directed perpendicular to the C-Ne internuclear axis (the  $2p_1$  orbital); the  $M_L = 0$  state's two determinants have both electrons in orbitals directed perpendicular to the C-Ne axis. Because Ne is a closed-shell species, any electron density directed toward it will produce a "repulsive" antibonding interaction. As a result, we expect the  $M_L = 1$  state to undergo a more repulsive interaction with the Ne atom than the  $M_L = 0$  state.

Although one may be tempted to 'guess' how the various  ${}^3P(M_L)$  states interact with a Ne atom by making an analogy between the three  $M_L$  states within the  ${}^3P$  level and the three orbitals that comprise a set of p-orbitals, such analogies are not generally valid. The wavefunctions that correspond to term symbols are N-electron functions; they describe how N spin-orbitals are occupied and, therefore, how N spin-orbitals will be affected by interaction with an approaching 'ligand' such as a Ne atom. The net effect of the ligand will depend on the occupancy of all N spin-orbitals.

To illustrate this point, consider how the  ${}^1S$  state of Carbon would be expected to interact with an approaching Ne atom. This term symbol's wavefunction  ${}^1S = 3^{-1/2} [|p_0 p_0 | - |p_{-1} p_1 | - |p_1 p_{-1} |]$  contains three determinants, each with a  $1/3$  probability factor. The first,  $|p_0 p_0 |$ , produces a repulsive interaction with the closed-shell Ne; the second and third,  $|p_{-1} p_1 |$  and  $|p_1 p_{-1} |$ , produce attractive interactions because they allow the Carbon's vacant  $p_0$  orbital to serve in a Lewis acid capacity and accept electron density from Ne. The net effect is likely to be an attractive interaction because of the equal weighting of these three determinants in the  ${}^1S$  wavefunction. This result could not have been 'guessed' by making making analogy with how an s-orbital interacts with a Ne atom; the  ${}^1S$  state and an s-orbital are distinctly different in this respect.

## Chapter 11

*One Must be Able to Evaluate the Matrix Elements Among Properly Symmetry Adapted N-Electron Configuration Functions for Any Operator, the Electronic Hamiltonian in Particular. The Slater-Condon Rules Provide this Capability*

### I. CSFs Are Used to Express the Full N-Electron Wavefunction

It has been demonstrated that a given electronic configuration can yield several space- and spin- adapted determinantal wavefunctions; such functions are referred to as configuration state functions (CSFs). These CSF wavefunctions are not the exact eigenfunctions of the many-electron Hamiltonian, H; they are simply functions which possess the space, spin, and permutational symmetry of the exact eigenstates. As such, they comprise an acceptable set of functions to use in, for example, a linear variational treatment of the true states.

In such variational treatments of electronic structure, the N-electron wavefunction is expanded as a sum over all CSFs that possess the desired spatial and spin symmetry:

$$\Psi = \sum_J C_J \Phi_J$$

Here, the  $\Phi_J$  represent the CSFs that are of the correct symmetry, and the  $C_J$  are their expansion coefficients to be determined in the variational calculation. If the spin-orbitals used to form the determinants, that in turn form the CSFs  $\{\Phi_J\}$ , are orthonormal one-electron functions (i.e.,  $\langle \phi_k | \phi_j \rangle = \delta_{k,j}$ ), then the CSFs can be shown to be orthonormal functions of N electrons

$$\langle \Phi_J | \Phi_K \rangle = \delta_{J,K}$$

In fact, the Slater determinants themselves also are orthonormal functions of N electrons whenever orthonormal spin-orbitals are used to form the determinants.

The above expansion of the full N-electron wavefunction is termed a "configuration-interaction" (CI) expansion. It is, in principle, a mathematically rigorous approach to expressing  $\Psi$  because the set of all determinants that can be formed from a complete set of spin-orbitals can be shown to be complete. In practice, one is limited to the number of orbitals that can be used and in the number of CSFs that can be included in the CI expansion. Nevertheless, the CI expansion method forms the basis of the most commonly used techniques in quantum chemistry.

In general, the optimal variational (or perturbative) wavefunction for any (i.e., the ground or excited) state will include contributions from spin-and space-symmetry adapted determinants derived from all possible configurations. For example, although the determinant with  $L = 1, S = 1, M_L = 1, M_S = 1$  arising from the  $1s^2 2s^2 2p^2$  configuration may contribute strongly to the true ground electronic state of the Carbon atom, there will be contributions from all configurations which can provide these  $L, S, M_L,$  and  $M_S$  values (e.g., the  $1s^2 2s^2 2p^1 3p^1$  and  $2s^2 2p^4$  configurations will also contribute, although the  $1s^2 2s^2 2p^1 3s^1$  and  $1s^2 2s^1 2p^2 3p^1$  will not because the latter two configurations are odd under inversion symmetry whereas the state under study is even).

The mixing of CSFs from many configurations to produce an optimal description of the true electronic states is referred to as configuration interaction (CI). Strong CI (i.e., mixing of CSFs with large amplitudes appearing for more than one dominant CSF) can occur, for example, when two CSFs from different electronic configurations have nearly the same Hamiltonian expectation value. For example, the  $1s^2 2s^2$  and  $1s^2 2p^2 \ ^1S$  configurations of Be and the analogous  $ns^2$  and  $np^2$  configurations of all alkaline earth atoms are close in energy because the  $ns$ - $np$  orbital energy splitting is small for these elements; the  $\sigma^2$  and  $\pi^2$  configurations of ethylene become equal in energy, and thus undergo strong CI mixing, as the  $CC$  bond is twisted by  $90^\circ$  in which case the  $\sigma$  and  $\pi^*$  orbitals become degenerate.

Within a variational treatment, the relative contributions of the spin-and space-symmetry adapted CSFs are determined by solving a secular problem for the eigenvalues ( $E_i$ ) and eigenvectors ( $C_i$ ) of the matrix representation  $H$  of the full many-electron Hamiltonian  $\mathbf{H}$  within this CSF basis:

$$\sum_L H_{K,L} C_{i,L} = E_i C_{i,K}.$$

The eigenvalue  $E_i$  gives the variational estimate for the energy of the  $i^{\text{th}}$  state, and the entries in the corresponding eigenvector  $C_{i,K}$  give the contribution of the  $K^{\text{th}}$  CSF to the  $i^{\text{th}}$  wavefunction  $\psi_i$  in the sense that

$$\psi_i = \sum_K C_{i,K} \psi_K,$$

where  $\psi_K$  is the  $K^{\text{th}}$  CSF.

II. The Slater-Condon Rules Give Expressions for the Operator Matrix Elements Among the CSFs

To form the  $H_{K,L}$  matrix, one uses the so-called Slater-Condon rules which express all non-vanishing determinantal matrix elements involving either one- or two- electron operators (one-electron operators are additive and appear as

$$F = \sum_i f(i);$$

two-electron operators are pairwise additive and appear as

$$G = \sum_{ij} g(i,j).$$

Because the CSFs are simple linear combinations of determinants with coefficients determined by space and spin symmetry, the  $H_{I,J}$  matrix in terms of determinants can be used to generate the  $H_{K,L}$  matrix over CSFs.

The Slater-Condon rules give the matrix elements between two determinants

$$|> = | 1 2 3 \dots N |$$

and

$$| '> = | '1 '2 '3 \dots 'N |$$

for any quantum mechanical operator that is a sum of one- and two- electron operators ( $F + G$ ). It expresses these matrix elements in terms of one- and two-electron integrals involving the spin-orbitals that appear in  $|>$  and  $| '>$  and the operators  $f$  and  $g$ .

As a first step in applying these rules, one must examine  $|>$  and  $| '>$  and determine by how many (if any) spin-orbitals  $|>$  and  $| '>$  differ. In so doing, one may have to reorder the spin-orbitals in one of the determinants to achieve maximal coincidence with those in the other determinant; it is essential to keep track of the number of permutations ( $N_p$ ) that one makes in achieving maximal coincidence. The results of the Slater-Condon rules given below are then multiplied by  $(-1)^{N_p}$  to obtain the matrix elements between the original  $|>$  and  $| '>$ . The final result does not depend on whether one chooses to permute  $|>$  or  $| '>$ .

Once maximal coincidence has been achieved, the Slater-Condon (SC) rules provide the following prescriptions for evaluating the matrix elements of any operator  $F + G$  containing a one-electron part  $F = \sum_i f(i)$  and a two-electron part  $G = \sum_{ij} g(i,j)$  (the Hamiltonian is, of course, a specific example of such an operator; the electric dipole

operator  $\sum_i e \mathbf{r}_i$  and the electronic kinetic energy  $-\hbar^2/2m_e \sum_i \nabla_i^2$  are examples of one-electron operators (for which one takes  $g = 0$ ); the electron-electron coulomb interaction  $\sum_{i>j} e^2/r_{ij}$  is a two-electron operator (for which one takes  $f = 0$ ):

## The Slater-Condon Rules

(i) If  $|\psi\rangle$  and  $|\psi'\rangle$  are identical, then

$$\langle \psi | F + G | \psi \rangle = \sum_i \langle i | f | i \rangle + \sum_{i < j} [ \langle i j | g | i j \rangle - \langle i j | g | j i \rangle ],$$

where the sums over  $i$  and  $j$  run over all spin-orbitals in  $|\psi\rangle$ ;

(ii) If  $|\psi\rangle$  and  $|\psi'\rangle$  differ by a single spin-orbital mismatch ( $p \rightarrow p'$ ),

$$\langle \psi | F + G | \psi' \rangle = \langle p | f | p' \rangle + \sum_j [ \langle p j | g | p' j \rangle - \langle p j | g | j p' \rangle ],$$

where the sum over  $j$  runs over all spin-orbitals in  $|\psi\rangle$  except  $p$ ;

(iii) If  $|\psi\rangle$  and  $|\psi'\rangle$  differ by two spin-orbitals ( $p \rightarrow p'$  and  $q \rightarrow q'$ ),

$$\langle \psi | F + G | \psi' \rangle = \langle p q | g | p' q' \rangle - \langle p q | g | q' p' \rangle$$

(note that the  $F$  contribution vanishes in this case);

(iv) If  $|\psi\rangle$  and  $|\psi'\rangle$  differ by three or more spin orbitals, then

$$\langle \psi | F + G | \psi' \rangle = 0;$$

(v) For the identity operator  $I$ , the matrix elements  $\langle \psi | I | \psi' \rangle = 0$  if  $|\psi\rangle$  and  $|\psi'\rangle$  differ by one or more spin-orbitals (i.e., the Slater determinants are orthonormal if their spin-orbitals are).

Recall that each of these results is subject to multiplication by a factor of  $(-1)^{N_p}$  to account for possible ordering differences in the spin-orbitals in  $|\psi\rangle$  and  $|\psi'\rangle$ .

In these expressions,

$$\langle i | f | j \rangle$$

is used to denote the one-electron integral

$$\int \psi_i^*(r) f(r) \psi_j(r) dr$$

and

$$\langle i j | g | k l \rangle \text{ (or in short hand notation } \langle i j | k l \rangle)$$

represents the two-electron integral

$$\int \psi_i^*(r) \psi_j(r') g(r,r') \psi_k(r) \psi_l(r') dr dr'$$

The notation  $\langle ij | kl \rangle$  introduced above gives the two-electron integrals for the  $g(r,r')$  operator in the so-called Dirac notation, in which the  $i$  and  $k$  indices label the spin-orbitals that refer to the coordinates  $r$  and the  $j$  and  $l$  indices label the spin-orbitals referring to coordinates  $r'$ . The  $r$  and  $r'$  denote  $r, \theta, \phi$  and  $r', \theta', \phi'$  (with  $\theta$  and  $\theta'$  being the polar or spin functions). The fact that  $r$  and  $r'$  are integrated and hence represent 'dummy' variables introduces index permutational symmetry into this list of integrals. For example,

$$\langle ij | kl \rangle = \langle ji | lk \rangle = \langle kl | ij \rangle^* = \langle lk | ji \rangle^*;$$

the final two equivalences are results of the Hermitian nature of  $g(r,r')$ .

It is also common to represent these same two-electron integrals in a notation referred to as Mulliken notation in which:

$$\int \psi_i^*(r) \psi_j(r') g(r,r') \psi_k(r) \psi_l(r') dr dr' = (ik | jl).$$

Here, the indices  $i$  and  $k$ , which label the spin-orbital having variables  $r$  are grouped together, and  $j$  and  $l$ , which label spin-orbitals referring to the  $r'$  variables appear together. The above permutational symmetries, when expressed in terms of the Mulliken integral list read:

$$(ik | jl) = (jl | ik) = (ki | lj)^* = (lj | ki)^*.$$

If the operators  $f$  and  $g$  do not contain any electron spin operators, then the spin integrations implicit in these integrals (all of the  $\psi_i$  are spin-orbitals, so each  $\psi_i$  is accompanied by an  $\alpha$  or  $\beta$  spin function and each  $\psi_i^*$  involves the adjoint of one of the  $\alpha$  or  $\beta$  spin functions) can be carried out as  $\langle \alpha | \alpha \rangle = 1$ ,  $\langle \alpha | \beta \rangle = 0$ ,  $\langle \beta | \alpha \rangle = 0$ ,  $\langle \beta | \beta \rangle = 1$ , thereby yielding integrals over spatial orbitals. These spin integration results follow immediately from the general properties of angular momentum eigenfunctions detailed in Appendix G; in particular, because  $\alpha$  and  $\beta$  are eigenfunctions of  $S_z$  with different eigenvalues, they must be orthogonal  $\langle \alpha | \beta \rangle = \langle \beta | \alpha \rangle = 0$ .

The essential results of the Slater-Condon rules are:

1. The full  $N!$  terms that arise in the  $N$ -electron Slater determinants do not have to be treated explicitly, nor do the  $N!(N! + 1)/2$  Hamiltonian matrix elements among the  $N!$  terms of one Slater determinant and the  $N!$  terms of the same or another Slater determinant.
2. All such matrix elements, for any one- and/or two-electron operator can be expressed in terms of one- or two-electron integrals over the spin-orbitals that appear in the determinants.
3. The integrals over orbitals are three or six dimensional integrals, regardless of how many electrons  $N$  there are.
4. These integrals over mo's can, through the LCAO-MO expansion, ultimately be expressed in terms of one- and two-electron integrals over the primitive atomic orbitals. It is only these ao-based integrals that can be evaluated explicitly (on high speed computers for all but the smallest systems).

### III. Examples of Applying the Slater-Condon Rules

It is wise to gain some experience using the SC rules, so let us consider a few illustrative example problems.

1. What is the contribution to the total energy of the  $^3P$  level of Carbon made by the two  $2p$  orbitals alone? Of course, the two  $1s$  and two  $2s$  spin-orbitals contribute to the total energy, but we artificially ignore all such contributions in this example to simplify the problem.

Because all nine of the  $^3P$  states have the same energy, we can calculate the energy of any one of them; it is therefore prudent to choose an "easy" one

$$^3P(M_L=1, M_S=1) = |p_1 \ p_0 \ |.$$

The energy of this state is  $\langle p_1 \ p_0 \ | H \ | p_1 \ p_0 \ \rangle$ . The SC rules tell us this equals:

$$I_{2p_1} + I_{2p_0} + \langle 2p_1 2p_0 | 2p_1 2p_0 \rangle - \langle 2p_1 2p_0 | 2p_0 2p_1 \rangle,$$

where the short hand notation  $I_j = \langle j | f | j \rangle$  is introduced.

If the contributions from the two  $1s$  and two  $2s$  spin-orbitals are now taken into account, one obtains a total energy that also contains  $2I_{1s} + 2I_{2s} + \langle 1s 1s | 1s 1s \rangle + 4\langle 1s 2s | 1s 2s \rangle - 2\langle 1s 2s | 2s 1s \rangle + \langle 2s 2s | 2s 2s \rangle + 2\langle 1s 2p_1 | 1s 2p_1 \rangle - \langle 1s 2p_1 | 2p_1 1s \rangle + 2\langle 1s 2p_0 | 1s 2p_0 \rangle - \langle 1s 2p_0 | 2p_0 1s \rangle + 2\langle 2s 2p_1 | 2s 2p_1 \rangle - \langle 2s 2p_1 | 2p_1 2s \rangle + 2\langle 2s 2p_0 | 2s 2p_0 \rangle - \langle 2s 2p_0 | 2p_0 2s \rangle$ .

2. Is the energy of another  $^3P$  state equal to the above state's energy? Of course, but it may prove informative to prove this.

Consider the  $M_S=0, M_L=1$  state whose energy is:

$$\begin{aligned}
 & 2^{-1/2} \langle [p_1 p_0 | + | p_1 p_0 | ] | H | [p_1 p_0 | + | p_1 p_0 | ] \rangle 2^{-1/2} \\
 &= 1/2 \{ I_{2p_1} + I_{2p_0} + \langle 2p_1 2p_0 | 2p_1 2p_0 \rangle + I_{2p_1} + I_{2p_0} + \langle 2p_1 2p_0 | 2p_1 2p_0 \rangle \} \\
 &+ 1/2 \{ - \langle 2p_1 2p_0 | 2p_0 2p_1 \rangle - \langle 2p_1 2p_0 | 2p_0 2p_1 \rangle \} \\
 &= I_{2p_1} + I_{2p_0} + \langle 2p_1 2p_0 | 2p_1 2p_0 \rangle - \langle 2p_1 2p_0 | 2p_0 2p_1 \rangle.
 \end{aligned}$$

Which is, indeed, the same as the other  $^3P$  energy obtained above.

3. What energy would the singlet state  $2^{-1/2} \langle [p_1 p_0 | - | p_1 p_0 | ]$  have?

The  $^3P M_S=0$  example can be used (changing the sign on the two determinants) to give

$$E = I_{2p_1} + I_{2p_0} + \langle 2p_1 2p_0 | 2p_1 2p_0 \rangle + \langle 2p_1 2p_0 | 2p_0 2p_1 \rangle.$$

Note, this is the  $M_L=1$  component of the  $^1D$  state; it is, of course, not a  $^1P$  state because no such state exists for two equivalent p electrons.

4. What is the CI matrix element coupling  $|1s^2 2s^2|$  and  $|1s^2 3s^2|$ ?

These two determinants differ by two spin-orbitals, so

$$\langle |1s \uparrow 1s \downarrow 2s \uparrow 2s \downarrow | H | 1s \uparrow 1s \downarrow 3s \uparrow 3s \downarrow \rangle = \langle 2s 2s | 3s 3s \rangle = \langle 2s 3s | 3s 2s \rangle$$

(note, this is an exchange-type integral).

5. What is the CI matrix element coupling  $|$  and  $|$ ?

These two determinants differ by two spin-orbitals, so

$$\langle \psi | H | \psi \rangle = \langle \psi | \psi \rangle = \langle \psi | \psi \rangle$$

(note, again this is an exchange-type integral).

6. What is the Hamiltonian matrix element coupling  $|\psi_1\rangle$  and  $2^{-1/2} [|\psi_1^* \psi_2\rangle - |\psi_2^* \psi_1\rangle]$ ?

The first determinant differs from the  $\psi_2$  determinant by one spin-orbital, as does the second (after it is placed into maximal coincidence by making one permutation), so

$$\begin{aligned} & \langle \psi_1 | H | 2^{-1/2} [|\psi_1^* \psi_2\rangle - |\psi_2^* \psi_1\rangle] \rangle \\ &= 2^{-1/2} [\langle \psi_1 | H | \psi_1^* \psi_2 \rangle + \langle \psi_1 | H | \psi_2^* \psi_1 \rangle] - (-1) 2^{-1/2} [\langle \psi_1 | H | \psi_1^* \psi_2 \rangle + \langle \psi_1 | H | \psi_2^* \psi_1 \rangle] \\ &= 2^{1/2} [\langle \psi_1 | H | \psi_1^* \psi_2 \rangle + \langle \psi_1 | H | \psi_2^* \psi_1 \rangle]. \end{aligned}$$

7. What is the element coupling  $|\psi_1\rangle$  and  $2^{-1/2} [|\psi_1^* \psi_2\rangle + |\psi_2^* \psi_1\rangle]$ ?

$$\begin{aligned} & \langle \psi_1 | H | 2^{-1/2} [|\psi_1^* \psi_2\rangle + |\psi_2^* \psi_1\rangle] \rangle \\ &= 2^{-1/2} [\langle \psi_1 | H | \psi_1^* \psi_2 \rangle + \langle \psi_1 | H | \psi_2^* \psi_1 \rangle] + (-1) 2^{-1/2} [\langle \psi_1 | H | \psi_1^* \psi_2 \rangle + \langle \psi_1 | H | \psi_2^* \psi_1 \rangle] = 0. \end{aligned}$$

This result should not surprise you because  $|\psi_1\rangle$  is an  $S=0$  singlet state while  $2^{-1/2} [|\psi_1^* \psi_2\rangle + |\psi_2^* \psi_1\rangle]$  is the  $M_S=0$  component of the  $S=1$  triplet state.

8. What is the  $\mathbf{r} = \sum_j e_j \mathbf{r}_j$  electric dipole matrix element between  $|\psi_1\rangle$  and  $2^{-1/2} [|\psi_1 \psi_0\rangle + |\psi_0 \psi_1\rangle]$ ? Is the second function a singlet or triplet? It is a singlet in disguise; by interchanging the  $\psi_0$  and  $\psi_1$  and thus introducing a  $(-1)$ , this function is clearly identified as  $2^{-1/2} [|\psi_1 \psi_0\rangle - |\psi_0 \psi_1\rangle]$  which is a singlet.

The first determinant differs from the latter two by one spin orbital in each case, so

$$\begin{aligned} & \langle \psi_1 \psi_1 | \mathbf{r} | 2^{-1/2} [|\psi_1 \psi_0\rangle + |\psi_0 \psi_1\rangle] \rangle = \\ & 2^{-1/2} [\langle \psi_1 | \mathbf{r} | \psi_0 \rangle + \langle \psi_1 | \mathbf{r} | \psi_0 \rangle] = 2^{1/2} \langle \psi_1 | \mathbf{r} | \psi_0 \rangle. \end{aligned}$$

9. What is the electric dipole matrix elements between the  $1^1 = |1 \ 1 \ 1 \ 1 \rangle$  state and the  $1^1 = 2^{-1/2}[|1 \ 1 \ -1 \ 1 \rangle + |1 \ -1 \ 1 \ 1 \rangle]$  state?

$$\langle 2^{-1/2}[|1 \ 1 \ -1 \ 1 \rangle + |1 \ -1 \ 1 \ 1 \rangle] | \mathbf{r} | 1 \ 1 \ 1 \ 1 \rangle$$

$$= 2^{-1/2}[\langle -1 | \mathbf{r} | 1 \rangle + \langle -1 | \mathbf{r} | 1 \rangle]$$

$$= 2^{1/2} \langle -1 | \mathbf{r} | 1 \rangle.$$

10. As another example of the use of the SC rules, consider the configuration interaction which occurs between the  $1s^2 2s^2$  and  $1s^2 2p^2 \ ^1S$  CSFs in the Be atom.

The CSFs corresponding to these two configurations are as follows:

$$1 = |1s \ 1s \ 2s \ 2s \rangle$$

and

$$2 = 1/\sqrt{3} [|1s \ 1s \ 2p_0 \ 2p_0 \rangle - |1s \ 1s \ 2p_1 \ 2p_{-1} \rangle$$

$$- |1s \ 1s \ 2p_{-1} \ 2p_1 \rangle].$$

The determinantal Hamiltonian matrix elements needed to evaluate the  $2 \times 2$   $H_{K,L}$  matrix appropriate to these two CSFs are evaluated via the SC rules. The first such matrix element is:

$$\langle |1s \ 1s \ 2s \ 2s \rangle | H | 1s \ 1s \ 2s \ 2s \rangle$$

$$= 2h_{1s} + 2h_{2s} + J_{1s,1s} + 4J_{1s,2s} + J_{2s,2s} - 2K_{1s,2s},$$

where

$$h_i = \langle i | -\hbar^2/2m_e \nabla^2 - 4e^2/r | i \rangle,$$

$$J_{i,j} = \langle i \ j | e^2/r_{12} | i \ j \rangle,$$

and

$$K_{ij} = \langle \psi_i | e^2/r_{12} | \psi_j \rangle$$

are the orbital-level one-electron, coulomb, and exchange integrals, respectively.

Coulomb integrals  $J_{ij}$  describe the coulombic interaction of one charge density ( $\psi_i^2$  above) with another charge density ( $\psi_j^2$  above); exchange integrals  $K_{ij}$  describe the interaction of an overlap charge density (i.e., a density of the form  $\psi_i \psi_j$ ) with itself ( $\psi_i \psi_j$  with  $\psi_i \psi_j$  in the above).

The spin functions  $\alpha$  and  $\beta$  which accompany each orbital in  $|1s \uparrow 1s \downarrow 2s \uparrow 2s \downarrow\rangle$  have been eliminated by carrying out the spin integrations as discussed above. Because  $H$  contains no spin operators, this is straightforward and amounts to keeping integrals  $\langle \psi_i | f | \psi_j \rangle$  only if  $\psi_i$  and  $\psi_j$  are of the same spin and integrals  $\langle \psi_i \psi_j | g | \psi_k \psi_l \rangle$  only if  $\psi_i$  and  $\psi_k$  are of the same spin and  $\psi_j$  and  $\psi_l$  are of the same spin. The physical content of the above energy (i.e., Hamiltonian expectation value) of the  $|1s \uparrow 1s \downarrow 2s \uparrow 2s \downarrow\rangle$  determinant is clear:  $2h_{1s} + 2h_{2s}$  is the sum of the expectation values of the one-electron (i.e., kinetic energy and electron-nuclear coulomb interaction) part of the Hamiltonian for the four occupied spin-orbitals;  $J_{1s,1s} + 4J_{1s,2s} + J_{2s,2s} - 2K_{1s,2s}$  contains the coulombic repulsions among all pairs of occupied spin-orbitals minus the exchange interactions among pairs of spin-orbitals with like spin.

The determinantal matrix elements linking  $\psi_1$  and  $\psi_2$  are as follows:

$$\langle |1s \uparrow 1s \downarrow 2s \uparrow 2s \downarrow | H | 1s \uparrow 1s \downarrow 2p_0 \uparrow 2p_0 \downarrow \rangle = \langle 2s \uparrow 2s \downarrow | 2p_0 \uparrow 2p_0 \downarrow \rangle,$$

$$\langle |1s \uparrow 1s \downarrow 2s \uparrow 2s \downarrow | H | 1s \uparrow 1s \downarrow 2p_1 \uparrow 2p_{-1} \downarrow \rangle = \langle 2s \uparrow 2s \downarrow | 2p_1 \uparrow 2p_{-1} \downarrow \rangle,$$

$$\langle |1s \uparrow 1s \downarrow 2s \uparrow 2s \downarrow | H | 1s \uparrow 1s \downarrow 2p_{-1} \uparrow 2p_1 \downarrow \rangle = \langle 2s \uparrow 2s \downarrow | 2p_{-1} \uparrow 2p_1 \downarrow \rangle,$$

where the Dirac convention has been introduced as a shorthand notation for the two-electron integrals (e.g.,  $\langle 2s \uparrow 2s \downarrow | 2p_0 \uparrow 2p_0 \downarrow \rangle$  represents  $\int \int 2s^*(r_1) 2s^*(r_2) e^2/r_{12} 2p_0(r_1) 2p_0(r_2) dr_1 dr_2$ ).

The three integrals shown above can be seen to be equal and to be of the exchange-integral form by expressing the integrals in terms of integrals over cartesian functions and recognizing identities due to the equivalence of the  $2p_x$ ,  $2p_y$ , and  $2p_z$  orbitals. For example,

$$\langle 2s \uparrow 2s \downarrow | 2p_1 \uparrow 2p_{-1} \downarrow \rangle = (1 - i)^2 \{ \langle 2s \uparrow 2s \downarrow | [2p_x + i 2p_y] [2p_x - i 2p_y] \rangle \} =$$

$$1/2 \{ \langle 2s \ 2s \mid x \ x \rangle + \langle 2s \ 2s \mid y \ y \rangle + i \langle 2s \ 2s \mid y \ x \rangle - i \langle 2s \ 2s \mid x \ y \rangle \} =$$

$$\langle 2s \ 2s \mid x \ x \rangle = K_{2s,x}$$

(here the two imaginary terms cancel and the two remaining real integrals are equal);

$$\langle 2s \ 2s \ 2p_0 \ 2p_0 \rangle = \langle 2s \ 2s \mid z \ z \rangle = \langle 2s \ 2s \mid x \ x \rangle = K_{2s,x}$$

(this is because  $K_{2s,z} = K_{2s,x} = K_{2s,y}$ );

$$\langle 2s \ 2s \mid 2p_{-1} 2p_1 \rangle = 1/2 \{ \langle 2s \ 2s \mid [2p_x - i 2p_y] [2p_x + i 2p_y] \rangle \} =$$

$$\langle 2s \ 2s \mid x \ x \rangle = \int \int 2s^*(r_1) 2s^*(r_2) e^2/r_{12} 2p_x(r_1) 2p_x(r_2) dr_1 dr_2 = K_{2s,x}.$$

These integrals are clearly of the exchange type because they involve the coulombic interaction of the  $2s \ 2p_{x,y, \text{or } z}$  overlap charge density with itself.

Moving on, the matrix elements among the three determinants in  $\mathcal{D}_2$  are given as follows:

$$\langle |1s \ 1s \ 2p_0 \ 2p_0 \rangle \mid H \mid |1s \ 1s \ 2p_0 \ 2p_0 \rangle \mid \rangle$$

$$= 2h_{1s} + 2h_{2p} + J_{1s,1s} + J_{2p_z,2p_z} + 4J_{1s,2p} - 2K_{1s,2p}$$

( $J_{1s,2p}$  and  $K_{1s,2p}$  are independent of whether the  $2p$  orbital is  $2p_x$ ,  $2p_y$ , or  $2p_z$ );

$$\langle |1s \ 1s \ 2p_1 \ 2p_{-1} \rangle \mid H \mid |1s \ 1s \ 2p_1 \ 2p_{-1} \rangle \mid \rangle$$

$$= 2h_{1s} + 2h_{2p} + J_{1s,1s} + 4J_{1s,2p} - 2K_{1s,2p} + \langle 2p_1 2p_{-1} \mid 2p_1 2p_{-1} \rangle;$$

$$\langle |1s \ 1s \ 2p_{-1} \ 2p_1 \rangle \mid H \mid |1s \ 1s \ 2p_{-1} \ 2p_1 \rangle \mid \rangle$$

$$= 2h_{1s} + 2h_{2p} + J_{1s,1s} + 4J_{1s,2p} - 2K_{1s,2p} + \langle 2p_{-1} 2p_1 \mid 2p_{-1} 2p_1 \rangle;$$

$$\langle |1s \ 1s \ 2p_0 \ 2p_0 \rangle \mid H \mid |1s \ 1s \ 2p_1 \ 2p_{-1} \rangle \mid \rangle = \langle 2p_0 2p_0 \mid 2p_1 2p_{-1} \rangle$$

$$\begin{aligned} \langle 1s \ 1s \ 2p_0 \ 2p_0 | H | 1s \ 1s \ 2p_{-1} \ 2p_1 \rangle &= \langle 2p_0 2p_0 | 2p_{-1} 2p_1 \rangle \\ \langle 1s \ 1s \ 2p_1 \ 2p_{-1} | H | 1s \ 1s \ 2p_{-1} \ 2p_1 \rangle &= \langle 2p_1 2p_{-1} | 2p_{-1} 2p_1 \rangle. \end{aligned}$$

Certain of these integrals can be recast in terms of cartesian integrals for which equivalences are easier to identify as follows:

$$\langle 2p_0 2p_0 | 2p_1 2p_{-1} \rangle = \langle 2p_0 2p_0 | 2p_{-1} 2p_1 \rangle = \langle z z | x x \rangle = K_{z,x};$$

$$\langle 2p_1 2p_{-1} | 2p_{-1} 2p_1 \rangle = \langle x x | y y \rangle + 1/2[\langle x x | x x \rangle - \langle x y | x y \rangle]$$

$$= K_{x,y} + 1/2 [ J_{x,x} - J_{x,y} ];$$

$$\langle 2p_1 2p_{-1} | 2p_1 2p_{-1} \rangle = \langle 2p_{-1} 2p_1 | 2p_{-1} 2p_1 \rangle = 1/2(J_{x,x} + J_{x,y}).$$

Finally, the 2x2 CI matrix corresponding to the CSFs  $\psi_1$  and  $\psi_2$  can be formed from the above determinantal matrix elements; this results in:

$$H_{11} = 2h_{1s} + 2h_{2s} + J_{1s,1s} + 4J_{1s,2s} + J_{2s,2s} - 2K_{1s,2s};$$

$$H_{12} = -K_{2s,x} / 3;$$

$$H_{22} = 2h_{1s} + 2h_{2p} + J_{1s,1s} + 4J_{1s,2p} - 2K_{1s,2p} + J_{z,z} - 2/3 K_{z,x}.$$

The lowest eigenvalue of this matrix provides this CI calculation's estimate of the ground-state  $^1S$  energy of Be; its eigenvector provides the CI amplitudes for  $\psi_1$  and  $\psi_2$  in this ground-state wavefunction. The other root of the 2x2 secular problem gives an approximation to another  $^1S$  state of higher energy, in particular, a state dominated by the  $3^{-1/2} [ |1s \ 1s \ 2p_0 \ 2p_0 \rangle - |1s \ 1s \ 2p_1 \ 2p_{-1} \rangle - |1s \ 1s \ 2p_{-1} \ 2p_1 \rangle ]$  CSF.

11. As another example, consider the matrix elements which arise in electric dipole transitions between two singlet electronic states:

$\langle \psi_1 | \mathbf{E} \cdot \sum_i \mathbf{e}_i \mathbf{r}_i | \psi_2 \rangle$ . Here  $\mathbf{E} \cdot \sum_i \mathbf{e}_i \mathbf{r}_i$  is the one-electron operator describing the interaction of an electric field of magnitude and polarization  $\mathbf{E}$  with the instantaneous dipole moment

of the electrons (the contribution to the dipole operator arising from the nuclear charges  $-\sum_a Z_a e^2 \mathbf{R}_a$  does not contribute because, when placed between  $\psi_1$  and  $\psi_2$ , this zero-electron operator yields a vanishing integral because  $\psi_1$  and  $\psi_2$  are orthogonal).

When the states  $\psi_1$  and  $\psi_2$  are described as linear combinations of CSFs as introduced earlier ( $\psi_i = \sum_K C_{iK} \Phi_K$ ), these matrix elements can be expressed in terms of CSF-based matrix elements  $\langle \Phi_K | \sum_i \mathbf{e} \mathbf{r}_i | \Phi_L \rangle$ . The fact that the electric dipole operator is a one-electron operator, in combination with the SC rules, guarantees that only states for which the dominant determinants differ by at most a single spin-orbital (i.e., those which are "singly excited") can be connected via electric dipole transitions through first order (i.e., in a one-photon transition to which the  $\langle \psi_1 | \sum_i \mathbf{e} \mathbf{r}_i | \psi_2 \rangle$  matrix elements pertain). It is for this reason that light with energy adequate to ionize or excite deep core electrons in atoms or molecules usually causes such ionization or excitation rather than double ionization or excitation of valence-level electrons; the latter are two-electron events.

In, for example, the  $\pi \rightarrow \pi^*$  excitation of an olefin, the ground and excited states are dominated by CSFs of the form (where all but the "active"  $\pi$  and  $\pi^*$  orbitals are not explicitly written):

$$\psi_1 = | \dots \pi \dots |$$

and

$$\psi_2 = 1/\sqrt{2} [ | \dots \pi^* \dots | - | \dots \pi \dots | ].$$

The electric dipole matrix element between these two CSFs can be found, using the SC rules, to be

$$e/\sqrt{2} [ \langle \dots \pi^* \dots | \mathbf{r} | \dots \pi \dots \rangle + \langle \dots \pi \dots | \mathbf{r} | \dots \pi^* \dots \rangle ] = \sqrt{2} e \langle \dots \pi^* \dots | \mathbf{r} | \dots \pi \dots \rangle.$$

Notice that in evaluating the second determinantal integral  $\langle \dots \pi \dots | \mathbf{e} \mathbf{r} | \dots \pi^* \dots \rangle$ , a sign change occurs when one puts the two determinants into maximum coincidence; this sign change then makes the minus sign in  $\psi_2$  yield a positive sign in the final result.

#### IV. Summary

In all of the above examples, the SC rules were used to reduce matrix elements of one- or two- electron operators between determinantal functions to one- or two- electron integrals over the orbitals which appear in the determinants. In any ab initio electronic structure computer program there must exist the capability to form symmetry-adapted CSFs and to evaluate, using these SC rules, the Hamiltonian and other operators' matrix elements among these CSFs in terms of integrals over the mos that appear in the CSFs. The SC rules provide not only the tools to compute quantitative matrix elements; they allow one to understand in qualitative terms the strengths of interactions among CSFs. In the following section, the SC rules are used to explain why chemical reactions in which the reactants and products have dominant CSFs that differ by two spin-orbital occupancies often display activation energies that exceed the reaction endoergicity.

## Chapter 12

*Along "reaction paths", configurations can be connected one-to-one according to their symmetries and energies. This is another part of the Woodward-Hoffmann rules*

### I. Concepts of Configuration and State Energies

#### A. Plots of CSF Energies Give Configuration Correlation Diagrams

The energy of a particular electronic state of an atom or molecule has been expressed in terms of Hamiltonian matrix elements, using the SC rules, over the various spin- and spatially-adapted determinants or CSFs which enter into the state wavefunction.

$$E = \sum_{I,J} C_I C_J \langle I | H | J \rangle$$

The diagonal matrix elements of  $H$  in the CSF basis multiplied by the appropriate CI amplitudes  $\langle I | H | I \rangle C_I C_I$  represent the energy of the  $I^{\text{th}}$  CSF weighted by the strength ( $C_I^2$ ) of that CSF in the wavefunction. The off-diagonal elements represent the effects of mixing among the CSFs; mixing is strongest whenever two or more CSFs have nearly the same energy (i.e.,  $\langle I | H | I \rangle \approx \langle J | H | J \rangle$ ) and there is strong coupling (i.e.,  $\langle I | H | J \rangle$  is large). Whenever the CSFs are widely separated in energy, each wavefunction is dominated by a single CSF.

#### B. CSFs Interact and Couple to Produce States and State Correlation Diagrams

Just as orbital energies connected according to their symmetries and plotted as functions of geometry constitute an orbital correlation diagram, plots of the diagonal CSF energies, connected according to symmetry, constitute a configuration correlation diagram (CCD). If, near regions where energies of CSFs of the same symmetry cross (according to the direct product rule of group theory discussed in Appendix E, only CSFs of the same symmetry mix because only they have non-vanishing  $\langle I | H | J \rangle$  matrix elements), CI mixing is allowed to couple the CSFs to give rise to "avoided crossings", then the CCD is converted into a so-called state correlation diagram (SCD).

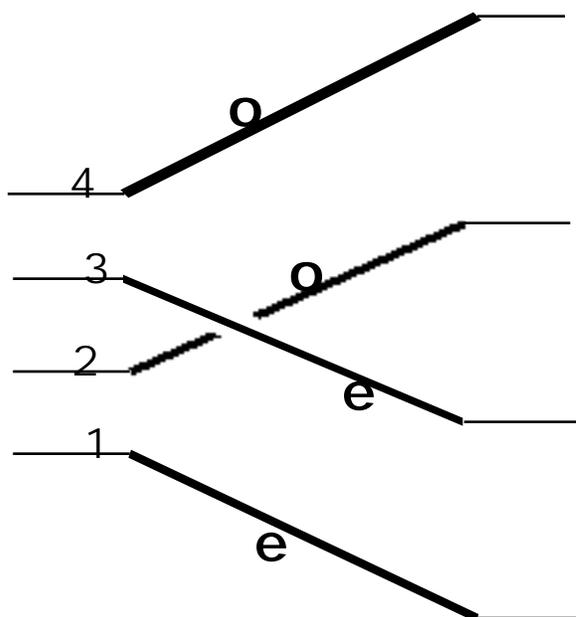
#### C. CSFs that Differ by Two Spin-Orbitals Interact Less Strongly than CSFs that Differ by One Spin-Orbital

The strengths of the couplings between pairs of CSFs whose energies cross are evaluated through the SC rules. CSFs that differ by more than two spin-orbital occupancies do not couple; the SC rules give vanishing Hamiltonian matrix elements for such pairs. Pairs that differ by two spin-orbitals (e.g.  $|\dots a \dots b \dots\rangle$  vs  $|\dots a' \dots b' \dots\rangle$ ) have interaction strengths determined by the two-electron integrals  $\langle ab | a'b' \rangle - \langle ab | b'a' \rangle$ . Pairs that differ by a single spin-orbital (e.g.  $|\dots a \dots \dots\rangle$  vs  $|\dots a' \dots \dots\rangle$ ) are coupled by the one- and two- electron parts of H:  $\langle a | f | b \rangle + \sum_j [\langle aj | bj \rangle - \langle aj | jb \rangle]$ . Usually, couplings among CSFs that differ by two spin-orbitals are much weaker than those among CSFs that differ by one spin-orbital. In the latter case, the full strength of H is brought to bear, whereas in the former, only the electron-electron coulomb potential is operative.

#### D. State Correlation Diagrams

In the SCD, the energies are connected by symmetry but the configurational nature as reflected in the  $C_I$  coefficients changes as one passes through geometries where crossings in the CCD occur. The SCD is the ultimate product of an orbital and configuration symmetry and energy analysis and gives one the most useful information about whether reactions will or will not encounter barriers on the ground and excited state surfaces.

As an example of the application of CCD's and SCD's, consider the disrotatory closing of 1,3-butadiene to produce cyclobutene. The OCD given earlier for this proposed reaction path is reproduced below.



Recall that the symmetry labels e and o refer to the symmetries of the orbitals under reflection through the one  $C_v$  plane that is preserved throughout the proposed disrotatory closing. Low-energy configurations (assuming one is interested in the thermal or low-lying photochemically excited-state reactivity of this system) for the reactant molecule and their overall space and spin symmetry are as follows:

(i)  $1^2 2^2 = 1e^2 1o^2$ ,  ${}^1\text{Even}$

(ii)  $1^2 2^1 3^1 = 1e^2 1o^1 2e^1$ ,  ${}^3\text{Odd}$  and  ${}^1\text{Odd}$ .

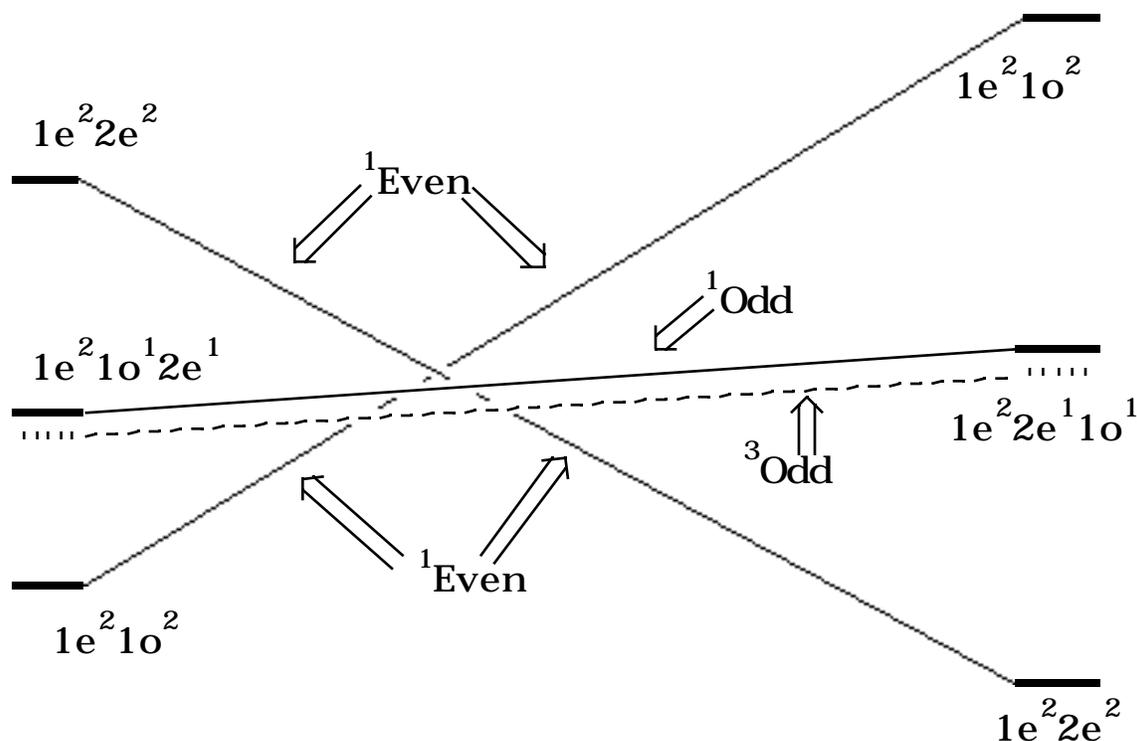
For the product molecule, on the other hand, the low-lying states are

(iii)  $2^2 = 1e^2 2e^2$ ,  ${}^1\text{Even}$

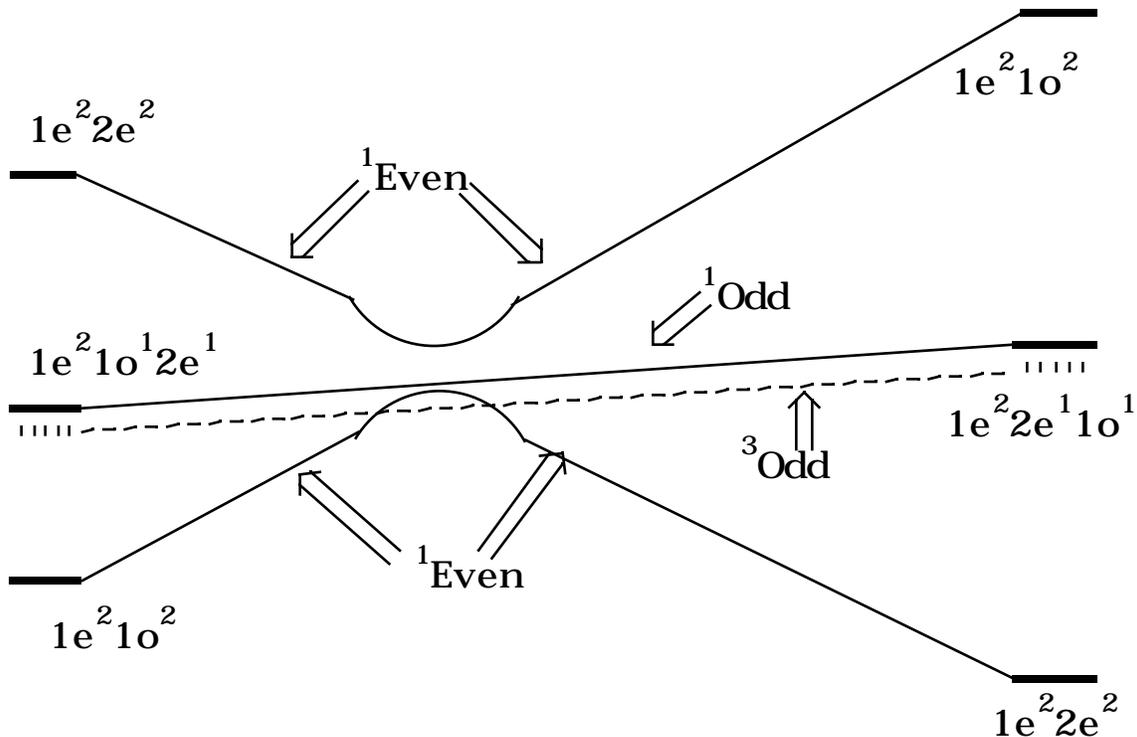
(iv)  $2^1 1^1 = 1e^2 2e^1 1o^1$ ,  ${}^3\text{Odd}$ ,  ${}^1\text{Odd}$ .

Notice that although the lowest energy configuration at the reactant geometry  $1^2 2^2 = 1e^2 1o^2$  and the lowest energy configuration at the product geometry  $2^2 = 1e^2 2e^2$  are both of  ${}^1\text{Even}$  symmetry, they are not the same configurations; they involve occupancy of different symmetry orbitals.

In constructing the CCD, one must trace the energies of all four of the above CSFs (actually there are more because the singlet and triplet excited CSFs must be treated independently) along the proposed reaction path. In doing so, one must realize that the  $1e^2 1o^2$  CSF has low energy on the reactant side of the CCD because it corresponds to  $1^2 2^2$  orbital occupancy, but on the product side, it corresponds to  $2^2 2^2$  orbital occupancy and is thus of very high energy. Likewise, the  $1e^2 2e^2$  CSF has low energy on the product side where it is  $2^2 2^2$  but high energy on the reactant side where it corresponds to  $1^2 3^2$ . The low-lying singly excited CSFs are  $1e^2 2e^1 1o^1$  at both reactant and product geometries; in the former case, they correspond to  $1^2 2^1 3^1$  occupancy and at the latter to  $2^1 1^1$  occupancy. Plotting the energies of these CSFs along the disrotatory reaction path results in the CCD shown below.



If the two  $1EVEN$  CSFs which cross are allowed to interact (the SC rules give their interaction strength in terms of the exchange integral  $\langle 1e^2 1o^2 | H | 1e^2 2e^2 \rangle = \langle 1o1o | 2e2e \rangle = K_{1o,2e}$ ) to produce states which are combinations of the two  $1EVEN$  CSFs, the following SCD results:



This SCD predicts that the thermal (i.e., on the ground electronic surface) disrotatory rearrangement of 1,3-butadiene to produce cyclobutene will experience a symmetry-imposed barrier which arises because of the avoided crossing of the two  ${}^1\text{Even}$  configurations; this avoidance occurs because the orbital occupancy pattern (i.e., the configuration) which is best for the ground state of the reactant is not identical to that of the product molecule. The SCD also predicts that there should be no symmetry-imposed barrier for the singlet or triplet excited-state rearrangement, although the reaction leading from excited 1,3-butadiene to excited cyclobutene may be endothermic on the grounds of bond strengths alone.

It is also possible to infer from the SCD that excitation of the lowest singlet state of 1,3-butadiene would involve a low quantum yield for producing cyclobutene and would, in fact, produce ground-state butadiene. As the reaction proceeds along the singlet surface this  ${}^1\text{Odd}$  state intersects the ground  ${}^1\text{Even}$  surface on the reactant side of the diagram; internal conversion (i.e., quenching from the  ${}^1\text{Odd}$  to the  ${}^1\text{Even}$  surfaces induced by using a vibration of odd symmetry to "digest" the excess energy (much like vibronic borrowing in spectroscopy) can lead to production of ground-state reactant molecules. Some fraction of such events will lead to the system remaining on the  ${}^1\text{Odd}$  surface until, further along the reaction path, the  ${}^1\text{Odd}$  surface again intersects the  ${}^1\text{Even}$  surface on the product side at which time quenching to produce ground-state products can occur.

Although, in principle, it is possible for some fraction of the events to follow the  $^1\text{O}dd$  surface beyond this second intersection and to thus lead to  $^1\text{O}dd$  product molecules that might fluoresce, quenching is known to be rapid in most polyatomic molecules; as a result, reactions which are chemiluminescent are rare. An appropriate introduction to the use of OCD's, CCD's, and SCD's as well as the radiationless processes that can occur in thermal and photochemical reactions is given in the text Energetic Principles of Chemical Reactions, J. Simons, Jones and Bartlett, Boston (1983).

## II. Mixing of Covalent and Ionic Configurations

As chemists, much of our intuition concerning chemical bonds is built on simple models introduced in undergraduate chemistry courses. The detailed examination of the  $\text{H}_2$  molecule via the valence bond and molecular orbital approaches forms the basis of our thinking about bonding when confronted with new systems. Let us examine this model system in further detail to explore the electronic states that arise by occupying two orbitals (derived from the two  $1s$  orbitals on the two hydrogen atoms) with two electrons.

In total, there exist six electronic states for all such two-orbital, two-electron systems. The heterolytic fragments  $\text{X} + \text{Y}:$  and  $\text{X}:: + \text{Y}$  produce two singlet states; the homolytic fragments  $\text{X}\cdot + \text{Y}\cdot$  produce one singlet state and a set of three triplet states having  $M_S = 1, 0,$  and  $-1$ . Understanding the relative energies of these six states, their bonding and antibonding characters, and which molecular state dissociates to which asymptote are important.

Before proceeding, it is important to clarify the notation (e.g.,  $\text{X}\cdot$ ,  $\text{Y}\cdot$ ,  $\text{X}, \text{Y}:$ , etc.), which is designed to be applicable to neutral as well as charged species. In all cases considered here, only two electrons play active roles in the bond formation. These electrons are represented by the dots. The symbols  $\text{X}\cdot$  and  $\text{Y}\cdot$  are used to denote species in which a single electron is attached to the respective fragment. By  $\text{X}::$ , we mean that both electrons are attached to the X- fragment;  $\text{Y}$  means that neither electron resides on the Y- fragment. Let us now examine the various bonding situations that can occur; these examples will help illustrate and further clarify this notation.

### A. The $\text{H}_2$ Case in Which Homolytic Bond Cleavage is Favored

To consider why the two-orbital two-electron single bond formation case can be more complex than often thought, let us consider the  $\text{H}_2$  system in more detail. In the molecular orbital description of  $\text{H}_2$ , both bonding  $\sigma_g$  and antibonding  $\sigma_u$  mos appear.

There are two electrons that can both occupy the  $\sigma_g$  mo to yield the ground electronic state  $H_2(1\sigma_g^+, \sigma_g^2)$ ; however, they can also occupy both orbitals to yield  $3\sigma_u^+(\sigma_g^1 \sigma_u^1)$  and  $1\sigma_u^+(\sigma_g^1 \sigma_u^1)$ , or both can occupy the  $\sigma_u$  mo to give the  $1\sigma_g^+(\sigma_u^2)$  state. As demonstrated explicitly below, these latter two states dissociate heterolytically to  $X + Y : = H^+ + H^-$ , and are sufficiently high in energy relative to  $X\cdot + Y\cdot = H + H$  that we ordinarily can ignore them. However, their presence and character are important in the development of a full treatment of the molecular orbital model for  $H_2$  and are essential to a proper treatment of cases in which heterolytic bond cleavage is favored.

## B. Cases in Which Heterolytic Bond Cleavage is Favored

For some systems one or both of the heterolytic bond dissociation asymptotes (e.g.,  $X + Y :$  or  $X : + Y$ ) may be lower in energy than the homolytic bond dissociation asymptote. Thus, the states that are analogues of the  $1\sigma_u^+(\sigma_g^1 \sigma_u^1)$  and  $1\sigma_g^+(\sigma_u^2)$  states of  $H_2$  can no longer be ignored in understanding the valence states of the  $XY$  molecules. This situation arises quite naturally in systems involving transition metals, where interactions between empty metal or metal ion orbitals and 2-electron donor ligands are ubiquitous.

Two classes of systems illustrate cases for which heterolytic bond dissociation lies lower than the homolytic products. The first involves transition metal dimer cations,  $M_2^+$ . Especially for metals to the right side of the periodic table, such cations can be considered to have ground-state electron configurations with  $2d^n d^{n+1}$  character, where the d electrons are not heavily involved in the bonding and the  $\sigma$  bond is formed primarily from the metal atom s orbitals. If the  $\sigma$  bond is homolytically broken, one forms  $X\cdot + Y\cdot = M(s^1 d^{n+1}) + M^+(s^1 d^n)$ . For most metals, this dissociation asymptote lies higher in energy than the heterolytic products  $X : + Y = M(s^2 d^n) + M^+(s^0 d^{n+1})$ , since the latter electron configurations correspond to the ground states for the neutrals and ions, respectively. A prototypical species which fits this bonding picture is  $Ni_2^+$ .

The second type of system in which heterolytic cleavage is favored arises with a metal-ligand complex having an atomic metal ion (with a  $s^0 d^{n+1}$  configuration) and a two electron donor,  $L :$ . A prototype is  $(Ag-C_6H_6)^+$  which was observed to photodissociate to form  $X\cdot + Y\cdot = Ag(2S, s^1 d^{10}) + C_6H_6^+(2B_1)$  rather than the lower energy (heterolytically cleaved) dissociation limit  $Y + X : = Ag^+(1S, s^0 d^{10}) + C_6H_6(1A_1)$ .

## C. Analysis of Two-Electron, Two-Orbital, Single-Bond Formation

## 1. Orbitals, Configurations and States

The resultant family of six electronic states can be described in terms of the six configuration state functions (CSFs) that arise when one occupies the pair of bonding and antibonding  $\sigma$  molecular orbitals with two electrons. The CSFs are combinations of Slater determinants formed to generate proper spin- and spatial symmetry- functions.

The spin- and spatial- symmetry adapted N-electron functions referred to as CSFs can be formed from one or more Slater determinants. For example, to describe the singlet CSF corresponding to the closed-shell  $\sigma^2$  orbital occupancy, a single Slater determinant

$$^1 \sigma^2(0) = | \sigma \bar{\sigma} | = (2)^{-1/2} \{ \sigma(1) \bar{\sigma}(2) - \bar{\sigma}(1) \sigma(2) \}$$

suffices. An analogous expression for the  $(\sigma^*)^2$  CSF is given by

$$^1 \sigma^{*2}(0) = | \sigma^* \bar{\sigma}^* | = (2)^{-1/2} \{ \sigma^*(1) \bar{\sigma}^*(2) - \bar{\sigma}^*(1) \sigma^*(2) \}.$$

Also, the  $M_S = 1$  component of the triplet state having  $\sigma^* \sigma$  orbital occupancy can be written as a single Slater determinant:

$$^3 \sigma^* \sigma(1) = | \sigma^* \sigma | = (2)^{-1/2} \{ \sigma^*(1) \sigma(2) - \sigma^*(2) \sigma(1) \},$$

as can the  $M_S = -1$  component of the triplet state

$$^3 \sigma^* \sigma(-1) = | \sigma \sigma^* | = (2)^{-1/2} \{ \sigma(1) \sigma^*(2) - \sigma(2) \sigma^*(1) \}.$$

However, to describe the singlet CSF and  $M_S = 0$  triplet CSF belonging to the  $\sigma^* \sigma$  occupancy, two Slater determinants are needed:

$$^1 \sigma^* \sigma(0) = \frac{1}{\sqrt{2}} [ \sigma^* \sigma - \sigma \sigma^* ]$$

is the singlet CSF and

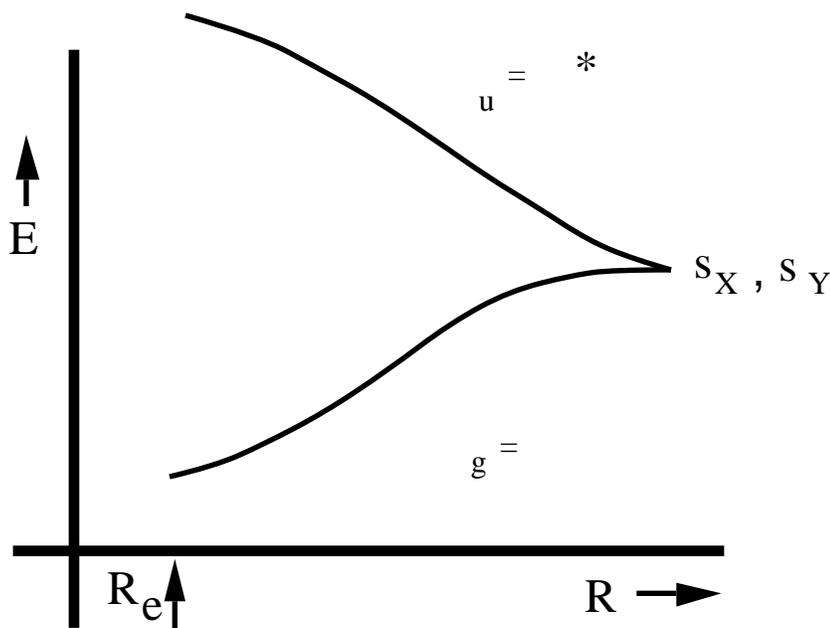
$$^3 \sigma^* \sigma(0) = \frac{1}{\sqrt{2}} [ \sigma^* \sigma + \sigma \sigma^* ]$$

is the triplet CSF. In each case, the spin quantum number  $S$ , its z-axis projection  $M_S$ , and the quantum number are given in the conventional  $^{2S+1} (M_S)$  notation.

## 2. Orbital, CSF, and State Correlation Diagrams

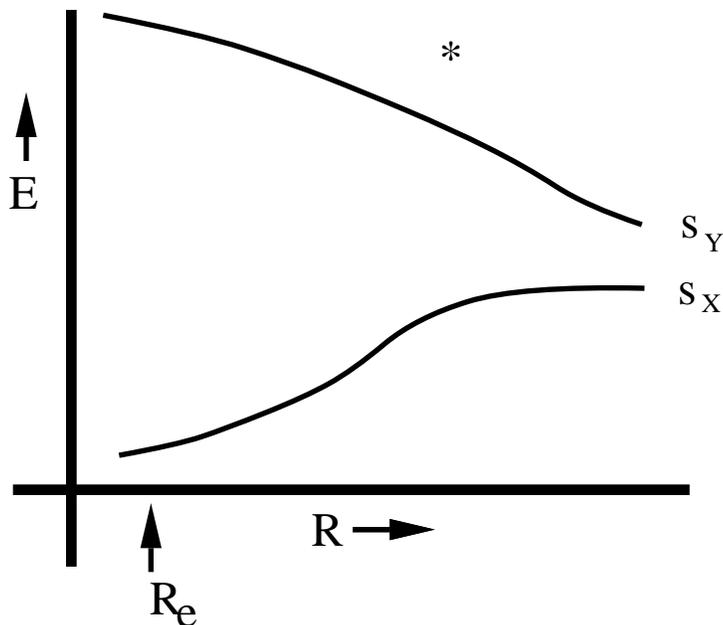
### i. Orbital Diagrams

The two orbitals of the constituent atoms or functional groups (denoted  $s_X$  and  $s_Y$  for convenience and in anticipation of considering groups X and Y that possess valence s orbitals) combine to form a bonding  $=_g$  molecular orbital and an antibonding  $^*_u$  molecular orbital (mo). As the distance  $R$  between the X and Y fragments is changed from near its equilibrium value of  $R_e$  and approaches infinity, the energies of the  $=_g$  and  $^*_u$  orbitals vary in a manner well known to chemists as depicted below.



Energies of the bonding  $=_g$  and antibonding  $^*_u$  orbitals as functions of interfragment distance;  $R_e$  denotes a distance near the equilibrium bond length for XY.

In the heteronuclear case, the  $s_X$  and  $s_Y$  orbitals still combine to form a bonding  $=_g$  and an antibonding  $^*_u$  orbital, although these orbitals no longer belong to  $g$  and  $u$  symmetry. The energies of these orbitals, for  $R$  values ranging from near  $R_e$  to  $R \rightarrow \infty$ , are depicted below.



Energies of the bonding and antibonding  $*$  orbitals as functions of internuclear distance. Here, X is more electronegative than Y.

For the homonuclear case, as  $R$  approaches  $\infty$ , the energies of the  $g$  and  $u$  orbitals become degenerate. Moreover, as  $R \rightarrow 0$ , the orbital energies approach those of the united atom. In the heteronuclear situation, as  $R$  approaches  $\infty$ , the energy of the  $g$  orbital approaches the energy of the  $s_X$  orbital, and the  $u$  orbital converges to the  $s_Y$  orbital energy. Unlike the homonuclear case, the  $g$  and  $u$  orbitals are not degenerate as  $R \rightarrow \infty$ . The energy "gap" between the  $g$  and  $u$  orbitals at  $R = \infty$  depends on the electronegativity difference between the groups X and Y. If this gap is small, it is expected that the behavior of this (slightly) heteronuclear system should approach that of the homonuclear  $X_2$  and  $Y_2$  systems. Such similarities are demonstrated in the next section.

## ii. Configuration and State Diagrams

The energy variation in these orbital energies gives rise to variations in the energies of the six CSFs and of the six electronic states that arise as combinations of these CSFs. The three singlet ( $^1 g(0)$ ,  $^1 u(0)$ , and  $^1 g^*(0)$ ) and three triplet ( $^3 g(1)$ ,  $^3 g(0)$  and  $^3 g(-1)$ ) CSFs are, by no means, the true electronic eigenstates of the system; they are simply spin and spatial angular momentum adapted antisymmetric spin-orbital products. In principle, the set of CSFs of the same symmetry must be combined to form the proper electronic eigenstates of the system:

$$= C \quad .$$

Within the approximation that the valence electronic states can be described adequately as combinations of the above valence CSFs, the three  $1 \sigma$ ,  $1 \sigma^*$ , and  $1 \sigma^{**}$  CSFs must be combined to form the three lowest energy valence electronic states of  $1 \sigma$  symmetry. For the homonuclear case, the  $1 \sigma^*$  CSF does not couple with the other two because it is of ungerade symmetry, while the other CSFs  $1 \sigma$  and  $1 \sigma^{**}$  have gerade symmetry and do combine.

The relative amplitudes  $C$  of the CSFs within each state are determined by solving the configuration-interaction (CI) secular problem:

$$\sum_J H_{JJ} C_J = E C$$

for the state energies  $E$  and state CI coefficient vectors  $C$ . Here,  $H$  is the electronic Hamiltonian of the molecule.

To understand the extent to which the  $1 \sigma$  and  $1 \sigma^{**}$  (and  $1 \sigma^*$  for heteronuclear cases) CSFs couple, it is useful to examine the energies

$H$  of these CSFs for the range of internuclear distances of interest  $R_e < R < \infty$ . Near  $R_e$ , where the energy of the  $2p$  orbital is substantially below that of the  $2s^*$  orbital, the  $2p \sigma$  CSF lies significantly below the  $2s^* \sigma$  CSF which, in turn lies below the  $2s^* \sigma^{**}$  CSF; the large energy splittings among these three CSFs simply reflecting the large gap between the  $2p$  and  $2s^*$  orbitals. The  $3p \sigma$  CSF generally lies below the corresponding  $1 \sigma^*$  CSF by an amount related to the exchange energy between the  $2p$  and  $2s^*$  orbitals.

As  $R \rightarrow \infty$ , the CSF energies  $H_{JJ}$  are more difficult to "intuit" because the  $2p$  and  $2s^*$  orbitals become degenerate (in the homonuclear case) or nearly so. To pursue this point and arrive at an energy ordering for the CSFs that is appropriate to the  $R \rightarrow \infty$  region, it is useful to express each of the above CSFs in terms of the atomic orbitals  $s_x$  and  $s_y$  that comprise  $2p$  and  $2s^*$ . To do so, the LCAO-MO expressions for  $2p \sigma$  and  $2s^* \sigma$ ,

$$2p \sigma = C [s_x + z s_y]$$

and

$$2s^* \sigma = C^* [z s_x - s_y],$$

are substituted into the Slater determinant definitions of the CSFs. Here C and C\* are the normalization constants. The parameter z is 1.0 in the homonuclear case and deviates from 1.0 in relation to the  $s_x$  and  $s_y$  orbital energy difference (if  $s_x$  lies below  $s_y$ , then  $z < 1.0$ ; if  $s_x$  lies above  $s_y$ ,  $z > 1.0$ ).

To simplify the analysis of the above CSFs, the familiar homonuclear case in which  $z = 1.0$  will be examined first. The process of substituting the above expressions for  $\psi_1^*$  and  $\psi_3^*$  into the Slater determinants that define the singlet and triplet CSFs can be illustrated as follows:

$$\begin{aligned} \psi_1^*(0) &= \frac{1}{\sqrt{2}} (s_x + s_y) (s_x + s_y) \\ &= C^2 [ s_x s_x + s_y s_y + s_x s_y + s_y s_x ] \end{aligned}$$

The first two of these atomic-orbital-based Slater determinants ( $s_x s_x$  and  $s_y s_y$ ) are denoted "ionic" because they describe atomic orbital occupancies, which are appropriate to the R region, that correspond to  $X\cdot + Y$  and  $X + Y\cdot$  valence bond structures, while  $s_x s_y$  and  $s_y s_x$  are called "covalent" because they correspond to  $X\cdot + Y\cdot$  structures.

In similar fashion, the remaining five CSFs may be expressed in terms of atomic-orbital-based Slater determinants. In so doing, use is made of the antisymmetry of the Slater determinants

$|\psi_1 \psi_2 \psi_3| = -|\psi_1 \psi_3 \psi_2|$ , which implies that any determinant in which two or more spin-orbitals are identical vanishes  $|\psi_1 \psi_2 \psi_2| = -|\psi_1 \psi_2 \psi_2| = 0$ . The result of decomposing the mo-based CSFs into their atomic orbital components is as follows:

$$\begin{aligned} \psi_1^{**}(0) &= \frac{1}{\sqrt{2}} (s_x + s_y) (s_x + s_y) \\ &= C^{*2} [ s_x s_x + s_y s_y \\ &\quad - s_x s_y - s_y s_x ] \\ \psi_1^*(0) &= \frac{1}{\sqrt{2}} [ s_x s_x - s_y s_y ] \\ &= CC^* \sqrt{2} [ s_x s_x - s_y s_y ] \\ \psi_3^*(1) &= \frac{1}{\sqrt{2}} (s_x - s_y) (s_x - s_y) \\ &= CC^* 2 s_y s_x \end{aligned}$$

$${}^3 \sigma^*(0) = \frac{1}{\sqrt{2}} [ \sigma^* + \sigma^* ]$$

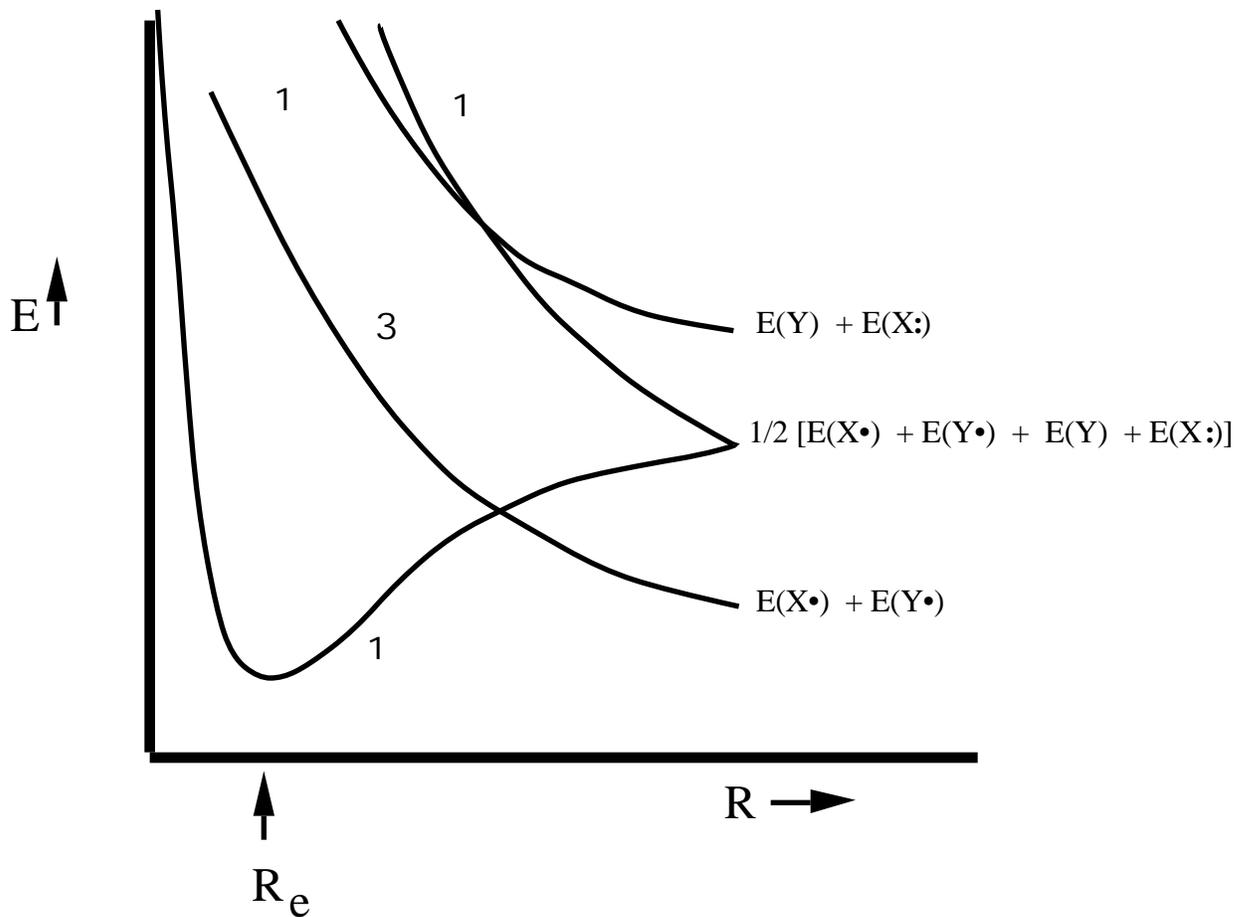
$$= CC^* \sqrt{2} [ s_y s_x - s_x s_y ]$$

$${}^3 \sigma^*(-1) = \sigma^*$$

$$= CC^* 2 s_y s_x$$

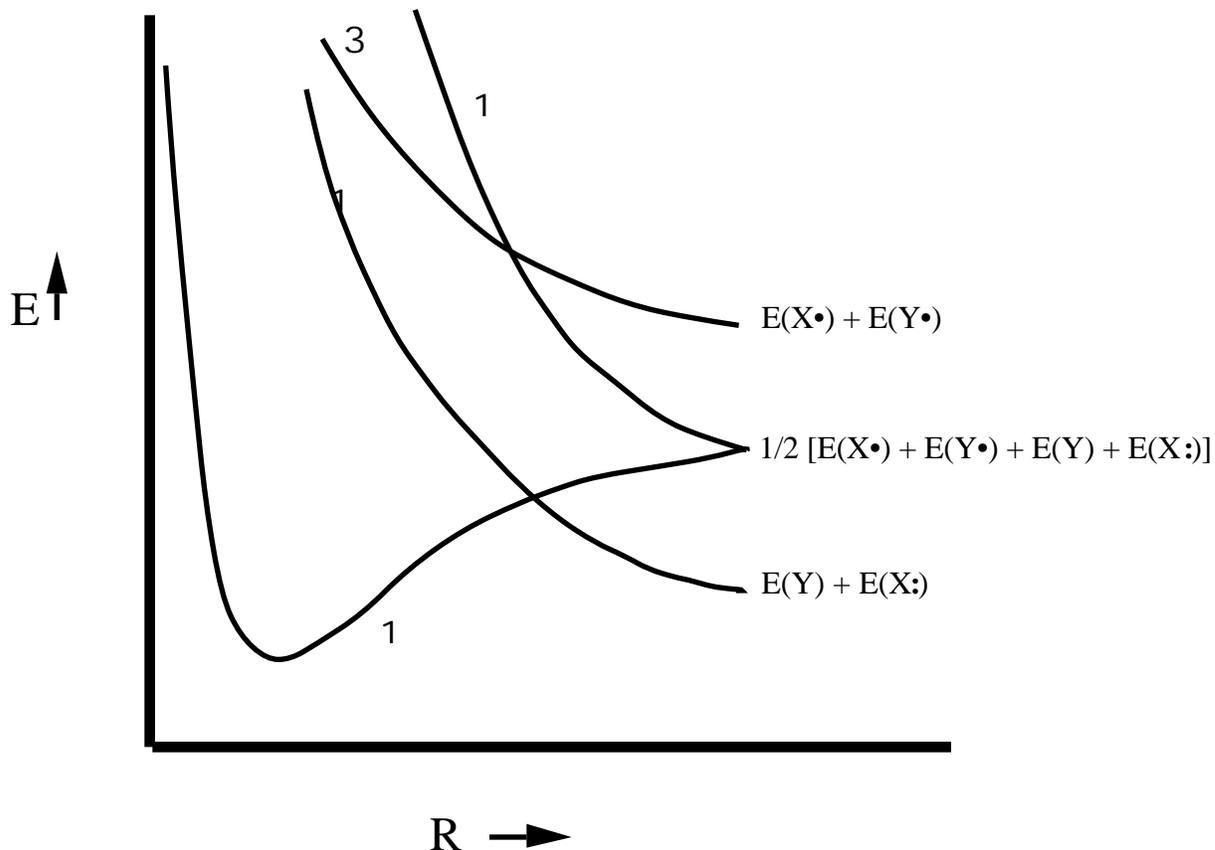
These decompositions of the six valence CSFs into atomic-orbital or valence bond components allow the  $R = \infty$  energies of the CSFs to be specified. For example, the fact that both  ${}^1 \sigma$  and  ${}^1 \sigma^*$  contain 50% ionic and 50% covalent structures implies that, as  $R \rightarrow \infty$ , both of their energies will approach the average of the covalent and ionic atomic energies  $1/2 [E(X\cdot) + E(Y\cdot) + E(Y) + E(X\cdot)]$ . The  ${}^1 \sigma^*$  CSF energy approaches the purely ionic value  $E(Y) + E(X\cdot)$  as  $R \rightarrow \infty$ . The energies of  ${}^3 \sigma^*(0)$ ,  ${}^3 \sigma^*(1)$  and  ${}^3 \sigma^*(-1)$  all approach the purely covalent value  $E(X\cdot) + E(Y\cdot)$  as  $R \rightarrow \infty$ .

The behaviors of the energies of the six valence CSFs as  $R$  varies are depicted below for situations in which the homolytic bond cleavage is energetically favored (i.e., for which  $E(X\cdot) + E(Y\cdot) < E(Y) + E(X\cdot)$ ).



Configuration correlation diagram for homonuclear case in which homolytic bond cleavage is energetically favored.

When heterolytic bond cleavage is favored, the configuration energies as functions of internuclear distance vary as shown below.



Configuration correlation diagram for a homonuclear case in which heterolytic bond cleavage is energetically favored.

It is essential to realize that the energies  $H$  of the CSFs do not represent the energies of the true electronic states  $E_K$ ; the CSFs are simply spin- and spatial-symmetry adapted antisymmetric functions that form a basis in terms of which to expand the true electronic states. For R-values at which the CSF energies are separated widely, the true  $E_K$  are rather well approximated by individual  $H$  values; such is the case near  $R_e$ .

For the homonuclear example, the  $^1$  and  $^1$  \*\* CSFs undergo CI coupling to form a pair of states of  $^1$  symmetry (the  $^1$  \* CSF cannot partake in this CI mixing because it is of ungerade symmetry; the  $^3$  \* states can not mix because they are of triplet spin symmetry). The CI mixing of the  $^1$  and  $^1$  \*\* CSFs is described in terms of a 2x2 secular problem

$$\begin{array}{cc} \begin{array}{cc} 1 & H & 1 \\ 1 & ** & H & 1 \end{array} & \begin{array}{cc} 1 & H & 1 & ** \\ 1 & ** & 1 & ** \end{array} & \begin{array}{c} A \\ B \end{array} & \begin{array}{c} A \\ B \end{array} \\ & & = E & \end{array}$$

The diagonal entries are the CSF energies depicted in the above two figures. Using the Slater-Condon rules, the off-diagonal coupling can be expressed in terms of an exchange integral between the  $1$  and  $1^{**}$  orbitals:

$$1 \quad H \quad 1 \quad ** = \int \psi_1^* \psi_{1^{**}} H \psi_1 \psi_{1^{**}} = \frac{1}{r_{12}} \int \psi_1^* \psi_{1^{**}} \psi_1 \psi_{1^{**}} = J$$

At  $R \rightarrow \infty$ , where the  $1$  and  $1^{**}$  CSFs are degenerate, the two solutions to the above CI secular problem are:

$$E_{\pm} = \frac{1}{2} [ E(X\bullet) + E(Y\bullet) + E(Y) + E(X\bullet) ] \pm \frac{1}{r_{12}} J$$

with respective amplitudes for the  $1$  and  $1^{**}$  CSFs given by

$$A_{\pm} = \pm \frac{1}{\sqrt{2}} ; \quad B_{\pm} = \mp \frac{1}{\sqrt{2}}$$

The first solution thus has

$$\psi_{-} = \frac{1}{\sqrt{2}} [ \psi_1 - \psi_{1^{**}} ]$$

which, when decomposed into atomic valence bond components, yields

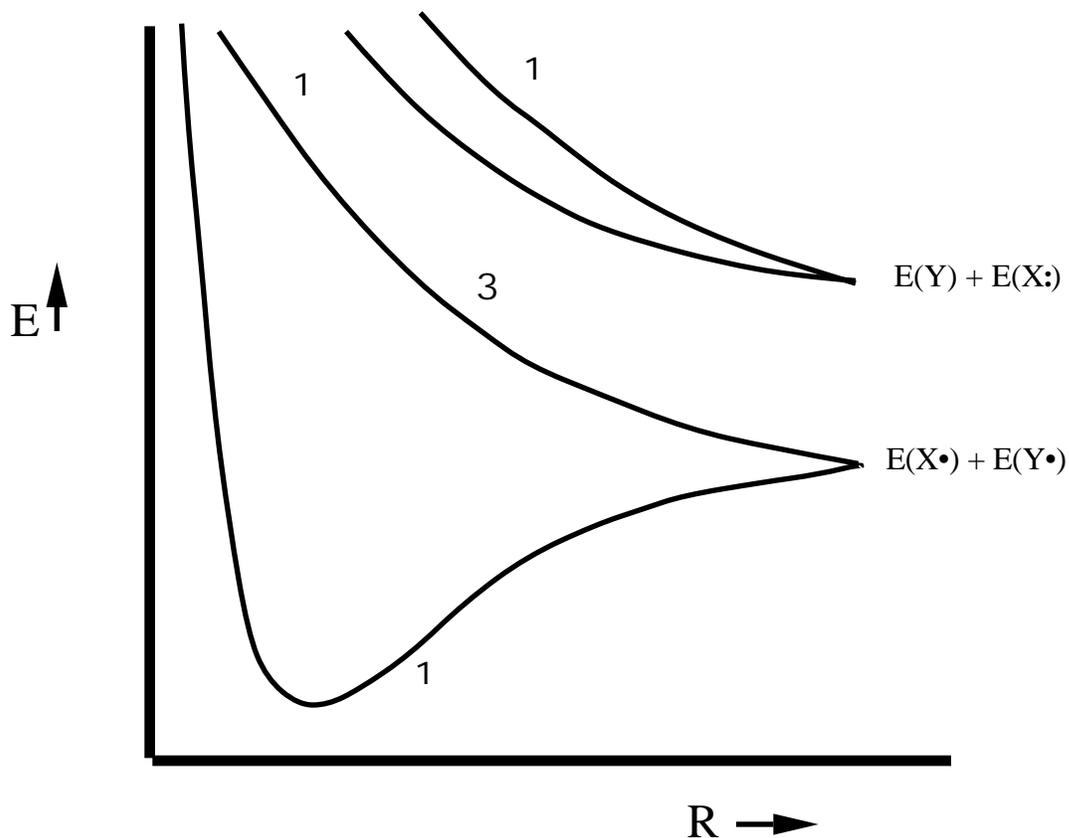
$$\psi_{-} = \frac{1}{\sqrt{2}} [ s_x s_y - s_x s_y ]$$

The other root has

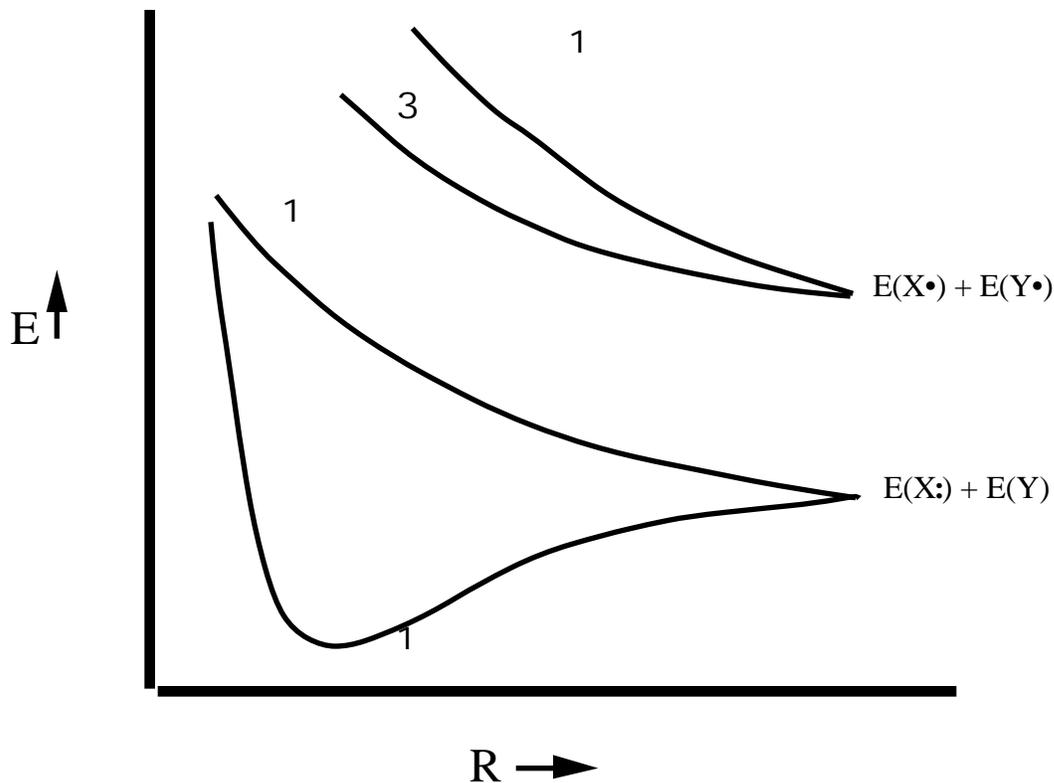
$$\begin{aligned} \psi_{+} &= \frac{1}{\sqrt{2}} [ \psi_1 + \psi_{1^{**}} ] \\ &= \frac{1}{\sqrt{2}} [ s_x s_x + s_y s_y ] \end{aligned}$$

Clearly,  $1$  and  $1^{**}$ , which both contain 50% ionic and 50% covalent parts, combine to produce  $\psi_{-}$  which is purely covalent and  $\psi_{+}$  which is purely ionic.

The above strong CI mixing of  $^1$  and  $^1$  \*\* as  $R$  qualitatively alters the configuration correlation diagrams shown above. Descriptions of the resulting valence singlet and triplet states are given below for homonuclear situations in which covalent products lie below and above ionic products, respectively. Note that in both cases, there exists a single attractive curve and five (n.b., the triplet state has three curves superposed) repulsive curves.



State correlation diagram for homonuclear case in which homolytic bond cleavage is energetically favored.



State correlation diagram for homonuclear case in which heterolytic bond cleavage is energetically favored.

If the energies of the  $s_x$  and  $s_y$  orbitals do not differ significantly (compared to the coulombic interactions between electron pairs), it is expected that the essence of the findings described above for homonuclear species will persist even for heteronuclear systems. A decomposition of the six CSFs listed above, using the heteronuclear molecular orbitals introduced earlier yields:

$${}^1(0) = C^2 \left[ \begin{array}{cc} s_x & s_x \\ +z & s_x & s_y \\ & +z & s_y & s_x \end{array} \right]$$

$${}^1{}^{**}(0) = C^{*2} \left[ \begin{array}{cc} z^2 & s_x & s_x \\ -z & s_x & s_y \\ & -z & s_y & s_x \end{array} \right]$$

$${}^1{}^*(0) = \frac{CC^*}{\sqrt{2}} \left[ \begin{array}{cc} 2z & s_x & s_x \\ -2z & s_y & s_y \\ + (z^2 - 1) & s_y & s_x \\ & + (z^2 - 1) & s_x & s_y \end{array} \right]$$

$${}^3 \sigma_g^{*}(0) = \frac{CC^*}{\sqrt{2}} (z^2 + 1) [ s_y s_x - s_x s_y ]$$

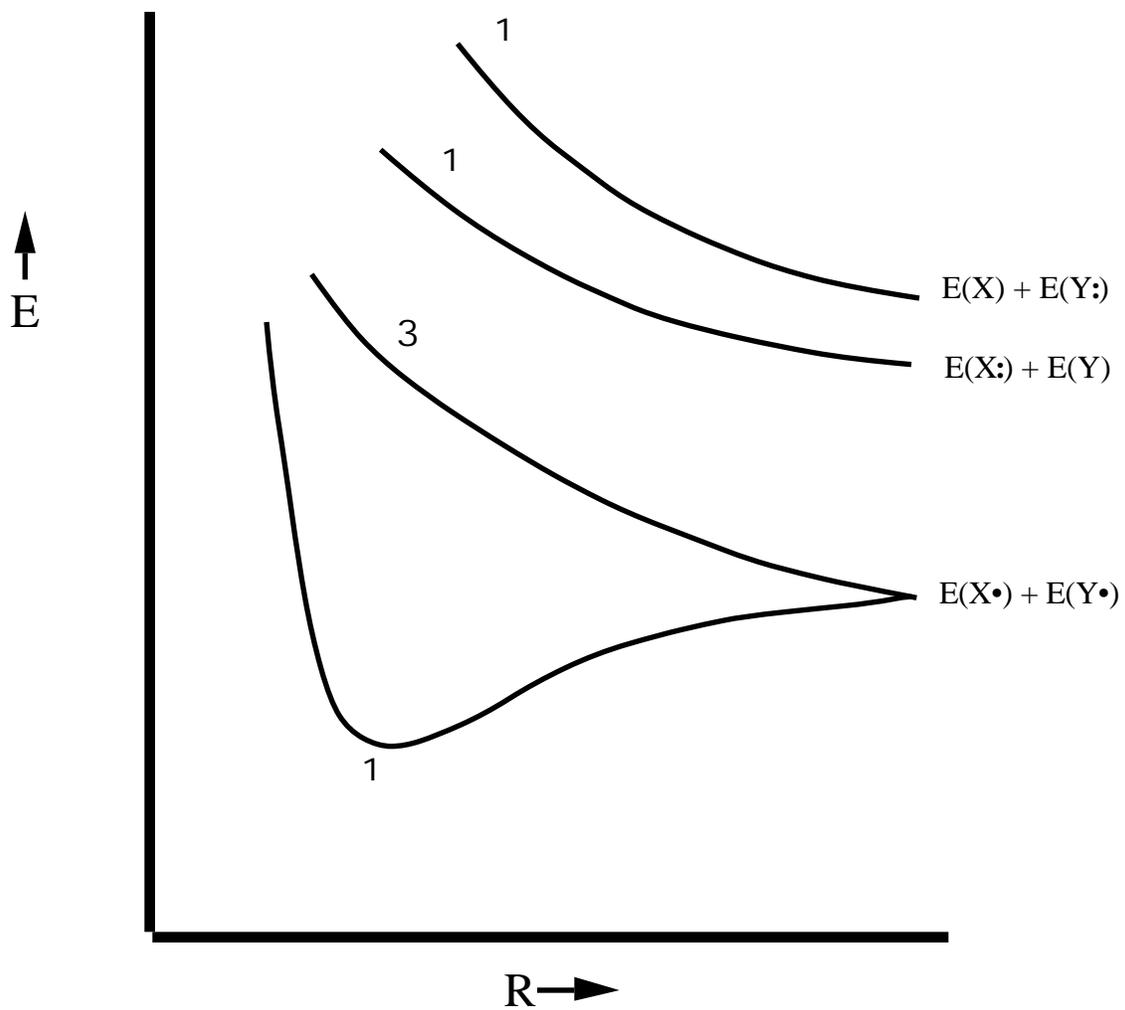
$${}^3 \sigma_g^{*}(1) = CC^* (z^2 + 1) s_y s_x$$

$${}^3 \sigma_g^{*}(-1) = CC^* (z^2 + 1) s_y s_x$$

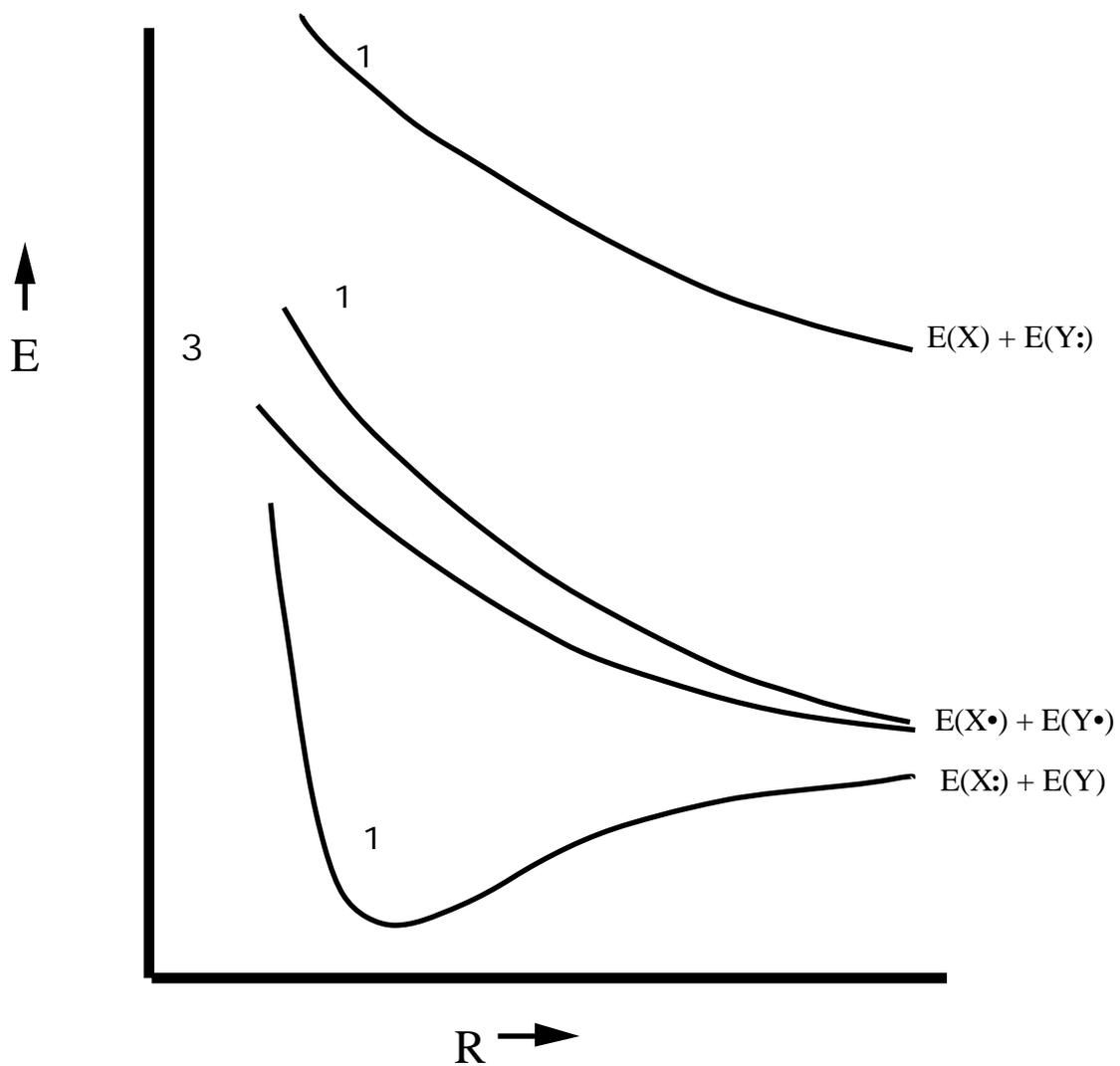
Clearly, the three  ${}^3 \sigma_g^{*}$  CSFs retain purely covalent R character even in the heteronuclear case. The  ${}^1 \sigma_g$ ,  ${}^1 \sigma_g^{**}$ , and  ${}^1 \sigma_g^{*}$  (all three of which can undergo CI mixing now) possess one covalent and two ionic components of the form  $s_x s_y + s_y s_x$ ,  $s_x s_x$ , and  $s_y s_y$ . The three singlet CSFs therefore can be combined to produce a singlet covalent product function  $s_x s_y + s_y s_x$  as well as both  $X + Y \cdot$  and  $X \cdot + Y$  ionic product wavefunctions

$s_y s_y$  and  $s_x s_x$ , respectively. In most situations, the energy ordering of the homolytic and heterolytic dissociation products will be either  $E(X\cdot) + E(Y\cdot) < E(X\cdot) + E(Y) < E(X) + E(Y\cdot)$  or  $E(X\cdot) + E(Y) < E(X\cdot) + E(Y\cdot) < E(X) + E(Y\cdot)$ .

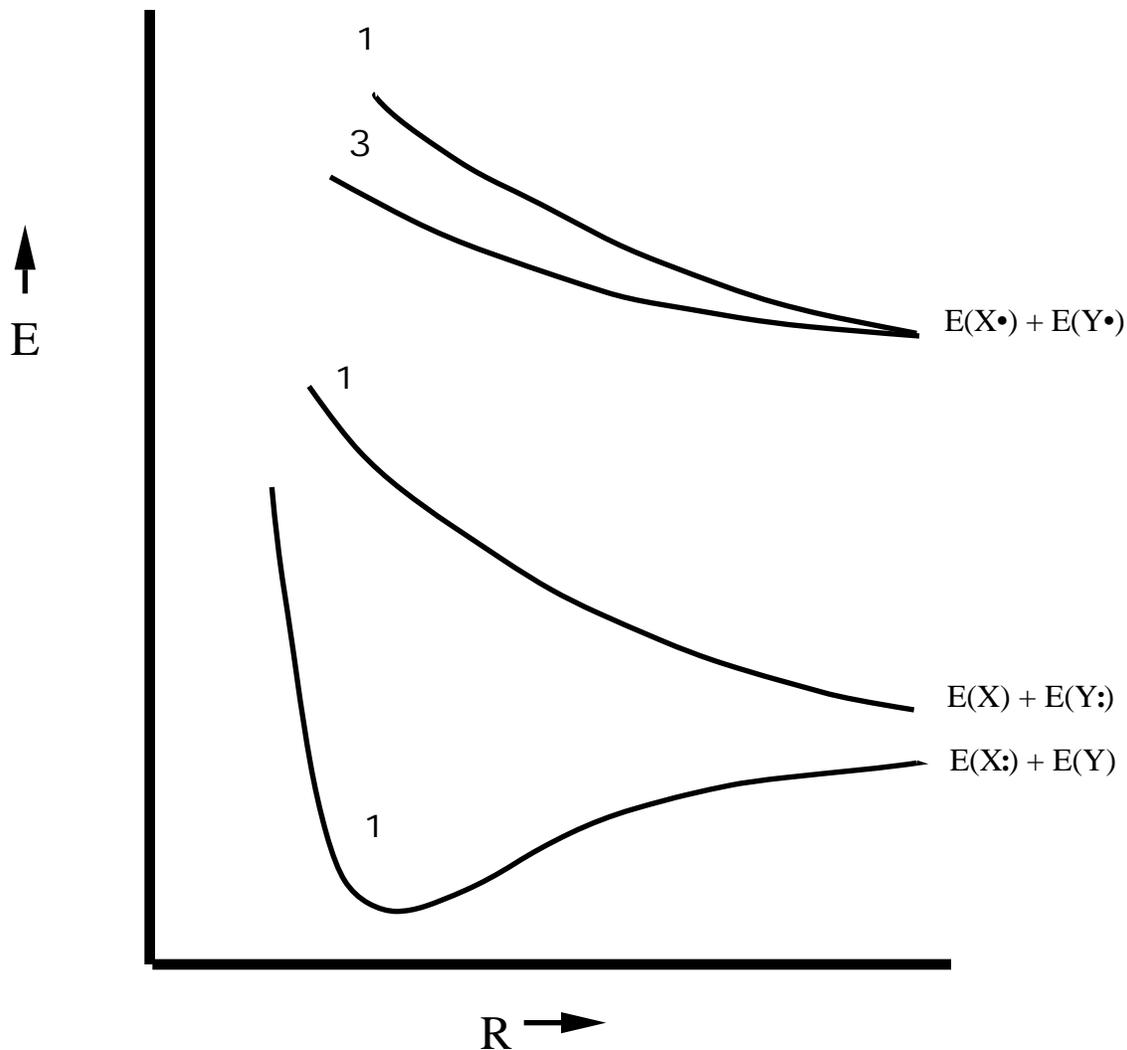
The extensions of the state correlation diagrams given above to the heteronuclear situations are described below.



State correlation diagram for heteronuclear case in which homolytic bond cleavage is energetically favored.



State correlation diagram for heteronuclear case in which heterolytic bond cleavage to one product is energetically favored but homolytic cleavage lies below the second heterolytic asymptote.



State correlation diagram for heteronuclear case in which both heterolytic bond cleavage products are energetically favored relative to homolytic cleavage.

Again note that only one curve is attractive and five are repulsive in all cases. In these heteronuclear cases, it is the mixing of the  $1$ ,  $1^*$ , and  $1^{**}$  CSFs, which varies with  $R$ , that determines which molecular state connects to which asymptote. As the energy ordering of the asymptotes varies, so do these correlations.

### 3. Summary

Even for the relatively simple two-electron, two-orbital single-bond interactions between a pair of atoms or functional groups, the correlations among energy-ordered molecular states and energy-ordered asymptotic states is complex enough to warrant considerations beyond what is taught in most undergraduate and beginning graduate inorganic and physical chemistry classes. In particular, the correlations that arise when one (or both) of the heterolytic bond dissociation asymptotes lies below the homolytic cleavage products are important to realize and keep in mind.

In all cases treated here, the three singlet states that arise produce one and only one attractive (bonding) potential energy curve; the other two singlet surfaces are repulsive. The three triplet surfaces are also repulsive. Of course, in arriving at these conclusions, we have considered only contributions to the inter-fragment interactions that arise from valence-orbital couplings; no consideration has been made of attractive or repulsive forces that result from one or both of the X- and Y- fragments possessing net charge. In the latter case, one must, of course, add to the qualitative potential surfaces described here any coulombic, charge-dipole, or charge-induced-dipole energies. Such additional factors can lead to attractive long-range interactions in typical ion-molecule complexes.

The necessity of the analysis developed above becomes evident when considering dissociation of diatomic transition metal ions. Most transition metal atoms have ground states with electron configurations of the form  $s^2d^n$  (for first-row metals, exceptions include Cr ( $s^1d^5$ ), Cu ( $s^1d^{10}$ ), and the  $s^1d^9$  state of Ni is basically isoenergetic with the  $s^2d^8$  ground state). The corresponding positive ions have ground states with  $s^1d^n$  (Sc, Ti, Mn, Fe) or  $s^0d^{n+1}$  (V, Cr, Co, Ni, Cu) electron configurations. For each of these elements, the alternate electron configuration leads to low-lying excited states.

One can imagine forming a  $M_2^+$  metal dimer ion with a configuration described as  $\sigma_g^2 d^{2n+1}$ , where the  $\sigma_g$  bonding orbital is formed primarily from the metal s orbitals and the d orbitals are largely nonbonding (as is particularly appropriate towards the right hand side of the periodic table). Cleavage of such a bond tends to occur heterolytically since this forms lower energy species,  $M(s^2d^n) + M^+(s^0d^{n+1})$ , than homolytic cleavage to  $M(s^1d^{n+1}) + M^+(s^1d^n)$ . For example,  $Co_2^+$  dissociates to  $Co(s^2d^7) + Co^+(s^0d^8)$  rather than to  $Co(s^1d^8) + Co^+(s^1d^7)$ ,<sup>2</sup> which lies 0.85 eV higher in energy.

Qualitative aspects of the above analysis for homonuclear transition metal dimer ions will persist for heteronuclear ions. For example, the ground-state dissociation asymptote for  $CoNi^+$  is the heterolytic cleavage products  $Co(s^2d^7) + Ni^+(s^0d^9)$ . The alternative heterolytic cleavage to form  $Co^+(s^0d^8) + Ni(s^2d^8)$  is 0.23 eV higher in energy, while homolytic cleavage can lead to  $Co^+(s^1d^7) + Ni(s^1d^9)$ , 0.45 eV higher, or  $Co(s^1d^8) + Ni^+(s^1d^8)$ , 1.47 eV higher. This is the situation illustrated in the last figure above.

### III. Various Types of Configuration Mixing

#### A. Essential CI

The above examples of the use of CCD's show that, as motion takes place along the proposed reaction path, geometries may be encountered at which it is essential to describe the electronic wavefunction in terms of a linear combination of more than one CSF:

$$\Psi = \sum_I C_I \Phi_I,$$

where the  $\Phi_I$  are the CSFs which are undergoing the avoided crossing. Such essential configuration mixing is often referred to as treating "essential CI".

#### B. Dynamical CI

To achieve reasonable chemical accuracy (e.g.,  $\pm 5$  kcal/mole) in electronic structure calculations it is necessary to use a multiconfigurational even in situations where no obvious strong configuration mixing (e.g., crossings of CSF energies) is present. For example, in describing the  $\sigma^2$  bonding electron pair of an olefin or the  $ns^2$  electron pair in alkaline earth atoms, it is important to mix in doubly excited CSFs of the form  $(\sigma^*)^2$  and  $np^2$ , respectively. The reasons for introducing such a CI-level treatment were treated for an alkaline earth atom earlier in this chapter.

Briefly, the physical importance of such doubly-excited CSFs can be made clear by using the identity:

$$C_1 | \dots \dots \rangle - C_2 | \dots \dots \rangle \\ = C_1/2 \{ | \dots (-x) \dots \rangle ( \dots +x \dots \rangle - | \dots (-x) \dots \rangle ( \dots +x \dots \rangle ) \},$$

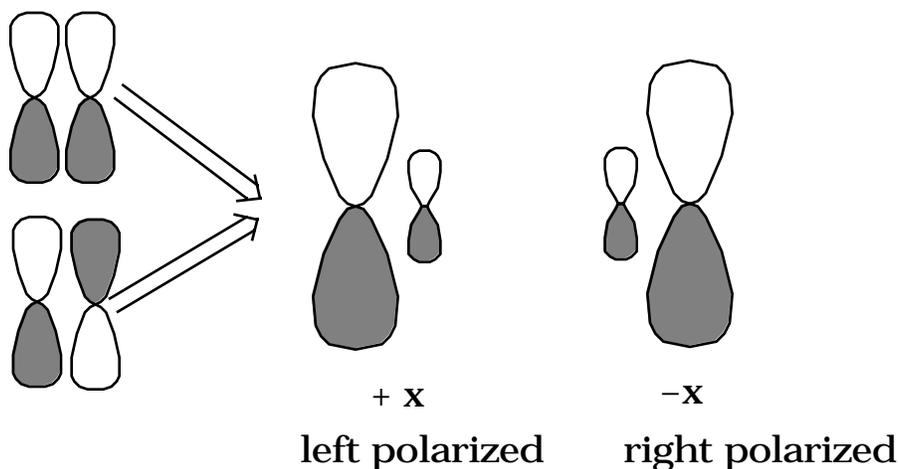
where

$$x = (C_2/C_1)^{1/2}.$$

This allows one to interpret the combination of two CSFs which differ from one another by a double excitation from one orbital ( $\sigma$ ) to another ( $\sigma^*$ ) as equivalent to a singlet coupling of

two different (non-orthogonal) orbitals ( $\psi_1$ ) and ( $\psi_2$ ). This picture is closely related to the so-called generalized valence bond (GVB) model that W. A. Goddard and his co-workers have developed (see, for example, W. A. Goddard and L. B. Harding, *Annu. Rev. Phys. Chem.* **29**, 363 (1978)). In the simplest embodiment of the GVB model, each electron pair in the atom or molecule is correlated by mixing in a CSF in which that electron pair is "doubly excited" to a correlating orbital. The direct product of all such pair correlations generates the GVB-type wavefunction. In the GVB approach, these electron correlations are not specified in terms of double excitations involving CSFs formed from orthonormal spin orbitals; instead, explicitly non-orthogonal GVB orbitals are used as described above, but the result is the same as one would obtain using the direct product of doubly excited CSFs.

In the olefin example mentioned above, the two non-orthogonal "polarized orbital pairs" involve mixing the  $\psi_1$  and  $\psi_2^*$  orbitals to produce two left-right polarized orbitals as depicted below:



In this case, one says that the  $\psi_1^2$  electron pair undergoes left-right correlation when the  $(\psi_1^* \psi_2)^2$  CSF is mixed into the CI wavefunction.

In the alkaline earth atom case, the polarized orbital pairs are formed by mixing the  $ns$  and  $np$  orbitals (actually, one must mix in equal amounts of  $p_1$ ,  $p_{-1}$ , and  $p_0$  orbitals to preserve overall  $^1S$  symmetry in this case), and give rise to angular correlation of the electron pair. Use of an  $(n+1)s^2$  CSF for the alkaline earth calculation would contribute in-out or radial correlation because, in this case, the polarized orbital pair formed from the  $ns$  and  $(n+1)s$  orbitals would be radially polarized.

The use of doubly excited CSFs is thus seen as a mechanism by which  $\psi_1$  can place electron pairs, which in the single-configuration picture occupy the same orbital, into

different regions of space (i.e., one into a member of the polarized orbital pair) thereby lowering their mutual coulombic repulsions. Such electron correlation effects are referred to as "dynamical electron correlation"; they are extremely important to include if one expects to achieve chemically meaningful accuracy (i.e.,  $\pm 5$  kcal/mole).

## Section 4 Molecular Rotation and Vibration

### Chapter 13

*Treating the full internal nuclear-motion dynamics of a polyatomic molecule is complicated. It is conventional to examine the rotational movement of a hypothetical "rigid" molecule as well as the vibrational motion of a non-rotating molecule, and to then treat the rotation-vibration couplings using perturbation theory.*

#### I. Rotational Motions of Rigid Molecules

In Chapter 3 and Appendix G the energy levels and wavefunctions that describe the rotation of rigid molecules are described. Therefore, in this Chapter these results will be summarized briefly and emphasis will be placed on detailing how the corresponding rotational Schrödinger equations are obtained and the assumptions and limitations underlying them.

##### A. Linear Molecules

###### 1. The Rotational Kinetic Energy Operator

As given in Chapter 3, the Schrödinger equation for the angular motion of a rigid (i.e., having fixed bond length  $R$ ) diatomic molecule is

$$\frac{\hbar^2}{2\mu} \left\{ (R^2 \sin^2 \theta)^{-1} \frac{\partial^2}{\partial \phi^2} + (R^2 \sin^2 \theta)^{-1} \frac{\partial^2}{\partial \theta^2} \right\} \psi = E \psi$$

or

$$L^2 / 2\mu R^2 = E \quad .$$

The Hamiltonian in this problem contains only the kinetic energy of rotation; no potential energy is present because the molecule is undergoing unhindered "free rotation". The angles  $\theta$  and  $\phi$  describe the orientation of the diatomic molecule's axis relative to a laboratory-fixed coordinate system, and  $\mu$  is the reduced mass of the diatomic molecule  $\mu = m_1 m_2 / (m_1 + m_2)$ .

###### 2. The Eigenfunctions and Eigenvalues

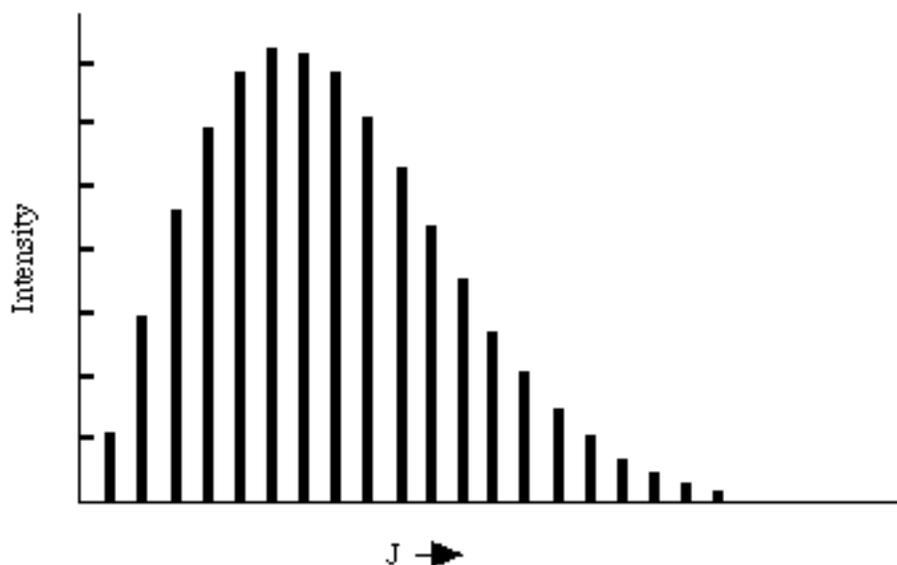
The eigenvalues corresponding to each eigenfunction are straightforward to find because  $H_{\text{rot}}$  is proportional to the  $L^2$  operator whose eigenvalues have already been determined. The resultant rotational energies are given as:

$$E_J = \hbar^2 J(J+1)/(2\mu R^2) = B J(J+1)$$

and are independent of  $M$ . Thus each energy level is labeled by  $J$  and is  $2J+1$ -fold degenerate (because  $M$  ranges from  $-J$  to  $J$ ). The rotational constant  $B$  (defined as  $\hbar^2/2\mu R^2$ ) depends on the molecule's bond length and reduced mass. Spacings between successive rotational levels (which are of spectroscopic relevance because angular momentum selection rules often restrict  $J$  to  $1,0$ , and  $-1$ ) are given by

$$E = B (J+1)(J+2) - B J(J+1) = 2B (J+1).$$

Within this "rigid rotor" model, the absorption spectrum of a rigid diatomic molecule should display a series of peaks, each of which corresponds to a specific  $J \Rightarrow J + 1$  transition. The energies at which these peaks occur should grow linearly with  $J$ . An example of such a progression of rotational lines is shown in the figure below.



The energies at which the rotational transitions occur appear to fit the  $E = 2B (J+1)$  formula rather well. The intensities of transitions from level  $J$  to level  $J+1$  vary strongly with  $J$  primarily because the population of molecules in the absorbing level varies with  $J$ .

These populations  $P_J$  are given, when the system is at equilibrium at temperature  $T$ , in terms of the degeneracy  $(2J+1)$  of the  $J^{\text{th}}$  level and the energy of this level  $B J(J+1)$  :

$$P_J = Q^{-1} (2J+1) \exp(-BJ(J+1)/kT),$$

where  $Q$  is the rotational partition function:

$$Q = \sum_J (2J+1) \exp(-BJ(J+1)/kT).$$

For low values of  $J$ , the degeneracy is low and the  $\exp(-BJ(J+1)/kT)$  factor is near unity. As  $J$  increases, the degeneracy grows linearly but the  $\exp(-BJ(J+1)/kT)$  factor decreases more rapidly. As a result, there is a value of  $J$ , given by taking the derivative of  $(2J+1) \exp(-BJ(J+1)/kT)$  with respect to  $J$  and setting it equal to zero,

$$2J_{\text{max}} + 1 = \sqrt{2kT/B}$$

at which the intensity of the rotational transition is expected to reach its maximum.

The eigenfunctions belonging to these energy levels are the spherical harmonics  $Y_{L,M}(\theta, \phi)$  which are normalized according to

$$\int_0^{2\pi} \int_0^\pi (Y_{L,M}^*(\theta, \phi) Y_{L',M'}(\theta, \phi) \sin \theta d\theta d\phi) = \delta_{L,L'} \delta_{M,M'}.$$

These functions are identical to those that appear in the solution of the angular part of Hydrogen-like atoms. The above energy levels and eigenfunctions also apply to the rotation of rigid linear polyatomic molecules; the only difference is that the moment of inertia  $I$  entering into the rotational energy expression is given by

$$I = \sum_a m_a R_a^2$$

where  $m_a$  is the mass of the  $a^{\text{th}}$  atom and  $R_a$  is its distance from the center of mass of the molecule. This moment of inertia replaces  $\mu R^2$  in the earlier rotational energy level expressions.

## B. Non-Linear Molecules

### 1. The Rotational Kinetic Energy Operator

The rotational kinetic energy operator for a rigid polyatomic molecule is shown in Appendix G to be

$$H_{\text{rot}} = J_a^2/2I_a + J_b^2/2I_b + J_c^2/2I_c$$

where the  $I_k$  ( $k = a, b, c$ ) are the three principal moments of inertia of the molecule (the eigenvalues of the moment of inertia tensor). This tensor has elements in a Cartesian coordinate system ( $K, K' = X, Y, Z$ ) whose origin is located at the center of mass of the molecule that are computed as:

$$I_{K,K} = \sum_j m_j (R_j^2 - R_{K,j}^2) \quad (\text{for } K = K')$$

$$I_{K,K'} = - \sum_j m_j R_{K,j} R_{K',j} \quad (\text{for } K \neq K').$$

The components of the quantum mechanical angular momentum operators along the three principal axes are:

$$J_a = -i\hbar \cos \theta \left[ \cot \theta \frac{\partial}{\partial \phi} - (\sin \theta)^{-1} \frac{\partial}{\partial \alpha} \right] - i\hbar \sin \theta \frac{\partial}{\partial \beta}$$

$$J_b = i\hbar \sin \theta \left[ \cot \theta \frac{\partial}{\partial \phi} - (\sin \theta)^{-1} \frac{\partial}{\partial \alpha} \right] - i\hbar \cos \theta \frac{\partial}{\partial \beta}$$

$$J_c = -i\hbar \frac{\partial}{\partial \phi}$$

The angles  $\alpha$ ,  $\beta$ , and  $\phi$  are the Euler angles needed to specify the orientation of the rigid molecule relative to a laboratory-fixed coordinate system. The corresponding square of the total angular momentum operator  $J^2$  can be obtained as

$$J^2 = J_a^2 + J_b^2 + J_c^2$$

$$= -\frac{2}{\sin^2 \theta} - \cot^2 \theta /$$

$$- (1/\sin^2 \theta) ( \frac{2}{\sin^2 \theta} + \frac{2}{\sin^2 \theta} - 2 \cos^2 \theta / ),$$

and the component along the lab-fixed Z axis  $J_Z$  is  $-\hbar^2 /$ .

## 2. The Eigenfunctions and Eigenvalues for Special Cases

### a. Spherical Tops

When the three principal moment of inertia values are identical, the molecule is termed a **spherical top**. In this case, the total rotational energy can be expressed in terms of the total angular momentum operator  $J^2$

$$H_{\text{rot}} = J^2/2I.$$

As a result, the eigenfunctions of  $H_{\text{rot}}$  are those of  $J^2$  (and  $J_a$  as well as  $J_Z$  both of which commute with  $J^2$  and with one another;  $J_Z$  is the component of  $\mathbf{J}$  along the lab-fixed Z-axis and commutes with  $J_a$  because  $J_Z = -\hbar^2 /$  and  $J_a = -\hbar^2 /$  act on different angles).

The energies associated with such eigenfunctions are

$$E(J,K,M) = \hbar^2 J(J+1)/2I^2,$$

for all K (i.e.,  $J_a$  quantum numbers) ranging from  $-J$  to  $J$  in unit steps and for all M (i.e.,  $J_Z$  quantum numbers) ranging from  $-J$  to  $J$ . Each energy level is therefore  $(2J + 1)^2$  degenerate because there are  $2J + 1$  possible K values and  $2J + 1$  possible M values for each J.

The eigenfunctions of  $J^2$ ,  $J_Z$  and  $J_a$ ,  $|J,M,K\rangle$  are given in terms of the set of rotation matrices  $D_{J,M,K}$ :

$$|J,M,K\rangle = \sqrt{\frac{2J+1}{8}} D_{J,M,K}^*(\alpha, \beta, \gamma)$$

which obey

$$J^2 |J,M,K\rangle = \hbar^2 J(J+1) |J,M,K\rangle,$$

$$J_a |J,M,K\rangle = \hbar K |J,M,K\rangle,$$

$$J_Z |J,M,K\rangle = \hbar M |J,M,K\rangle.$$

### b. Symmetric Tops

Molecules for which two of the three principal moments of inertia are equal are called **symmetric tops**. Those for which the unique moment of inertia is smaller than the other two are termed **prolate** symmetric tops; if the unique moment of inertia is larger than the others, the molecule is an **oblate** symmetric top.

Again, the rotational kinetic energy, which is the full rotational Hamiltonian, can be written in terms of the total rotational angular momentum operator  $J^2$  and the component of angular momentum along the axis with the unique principal moment of inertia:

$$H_{\text{rot}} = J^2/2I + J_a^2\{1/2I_a - 1/2I\}, \text{ for prolate tops}$$

$$H_{\text{rot}} = J^2/2I + J_c^2\{1/2I_c - 1/2I\}, \text{ for oblate tops.}$$

As a result, the eigenfunctions of  $H_{\text{rot}}$  are those of  $J^2$  and  $J_a$  or  $J_c$  (and of  $J_Z$ ), and the corresponding energy levels are:

$$E(J,K,M) = \hbar^2 J(J+1)/2I^2 + \hbar^2 K^2 \{1/2I_a - 1/2I\},$$

for prolate tops

$$E(J,K,M) = \hbar^2 J(J+1)/2I^2 + \hbar^2 K^2 \{1/2I_c - 1/2I\},$$

for oblate tops, again for  $K$  and  $M$  (i.e.,  $J_a$  or  $J_c$  and  $J_Z$  quantum numbers, respectively) ranging from  $-J$  to  $J$  in unit steps. Since the energy now depends on  $K$ , these levels are only  $2J + 1$  degenerate due to the  $2J + 1$  different  $M$  values that arise for each  $J$  value. The eigenfunctions  $|J, M, K\rangle$  are the same rotation matrix functions as arise for the spherical-top case.

### c. Asymmetric Tops

The rotational eigenfunctions and energy levels of a molecule for which all three principal moments of inertia are distinct (a so-called **asymmetric top**) can not easily be expressed in terms of the angular momentum eigenstates and the J, M, and K quantum numbers. However, given the three principal moments of inertia  $I_a$ ,  $I_b$ , and  $I_c$ , a matrix representation of each of the three contributions to the rotational Hamiltonian

$$H_{\text{rot}} = J_a^2/2I_a + J_b^2/2I_b + J_c^2/2I_c$$

can be formed within a basis set of the  $\{|J, M, K\rangle\}$  rotation matrix functions. This matrix will not be diagonal because the  $|J, M, K\rangle$  functions are not eigenfunctions of the asymmetric top  $H_{\text{rot}}$ . However, the matrix can be formed in this basis and subsequently brought to diagonal form by finding its eigenvectors  $\{C_{n, J, M, K}\}$  and its eigenvalues  $\{E_n\}$ . The vector coefficients express the asymmetric top eigenstates as

$$|n\rangle = \sum_{J, M, K} C_{n, J, M, K} |J, M, K\rangle.$$

Because the total angular momentum  $J^2$  still commutes with  $H_{\text{rot}}$ , each such eigenstate will contain only one J-value, and hence  $|n\rangle$  can also be labeled by a J quantum number:

$$|n, J\rangle = \sum_{M, K} C_{n, J, M, K} |J, M, K\rangle.$$

To form the only non-zero matrix elements of  $H_{\text{rot}}$  within the  $|J, M, K\rangle$  basis, one can use the following properties of the rotation-matrix functions (see, for example, Zare's book on Angular Momentum):

$$\langle J, M, K | J_a^2 | J, M, K \rangle = \langle J, M, K | J_b^2 | J, M, K \rangle$$

$$= 1/2 \langle J, M, K | J^2 - J_c^2 | J, M, K \rangle = \hbar^2 [J(J+1) - K^2],$$

$$\langle J, M, K | J_c^2 | J, M, K \rangle = \hbar^2 K^2,$$

$$\langle J, M, K | J_a^2 | J, M, K \pm 2 \rangle = - \langle J, M, K | J_b^2 | J, M, K \pm 2 \rangle$$

$$= \hbar^2 [J(J+1) - K(K \pm 1)]^{1/2} [J(J+1) - (K \pm 1)(K \pm 2)]^{1/2}$$

$$\langle J, M, K | J_c^2 | J, M, K \pm 2 \rangle = 0.$$

Each of the elements of  $J_c^2$ ,  $J_a^2$ , and  $J_b^2$  must, of course, be multiplied, respectively, by  $1/2I_c$ ,  $1/2I_a$ , and  $1/2I_b$  and summed together to form the matrix representation of  $H_{\text{rot}}$ . The diagonalization of this matrix then provides the asymmetric top energies and wavefunctions.

## II. Vibrational Motion Within the Harmonic Approximation

The simple harmonic motion of a diatomic molecule was treated in Chapter 1, and will not be repeated here. Instead, emphasis is placed on polyatomic molecules whose electronic energy's dependence on the  $3N$  Cartesian coordinates of its  $N$  atoms can be written (approximately) in terms of a Taylor series expansion about a stable local minimum. We therefore assume that the molecule of interest exists in an electronic state for which the geometry being considered is stable (i.e., not subject to spontaneous geometrical distortion).

The Taylor series expansion of the electronic energy is written as:

$$V(q_k) = V(0) + \sum_k \left( \frac{\partial V}{\partial q_k} \right) q_k + \frac{1}{2} \sum_{j,k} q_j H_{j,k} q_k + \dots$$

where  $V(0)$  is the value of the electronic energy at the stable geometry under study,  $q_k$  is the displacement of the  $k_{\text{th}}$  Cartesian coordinate away from this starting position,  $\left( \frac{\partial V}{\partial q_k} \right)$  is the gradient of the electronic energy along this direction, and the  $H_{j,k}$  are the second derivative or Hessian matrix elements along these directions  $H_{j,k} = \left( \frac{\partial^2 V}{\partial q_j \partial q_k} \right)$ . If the starting geometry corresponds to a stable species, the gradient terms will all vanish (meaning this geometry corresponds to a minimum, maximum, or saddle point), and the Hessian matrix will possess  $3N - 5$  (for linear species) or  $3N - 6$  (for non-linear molecules) positive eigenvalues and 5 or 6 zero eigenvalues (corresponding to 3 translational and 2 or 3 rotational motions of the molecule). If the Hessian has one negative eigenvalue, the geometry corresponds to a transition state (these situations are discussed in detail in Chapter 20).

From now on, we assume that the geometry under study corresponds to that of a stable minimum about which vibrational motion occurs. The treatment of unstable geometries is of great importance to chemistry, but this Chapter deals with vibrations of stable species. For a good treatment of situations under which geometrical instability is expected to occur, see Chapter 2 of the text Energetic Principles of Chemical Reactions by

J. Simons. A discussion of how local minima and transition states are located on electronic energy surfaces is provided in Chapter 20 of the present text.

## A. The Newton Equations of Motion for Vibration

### 1. The Kinetic and Potential Energy Matrices

Truncating the Taylor series at the quadratic terms (assuming these terms dominate because only small displacements from the equilibrium geometry are of interest), one has the so-called **harmonic** potential:

$$V(q_k) = V(0) + \frac{1}{2} \sum_{j,k} q_j H_{j,k} q_k.$$

The classical mechanical equations of motion for the  $3N$   $\{q_k\}$  coordinates can be written in terms of the above potential energy and the following kinetic energy function:

$$T = \frac{1}{2} \sum_j m_j \dot{q}_j^2,$$

where  $\dot{q}_j$  denotes the time rate of change of the coordinate  $q_j$  and  $m_j$  is the mass of the atom on which the  $j^{\text{th}}$  Cartesian coordinate resides. The Newton equations thus obtained are:

$$m_j \ddot{q}_j = - \sum_k H_{j,k} q_k$$

where the force along the  $j^{\text{th}}$  coordinate is given by minus the derivative of the potential  $V$  along this coordinate ( $-\partial V / \partial q_j = - \sum_k H_{j,k} q_k$ ) within the harmonic approximation.

These classical equations can more compactly be expressed in terms of the time evolution of a set of so-called **mass weighted** Cartesian coordinates defined as:

$$x_j = q_j (m_j)^{1/2},$$

in terms of which the Newton equations become

$$\ddot{x}_j = - \sum_k H'_{j,k} x_k$$

and the mass-weighted Hessian matrix elements are

$$H'_{j,k} = H_{j,k} (m_j m_k)^{-1/2}.$$

## 2. The Harmonic Vibrational Energies and Normal Mode Eigenvectors

Assuming that the  $x_j$  undergo some form of sinusoidal time evolution:

$$x_j(t) = x_j(0) \cos(\omega_j t),$$

and substituting this into the Newton equations produces a matrix eigenvalue equation:

$$-\omega_j^2 x_j = -\sum_k H'_{j,k} x_k$$

in which the eigenvalues are the squares of the so-called **normal mode** vibrational frequencies and the eigenvectors give the amplitudes of motion along each of the  $3N$  mass weighted Cartesian coordinates that belong to each mode.

Within this harmonic treatment of vibrational motion, the total vibrational energy of the molecule is given as

$$E(v_1, v_2, \dots, v_{3N-5 \text{ or } 6}) = \sum_{j=1}^{3N-5 \text{ or } 6} \hbar \omega_j (v_j + 1/2)$$

as sum of  $3N-5$  or  $3N-6$  independent contributions one for each normal mode. The corresponding total vibrational wavefunction

$$\psi(x_1, x_2, \dots, x_{3N-5 \text{ or } 6}) = \prod_j \psi_{v_j}(x_j)$$

as a product of  $3N-5$  or  $3N-6$  harmonic oscillator functions  $\psi_{v_j}(x_j)$  are for each normal mode within this picture, the energy gap between one vibrational level and another in which one of the  $v_j$  quantum numbers is increased by unity (the origin of this "selection rule" is discussed in Chapter 15) is

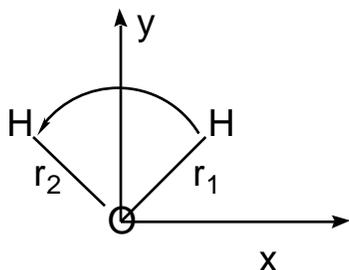
$$E_{v_j+1} - E_{v_j} = \hbar \omega_j$$

The harmonic model thus predicts that the "fundamental" ( $v=0 \rightarrow v=1$ ) and "hot band" ( $v=1 \rightarrow v=2$ ) transition should occur at the same energy, and the overtone ( $v=0 \rightarrow v=2$ ) transitions should occur at exactly twice this energy.

## B. The Use of Symmetry

### 1. Symmetry Adapted Modes

It is often possible to simplify the calculation of the normal mode frequencies and eigenvectors by exploiting molecular point group symmetry. For molecules that possess symmetry, the electronic potential  $V(q_j)$  displays symmetry with respect to displacements of symmetry equivalent Cartesian coordinates. For example, consider the water molecule at its  $C_{2v}$  equilibrium geometry as illustrated in the figure below. A very small movement of the  $H_2O$  molecule's left H atom in the positive x direction ( $x_L$ ) produces the same change in  $V$  as a correspondingly small displacement of the right H atom in the negative x direction ( $-x_R$ ). Similarly, movement of the left H in the positive y direction ( $y_L$ ) produces an energy change identical to movement of the right H in the positive y direction ( $y_R$ ).



The equivalence of the pairs of Cartesian coordinate displacements is a result of the fact that the displacement vectors are connected by the point group operations of the  $C_{2v}$  group. In particular, reflection of  $x_L$  through the  $yz$  plane produces  $-x_R$ , and reflection of  $y_L$  through this same plane yields  $y_R$ .

More generally, it is possible to combine sets of Cartesian displacement coordinates  $\{q_k\}$  into so-called symmetry adapted coordinates  $\{Q_{\alpha j}\}$ , where the index  $\alpha$  labels the irreducible representation and  $j$  labels the particular combination of that symmetry. These symmetry adapted coordinates can be formed by applying the point group projection operators to the individual Cartesian displacement coordinates.

To illustrate, again consider the H<sub>2</sub>O molecule in the coordinate system described above. The 3N = 9 mass weighted Cartesian displacement coordinates (X<sub>L</sub>, Y<sub>L</sub>, Z<sub>L</sub>, X<sub>O</sub>, Y<sub>O</sub>, Z<sub>O</sub>, X<sub>R</sub>, Y<sub>R</sub>, Z<sub>R</sub>) can be symmetry adapted by applying the following four projection operators:

$$P_{A_1} = 1 + yz + xy + C_2$$

$$P_{b_1} = 1 + yz - xy - C_2$$

$$P_{b_2} = 1 - yz + xy - C_2$$

$$P_{a_2} = 1 - yz - xy + C_2$$

to each of the 9 original coordinates. Of course, one will not obtain 9 x 4 = 36 independent symmetry adapted coordinates in this manner; many identical combinations will arise, and only 9 will be independent.

The independent combination of a<sub>1</sub> symmetry (normalized to produce vectors of unit length) are

$$Q_{a_1,1} = 2^{-1/2} [X_L - X_R]$$

$$Q_{a_1,2} = 2^{-1/2} [Y_L + Y_R]$$

$$Q_{a_1,3} = [Y_O]$$

Those of b<sub>2</sub> symmetry are

$$Q_{b_2,1} = 2^{-1/2} [X_L + X_R]$$

$$Q_{b_2,2} = 2^{-1/2} [Y_L - Y_R]$$

$$Q_{b_2,3} = [X_O],$$

and the combinations

$$Q_{b_1,1} = 2^{-1/2} [Z_L + Z_R]$$

$$Q_{b_1,2} = [Z_O]$$

are of  $b_1$  symmetry, whereas

$$Q_{a_2,1} = 2^{-1/2} [Z_L - Z_R]$$

is of  $a_2$  symmetry.

## 2. Point Group Symmetry of the Harmonic Potential

These nine  $Q_{j,j}$  are expressed as unitary transformations of the original mass weighted Cartesian coordinates:

$$Q_{j,j} = \sum_k C_{j,k} X_k$$

These transformation coefficients  $\{C_{j,k}\}$  can be used to carry out a unitary transformation of the 9x9 mass-weighted Hessian matrix. In so doing, we need only form blocks

$$H_{j,l} = \sum_k \sum_{k'} C_{j,k} H_{k,k'} (m_k m_{k'})^{-1/2} C_{l,k'}$$

within which the symmetries of the two modes are identical. The off-diagonal elements

$$H_{j,l} = \sum_k \sum_{k'} C_{j,k} H_{k,k'} (m_k m_{k'})^{-1/2} C_{l,k'}$$

vanish because the potential  $V(q_j)$  (and the full vibrational Hamiltonian  $H = T + V$ ) commutes with the  $C_{2v}$  point group symmetry operations.

As a result, the 9x9 mass-weighted Hessian eigenvalue problem can be subdivided into two 3x3 matrix problems (of  $a_1$  and  $b_2$  symmetry), one 2x2 matrix of  $b_1$  symmetry

and one 1x1 matrix of  $a_2$  symmetry. For example, the  $a_1$  symmetry block  $H_{j,1}^{a_1}$  is formed as follows:

$$\begin{array}{ccccccc}
 \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 & & & & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\
 & m^{-1}H \frac{2v}{x_L^2} & & m^{-1}H \frac{2v}{x_L x_R} & & (m_H m_O)^{-1/2} \frac{2v}{x_L x_O} & & & \\
 \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & m^{-1}H \frac{2v}{x_R x_L} & & m^{-1}H \frac{2v}{x_R^2} & & (m_H m_O)^{-1/2} \frac{2v}{x_R x_O} & \\
 & & & (m_H m_O)^{-1/2} \frac{2v}{x_O x_L} & & (m_H m_O)^{-1/2} \frac{2v}{x_O x_R} & & m^{-1}O \frac{2v}{x_O^2} & \\
 0 & 0 & 1 & & & & & & 0 & 0 & 1
 \end{array}$$

The  $b_2$ ,  $b_1$  and  $a_2$  blocks are formed in a similar manner. The eigenvalues of each of these blocks provide the squares of the harmonic vibrational frequencies, the eigenvectors provide the normal mode displacements as linear combinations of the symmetry adapted  $\{Q_j\}$ .

Regardless of whether symmetry is used to block diagonalize the mass-weighted Hessian, six (for non-linear molecules) or five (for linear species) of the eigenvalues will equal zero. The eigenvectors belonging to these zero eigenvalues describe the 3 translations and 2 or 3 rotations of the molecule. For example,

$$\frac{1}{\sqrt{3}} [X_L + X_R + X_O]$$

$$\frac{1}{\sqrt{3}} [Y_L + Y_R + Y_O]$$

$$\frac{1}{\sqrt{3}} [Z_L + Z_R + Z_O]$$

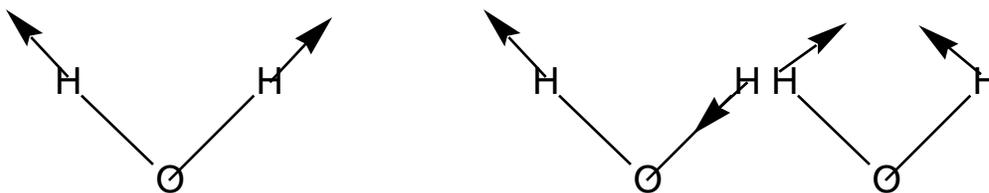
are three translation eigenvectors of  $b_2$ ,  $a_1$  and  $b_1$  symmetry, and

$$\frac{1}{\sqrt{2}} (Z_L - Z_R)$$

is a rotation (about the Y-axis in the figure shown above) of  $a_2$  symmetry. This rotation vector can be generated by applying the  $a_2$  projection operator to  $Z_L$  or to  $Z_R$ . The fact that rotation about the Y-axis is of  $a_2$  symmetry is indicated in the right-hand column of the  $C_{2v}$

character table of Appendix E via the symbol  $R_Z$  (n.b., care must be taken to realize that the axis convention used in the above figure is different than that implied in the character table; the latter has the  $Z$ -axis out of the molecular plane, while the figure calls this the  $X$ -axis). The other two rotations are of  $b_1$  and  $b_2$  symmetry (see the  $C_{2v}$  character table of Appendix E) and involve spinning of the molecule about the  $X$ - and  $Z$ - axes of the figure drawn above, respectively.

So, of the 9 cartesian displacements, 3 are of  $a_1$  symmetry, 3 of  $b_2$ , 2 of  $b_1$ , and 1 of  $a_2$ . Of these, there are three translations ( $a_1$ ,  $b_2$ , and  $b_1$ ) and three rotations ( $b_2$ ,  $b_1$ , and  $a_2$ ). This leaves two vibrations of  $a_1$  and one of  $b_2$  symmetry. For the  $H_2O$  example treated here, the three non zero eigenvalues of the mass-weighted Hessian are therefore of  $a_1$ ,  $b_2$ , and  $a_1$  symmetry. They describe the symmetric and asymmetric stretch vibrations and the bending mode, respectively as illustrated below.



The method of vibrational analysis presented here can work for any polyatomic molecule. One knows the mass-weighted Hessian and then computes the non-zero eigenvalues which then provide the squares of the normal mode vibrational frequencies. Point group symmetry can be used to block diagonalize this Hessian and to label the vibrational modes according to symmetry.

### III. Anharmonicity

The electronic energy of a molecule, ion, or radical at geometries near a stable structure can be expanded in a Taylor series in powers of displacement coordinates as was done in the preceding section of this Chapter. This expansion leads to a picture of uncoupled harmonic vibrational energy levels

$$E(v_1 \cdots v_{3N-5 \text{ or } 6}) = \prod_{j=1}^{3N-5 \text{ or } 6} \hbar \omega_j (v_j + 1/2)$$

and wavefunctions

$$\psi(x_1 \cdots x_{3N-5 \text{ or } 6}) = \prod_{j=1}^{3N-5 \text{ or } 6} \psi_j(x_j).$$

The spacing between energy levels in which one of the normal-mode quantum numbers increases by unity

$$E_{v_j} = E(\cdots v_j+1 \cdots) - E(\cdots v_j \cdots) = \hbar \omega_j$$

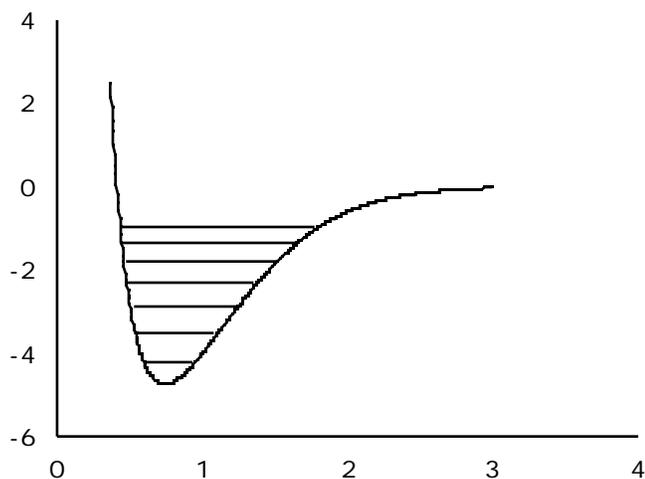
is predicted to be independent of the quantum number  $v_j$ . This picture of evenly spaced energy levels

$$E_0 = E_1 = E_2 = \cdots$$

is an incorrect aspect of the harmonic model of vibrational motion, and is a result of the quadratic model for the potential energy surface  $V(x_j)$ .

#### A. The Expansion of $E(v)$ in Powers of $(v+1/2)$ .

Experimental evidence clearly indicates that significant deviations from the harmonic oscillator energy expression occur as the quantum number  $v_j$  grows. In Chapter 1 these deviations were explained in terms of the diatomic molecule's true potential  $V(R)$  deviating strongly from the harmonic  $1/2k (R-R_e)^2$  potential at higher energy (and hence larger  $R-R_e$ ) as shown in the following figure.



At larger bond lengths, the true potential is "softer" than the harmonic potential, and eventually reaches its asymptote which lies at the dissociation energy  $D_e$  above its minimum. This negative deviation of the true  $V(R)$  from  $1/2 k(R-R_e)^2$  causes the true vibrational energy levels to lie below the harmonic predictions.

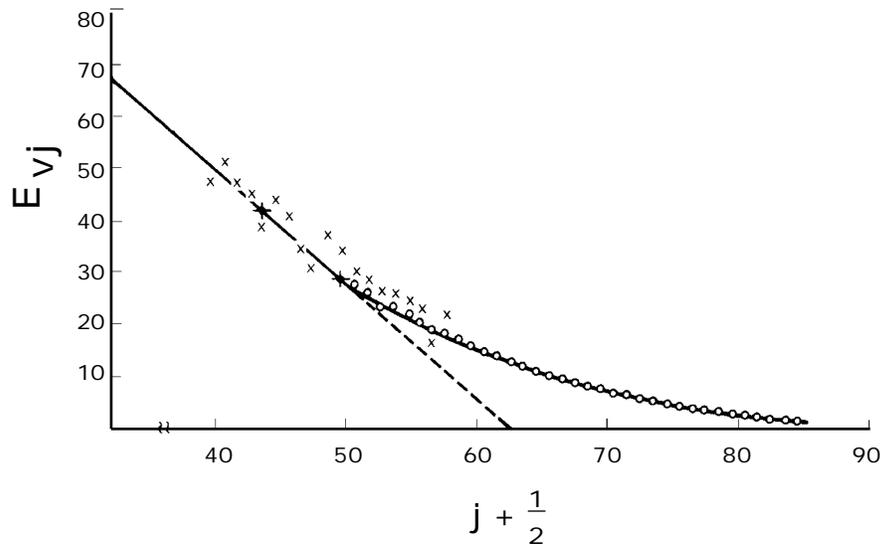
It is convention to express the experimentally observed vibrational energy levels, along each of the  $3N-5$  or  $6$  independent modes, as follows:

$$E(v_j) = \hbar \left[ \omega_j (v_j + 1/2) - (x_j) \omega_j (v_j + 1/2)^2 + (y_j) \omega_j (v_j + 1/2)^3 + (z_j) \omega_j (v_j + 1/2)^4 + \dots \right]$$

The first term is the harmonic expression. The next is termed the first anharmonicity; it (usually) produces a negative contribution to  $E(v_j)$  that varies as  $(v_j + 1/2)^2$ . The spacings between successive  $v_j$  and  $v_j + 1$  energy levels is then given by:

$$\begin{aligned} E_{v_j} &= E(v_j + 1) - E(v_j) \\ &= \hbar \left[ \omega_j - 2(x_j) \omega_j (v_j + 1) + \dots \right] \end{aligned}$$

A plot of the spacing between neighboring energy levels versus  $v_j$  should be linear for values of  $v_j$  where the harmonic and first overtone terms dominate. The slope of such a plot is expected to be  $-2\hbar(x_j)\omega_j$  and the small  $-v_j$  intercept should be  $\hbar[\omega_j - 2(x_j)\omega_j]$ . Such a plot of experimental data, which clearly can be used to determine the  $\omega_j$  and  $(x_j)$  parameter of the vibrational mode of study, is shown in the figure below.



## B. The Birge-Sponer Extrapolation

These so-called Birge-Sponer plots can also be used to determine dissociation energies of molecules. By linearly extrapolating the plot of experimental  $E_{v_j}$  values to large  $v_j$  values, one can find the value of  $v_j$  at which the spacing between neighboring vibrational levels goes to zero. This value  $v_{j, \max}$  specifies the quantum number of the last bound vibrational level for the particular potential energy function  $V(x_j)$  of interest. The dissociation energy  $D_e$  can then be computed by adding to  $1/2\hbar \nu_j$  (the zero point energy along this mode) the sum of the spacings between neighboring vibrational energy levels from  $v_j = 0$  to  $v_j = v_{j, \max}$ :

$$D_e = 1/2\hbar \nu_j + \sum_{v_j = 0}^{v_{j, \max}} E_{v_j}.$$

Since experimental data are not usually available for the entire range of  $v_j$  values (from 0 to  $v_{j, \max}$ ), this sum must be computed using the anharmonic expression for  $E_{v_j}$ :

$$E_{v_j} = \hbar \left[ \nu_j - 2(x/\nu_j)v_j(v_j + 1/2) + \dots \right].$$

Alternatively, the sum can be computed from the Birge-Sponer graph by measuring the area under the straight-line fit to the graph of  $E_{v_j}$  or  $\nu_j$  from  $v_j = 0$  to  $v_j = v_{j,\max}$ .

*This completes our introduction to the subject of rotational and vibrational motions of molecules (which applies equally well to ions and radicals). The information contained in this Section is used again in Section 5 where photon-induced transitions between pairs of molecular electronic, vibrational, and rotational eigenstates are examined. More advanced treatments of the subject matter of this Section can be found in the text by Wilson, Decius, and Cross, as well as in Zare's text on angular momentum.*

## Section 5 Time Dependent Processes

### Chapter 14

*The interaction of a molecular species with electromagnetic fields can cause transitions to occur among the available molecular energy levels (electronic, vibrational, rotational, and nuclear spin). Collisions among molecular species likewise can cause transitions to occur. Time-dependent perturbation theory and the methods of molecular dynamics can be employed to treat such transitions.*

#### I. The Perturbation Describing Interactions With Electromagnetic Radiation

The full N-electron non-relativistic Hamiltonian H discussed earlier in this text involves the kinetic energies of the electrons and of the nuclei and the mutual coulombic interactions among these particles

$$H = \sum_{a=1, M} \left( -\frac{\hbar^2}{2m_a} \right) \nabla_a^2 + \sum_j \left[ \left( -\frac{\hbar^2}{2m_e} \right) \nabla_j^2 - \sum_a Z_a e^2 / r_{j,a} \right] \\ + \sum_{j < k} e^2 / r_{j,k} + \sum_{a < b} Z_a Z_b e^2 / R_{a,b}.$$

When an electromagnetic field is present, this is not the correct Hamiltonian, but it can be modified straightforwardly to obtain the proper H.

#### A. The Time-Dependent Vector $\mathbf{A}(\mathbf{r}, t)$ Potential

The only changes required to achieve the Hamiltonian that describes the same system in the presence of an electromagnetic field are to replace the momentum operators  $\mathbf{P}_a$  and  $\mathbf{p}_j$  for the nuclei and electrons, respectively, by  $(\mathbf{P}_a - Z_a e/c \mathbf{A}(R_{a,t}))$  and  $(\mathbf{p}_j - e/c \mathbf{A}(r_j, t))$ . Here  $Z_a e$  is the charge on the  $a^{\text{th}}$  nucleus,  $-e$  is the charge of the electron, and  $c$  is the speed of light.

The vector potential  $\mathbf{A}$  depends on time  $t$  and on the spatial location  $\mathbf{r}$  of the particle in the following manner:

$$\mathbf{A}(\mathbf{r}, t) = 2 \mathbf{A}_0 \cos (\omega t - \mathbf{k} \cdot \mathbf{r}).$$

The circular frequency of the radiation  $\omega$  (radians per second) and the wave vector  $\mathbf{k}$  (the magnitude of  $\mathbf{k}$  is  $|\mathbf{k}| = 2\pi / \lambda$ , where  $\lambda$  is the wavelength of the light) control the temporal

and spatial oscillations of the photons. The vector  $\mathbf{A}_0$  characterizes the strength (through the magnitude of  $\mathbf{A}_0$ ) of the field as well as the direction of the  $\mathbf{A}$  potential; the direction of propagation of the photons is given by the unit vector  $\mathbf{k}/|\mathbf{k}|$ . The factor of 2 in the definition of  $\mathbf{A}$  allows one to think of  $\mathbf{A}_0$  as measuring the strength of both  $\exp(i(\omega t - \mathbf{k} \cdot \mathbf{r}))$  and  $\exp(-i(\omega t - \mathbf{k} \cdot \mathbf{r}))$  components of the  $\cos(\omega t - \mathbf{k} \cdot \mathbf{r})$  function.

### B. The Electric $\mathbf{E}(\mathbf{r},t)$ and Magnetic $\mathbf{H}(\mathbf{r},t)$ Fields

The electric  $\mathbf{E}(\mathbf{r},t)$  and magnetic  $\mathbf{H}(\mathbf{r},t)$  fields of the photons are expressed in terms of the vector potential  $\mathbf{A}$  as

$$\mathbf{E}(\mathbf{r},t) = -\frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} = -\frac{1}{c} \dot{\mathbf{A}} = -\frac{1}{c} \dot{\mathbf{A}}_0 \sin(\omega t - \mathbf{k} \cdot \mathbf{r})$$

$$\mathbf{H}(\mathbf{r},t) = \nabla \times \mathbf{A} = \mathbf{k} \times \mathbf{A}_0 \cos(\omega t - \mathbf{k} \cdot \mathbf{r}).$$

The  $\mathbf{E}$  field lies parallel to the  $\mathbf{A}_0$  vector, and the  $\mathbf{H}$  field is perpendicular to  $\mathbf{A}_0$ ; both are perpendicular to the direction of propagation of the light  $\mathbf{k}/|\mathbf{k}|$ .  $\mathbf{E}$  and  $\mathbf{H}$  have the same phase because they both vary with time and spatial location as  $\sin(\omega t - \mathbf{k} \cdot \mathbf{r})$ . The relative orientations of these vectors are shown below.



### C. The Resulting Hamiltonian

Replacing the nuclear and electronic momenta by the modifications shown above in the kinetic energy terms of the full electronic and nuclear-motion hamiltonian results in the following additional factors appearing in H:

$$H_{\text{int}} = \sum_j \left\{ (ie \hbar / m_e c) \mathbf{A}(r_j, t) \cdot \mathbf{p}_j + (e^2 / 2m_e c^2) |\mathbf{A}(r_j, t)|^2 \right\} \\ + \sum_a \left\{ (i Z_a e \hbar / m_a c) \mathbf{A}(R_a, t) \cdot \mathbf{p}_a + (Z_a^2 e^2 / 2m_a c^2) |\mathbf{A}(R_a, t)|^2 \right\}.$$

These so-called interaction perturbations  $H_{\text{int}}$  are what induces transitions among the various electronic/vibrational/rotational states of a molecule. The one-electron additive nature of  $H_{\text{int}}$  plays an important role in determining the kind of transitions that  $H_{\text{int}}$  can induce. For example, it causes the most intense electronic transitions to involve excitation of a single electron from one orbital to another (recall the Slater-Condon rules).

## II. Time-Dependent Perturbation Theory

### A. The Time-Dependent Schrödinger Equation

The mathematical machinery needed to compute the rates of transitions among molecular states induced by such a time-dependent perturbation is contained in time-dependent perturbation theory (TDPT). The development of this theory proceeds as follows. One first assumes that one has in-hand all of the eigenfunctions  $\{ \psi_k \}$  and eigenvalues  $\{ E_k^0 \}$  that characterize the Hamiltonian  $H^0$  of the molecule in the absence of the external perturbation:

$$H^0 \psi_k = E_k^0 \psi_k.$$

One then writes the time-dependent Schrödinger equation

$$i \hbar \partial / \partial t \psi = (H^0 + H_{\text{int}}) \psi$$

in which the full Hamiltonian is explicitly divided into a part that governs the system in the absence of the radiation field and  $H_{\text{int}}$  which describes the interaction with the field.

### B. Perturbative Solution

By treating  $H^0$  as of zeroth order (in the field strength  $|\mathbf{A}_0|$ ), expanding order-by-order in the field-strength parameter:

$$= 0 + 1 + 2 + 3 + \dots,$$

realizing that  $H_{\text{int}}$  contains terms that are both first- and second- order in  $|\mathbf{A}_0|$

$$H^1_{\text{int}} = \sum_j \left\{ (ie \hbar / m_e c) \mathbf{A}(r_j, t) \cdot \mathbf{p}_j \right\}$$

$$+ \sum_a \left\{ (i Z_a e \hbar / m_a c) \mathbf{A}(R_a, t) \cdot \mathbf{p}_a \right\},$$

$$H^2_{\text{int}} = \sum_j \left\{ (e^2 / 2m_e c^2) |\mathbf{A}(r_j, t)|^2 \right\}$$

$$+ \sum_a \left\{ (Z_a^2 e^2 / 2m_a c^2) |\mathbf{A}(R_a, t)|^2 \right\},$$

and then collecting together all terms of like power of  $|\mathbf{A}_0|$ , one obtains the set of time-dependent perturbation theory equations. The lowest order such equations read:

$$i \hbar \frac{d}{dt} \psi^0 = H^0 \psi^0$$

$$i \hbar \frac{d}{dt} \psi^1 = (H^0 + H^1_{\text{int}}) \psi^0$$

$$i \hbar \frac{d}{dt} \psi^2 = (H^0 + H^2_{\text{int}}) \psi^0 + H^1_{\text{int}} \psi^1.$$

The zeroth order equations can easily be solved because  $H^0$  is independent of time. Assuming that at  $t = -\infty$ ,  $\psi = \psi_i$  (we use the index  $i$  to denote the initial state), this solution is:

$$\psi^0 = \psi_i \exp(-i E_i^0 t / \hbar).$$

The first-order correction to  $\psi^0$ ,  $\psi^1$  can be found by (i) expanding  $\psi^1$  in the complete set of zeroth-order states  $\{ \psi_f \}$ :

$$\psi^1 = \sum_f \langle \psi_f | \psi^1 \rangle \psi_f = \sum_f C_f^1 \psi_f,$$

(ii) using the fact that

$$H^0 \psi_f = E_f^0 \psi_f,$$

and (iii) substituting all of this into the equation that  $\psi^{(1)}$  obeys. The resultant equation for the coefficients that appear in the first-order equation can be written as

$$i \hbar \frac{dC_f^{(1)}}{dt} = \sum_k \{E_k^{(0)} C_k^{(1)}\}_{f,k} + \langle f | H_{int}^{(1)} | i \rangle \exp(-i E_i^{(0)} t / \hbar),$$

or

$$i \hbar \frac{dC_f^{(1)}}{dt} = E_f^{(0)} C_f^{(1)} + \langle f | H_{int}^{(1)} | i \rangle \exp(-i E_i^{(0)} t / \hbar).$$

Defining

$$C_f^{(1)}(t) = D_f^{(1)}(t) \exp(-i E_f^{(0)} t / \hbar),$$

this equation can be cast in terms of an easy-to-solve equation for the  $D_f^{(1)}$  coefficients:

$$i \hbar \frac{dD_f^{(1)}}{dt} = \langle f | H_{int}^{(1)} | i \rangle \exp(i [E_f^{(0)} - E_i^{(0)}] t / \hbar).$$

Assuming that the electromagnetic field  $\mathbf{A}(\mathbf{r},t)$  is turned on at  $t=0$ , and remains on until  $t = T$ , this equation for  $D_f^{(1)}$  can be integrated to yield:

$$D_f^{(1)}(t) = (i \hbar)^{-1} \int_0^T \langle f | H_{int}^{(1)} | i \rangle \exp(i [E_f^{(0)} - E_i^{(0)}] t' / \hbar) dt'.$$

## C. Application to Electromagnetic Perturbations

### 1. First-Order Fermi-Wentzel "Golden Rule"

Using the earlier expressions for  $H_{int}^{(1)}$  and for  $\mathbf{A}(\mathbf{r},t)$

$$H_{int}^{(1)} = \sum_j \left\{ \left( \frac{i e \hbar}{m_e c} \right) \mathbf{A}(\mathbf{r}_j, t) \cdot \mathbf{p}_j \right\}$$

$$+ \sum_a \left\{ \left( \frac{i Z_a e \hbar}{m_a c} \right) \mathbf{A}(\mathbf{R}_a, t) \cdot \mathbf{p}_a \right\}$$

and

$$2 \mathbf{A}_0 \cos(\omega t - \mathbf{k} \cdot \mathbf{r}) = \mathbf{A}_0 \{ \exp[i(\omega t - \mathbf{k} \cdot \mathbf{r})] + \exp[-i(\omega t - \mathbf{k} \cdot \mathbf{r})] \},$$

it is relatively straightforward to carry out the above time integration to achieve a final expression for  $D_f^1(t)$ , which can then be substituted into  $C_f^1(t) = D_f^1(t) \exp(-i E_f^0 t / \hbar)$  to obtain the final expression for the first-order estimate of the probability amplitude for the molecule appearing in the state  $f \exp(-i E_f^0 t / \hbar)$  after being subjected to electromagnetic radiation from  $t = 0$  until  $t = T$ . This final expression reads:

$$\begin{aligned} C_f^1(T) = & (i \hbar)^{-1} \exp(-i E_f^0 T / \hbar) \{ \langle f | j \{ (ie \hbar / m_e c) \exp[-i \mathbf{k} \cdot \mathbf{r}_j] \mathbf{A}_0 \cdot \mathbf{e}_j \\ & + \sum_a (i Z_a e \hbar / m_a c) \exp[-i \mathbf{k} \cdot \mathbf{R}_a] \mathbf{A}_0 \cdot \mathbf{e}_a | i \rangle \} \frac{\exp(i(\omega_{fi} + \omega) T) - 1}{i(\omega_{fi} + \omega)} \\ & + (i \hbar)^{-1} \exp(-i E_f^0 T / \hbar) \{ \langle f | j \{ (ie \hbar / m_e c) \exp[i \mathbf{k} \cdot \mathbf{r}_j] \mathbf{A}_0 \cdot \mathbf{e}_j \\ & + \sum_a (i Z_a e \hbar / m_a c) \exp[i \mathbf{k} \cdot \mathbf{R}_a] \mathbf{A}_0 \cdot \mathbf{e}_a | i \rangle \} \frac{\exp(i(-\omega_{fi} + \omega) T) - 1}{i(-\omega_{fi} + \omega)}, \end{aligned}$$

where

$$\omega_{fi} = [E_f^0 - E_i^0] / \hbar$$

is the resonance frequency for the transition between "initial" state  $i$  and "final" state  $f$ .

Defining the time-independent parts of the above expression as

$$\begin{aligned} f_{f,i} = & \langle f | j \{ (e \hbar / m_e c) \exp[-i \mathbf{k} \cdot \mathbf{r}_j] \mathbf{A}_0 \cdot \mathbf{e}_j \\ & + \sum_a (Z_a e \hbar / m_a c) \exp[-i \mathbf{k} \cdot \mathbf{R}_a] \mathbf{A}_0 \cdot \mathbf{e}_a | i \rangle, \end{aligned}$$

this result can be written as

$$\begin{aligned} C_f^1(T) = & \exp(-i E_f^0 T / \hbar) \{ f_{f,i} \frac{\exp(i(\omega_{fi} + \omega) T) - 1}{i(\omega_{fi} + \omega)} \\ & + f_{f,i} \frac{\exp(-i(\omega_{fi} - \omega) T) - 1}{-i(\omega_{fi} - \omega)} \}. \end{aligned}$$

The modulus squared  $|C_f^1(T)|^2$  gives the probability of finding the molecule in the final state  $f$  at time  $T$ , given that it was in  $i$  at time  $t = 0$ . If the light's frequency  $\omega$  is tuned close to the transition frequency  $\omega_{fi}$  of a particular transition, the term whose denominator contains  $(\omega - \omega_{fi})$  will dominate the term with  $(\omega + \omega_{fi})$  in its denominator. Within this "near-resonance" condition, the above probability reduces to:

$$\begin{aligned} |C_f^1(T)|^2 &= 2 |V_{fi}|^2 \frac{(1 - \cos((\omega - \omega_{fi})T))}{(\omega - \omega_{fi})^2} \\ &= 4 |V_{fi}|^2 \frac{\sin^2(1/2(\omega - \omega_{fi})T)}{(\omega - \omega_{fi})^2} . \end{aligned}$$

This is the final result of the first-order time-dependent perturbation theory treatment of light-induced transitions between states  $i$  and  $f$ .

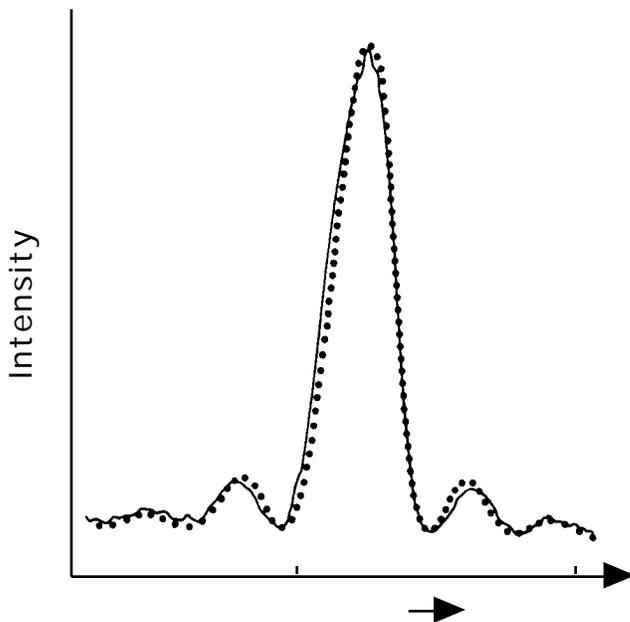
The so-called sinc- function

$$\frac{\sin^2(1/2(\omega - \omega_{fi})T)}{(\omega - \omega_{fi})^2}$$

as shown in the figure below is strongly peaked near  $\omega = \omega_{fi}$ , and displays secondary maxima (of decreasing amplitudes) near  $\omega = \omega_{fi} + 2n\pi/T$ ,  $n = 1, 2, \dots$ . In the  $T \rightarrow \infty$  limit, this function becomes narrower and narrower, and the area under it

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{\sin^2(1/2(\omega - \omega_{fi})T)}{(\omega - \omega_{fi})^2} d\omega &= T/2 \int_{-\infty}^{\infty} \frac{\sin^2(1/2(x - \omega_{fi})T)}{1/4T^2(x - \omega_{fi})^2} dx \\ &= T/2 \int_{-\infty}^{\infty} \frac{\sin^2(x)}{x^2} dx = T/2 \int_{-\infty}^{\infty} \frac{\sin^2(x)}{x^2} dx \end{aligned}$$

grows with T. Physically, this means that when the molecules are exposed to the light source for long times (large T), the sinc function emphasizes values near  $\omega_i$  (i.e., the on-resonance values). These properties of the sinc function will play important roles in what follows.



In most experiments, light sources have a "spread" of frequencies associated with them; that is, they provide photons of various frequencies. To characterize such sources, it is common to introduce the spectral source function  $g(\omega) d\omega$  which gives the probability that the photons from this source have frequency somewhere between  $\omega$  and  $\omega + d\omega$ . For narrow-band lasers,  $g(\omega)$  is a sharply peaked function about some "nominal" frequency  $\omega_0$ ; broader band light sources have much broader  $g(\omega)$  functions.

When such non-monochromatic light sources are used, it is necessary to average the above formula for  $|C_f^1(T)|^2$  over the  $g(\omega) d\omega$  probability function in computing the probability of finding the molecule in state  $f$  after time T, given that it was in  $i$  up until  $t = 0$ , when the light source was turned on. In particular, the proper expression becomes:

$$|C_f^1(T)|_{\text{ave}}^2 = 4 |C_{f,i}^1|^2 \int g(\omega) \frac{\sin^2(1/2(\omega - \omega_{f,i})T)}{(\omega - \omega_{f,i})^2} d\omega$$

$$= 2 |f_{i,i}|^2 T \int_{-\infty}^{\infty} g(\omega) \frac{\sin^2(1/2(\omega - \omega_{f,i})T)}{1/4T^2(\omega - \omega_{f,i})^2} d\omega \quad T/2 .$$

If the light-source function is "tuned" to peak near  $\omega = \omega_{f,i}$ , and if  $g(\omega)$  is much broader (in  $\omega$ -space) than the  $\frac{\sin^2(1/2(\omega - \omega_{f,i})T)}{(\omega - \omega_{f,i})^2}$  function,  $g(\omega)$  can be replaced by its value at the peak of the  $\frac{\sin^2(1/2(\omega - \omega_{f,i})T)}{(\omega - \omega_{f,i})^2}$  function, yielding:

$$|C_f^1(T)|_{\text{ave}}^2 = 2 g(\omega_{f,i}) |f_{f,i}|^2 T \int_{-\infty}^{\infty} \frac{\sin^2(1/2(\omega - \omega_{f,i})T)}{1/4T^2(\omega - \omega_{f,i})^2} d\omega \quad T/2$$

$$= 2 g(\omega_{f,i}) |f_{f,i}|^2 T \int_{-\infty}^{\infty} \frac{\sin^2(x)}{x^2} dx = 2 g(\omega_{f,i}) |f_{f,i}|^2 T .$$

The fact that the probability of excitation from  $\omega_i$  to  $\omega_f$  grows linearly with the time  $T$  over which the light source is turned on implies that the rate of transitions between these two states is constant and given by:

$$R_{i,f} = 2 g(\omega_{f,i}) |f_{f,i}|^2 ;$$

this is the so-called first-order Fermi-Wentzel "**golden rule**" expression for such transition rates. It gives the rate as the square of a transition matrix element between the two states involved, of the first order perturbation multiplied by the light source function  $g(\omega)$  evaluated at the transition frequency  $\omega_{f,i}$ .

## 2. Higher Order Results

Solution of the second-order time-dependent perturbation equations,

$$i \hbar \frac{d}{dt} c_n = (H^0 c_n + H_{\text{int}}^2 c_n + H_{\text{int}}^1 c_m) \quad (1)$$

which will not be treated in detail here, gives rise to two distinct types of contributions to the transition probabilities between  $i$  and  $f$ :

i. There will be matrix elements of the form

$$\langle f | j \{ (e^2/2m_e c^2) |\mathbf{A}(\mathbf{r}_j, t)|^2 \} + a \{ (Z_a^2 e^2/2m_a c^2) |\mathbf{A}(\mathbf{R}_a, t)|^2 \} | i \rangle$$

arising when  $H_{int}^2$  couples  $i$  to  $f$ .

ii. There will be matrix elements of the form

$$\langle k | f | j \{ (ie \hbar / m_e c) \mathbf{A}(\mathbf{r}_j, t) \cdot \mathbf{j} \} + a \{ (i Z_a e \hbar / m_a c) \mathbf{A}(\mathbf{R}_a, t) \cdot \mathbf{a} \} | k \rangle$$

$$\langle k | j \{ (ie \hbar / m_e c) \mathbf{A}(\mathbf{r}_j, t) \cdot \mathbf{j} \} + a \{ (i Z_a e \hbar / m_a c) \mathbf{A}(\mathbf{R}_a, t) \cdot \mathbf{a} \} | i \rangle$$

arising from expanding  $H_{int}^1 = \sum_k C_k^1 H_{int}^1 | k \rangle$  and using the earlier result for the first-order amplitudes  $C_k^1$ . Because both types of second-order terms vary quadratically with the  $\mathbf{A}(\mathbf{r}, t)$  potential, and because  $\mathbf{A}$  has time dependence of the form  $\cos(\omega t - \mathbf{k} \cdot \mathbf{r})$ , these terms contain portions that vary with time as  $\cos(2\omega t)$ . As a result, transitions between initial and final states  $i$  and  $f$  whose transition frequency is  $\omega_{fi}$  can be induced when  $2\omega = \omega_{fi}$ ; in this case, one speaks of coherent two-photon induced transitions in which the electromagnetic field produces a perturbation that has twice the frequency of the "nominal" light source frequency  $\omega$ .

#### D. The "Long-Wavelength" Approximation

To make progress in further analyzing the first-order results obtained above, it is useful to consider the wavelength  $\lambda$  of the light used in most visible/ultraviolet, infrared, or microwave spectroscopic experiments. Even the shortest such wavelengths (ultraviolet) are considerably longer than the spatial extent of all but the largest molecules (i.e., polymers and biomolecules for which the approximations we introduce next are not appropriate).

In the definition of the essential coupling matrix element  $f_{fi}$

$$f_{fi} = \langle f | j \{ (e / m_e c) \exp[-i\mathbf{k} \cdot \mathbf{r}_j] \mathbf{A}_0 \cdot \mathbf{j} \} + a \{ (Z_a e / m_a c) \exp[-i\mathbf{k} \cdot \mathbf{R}_a] \mathbf{A}_0 \cdot \mathbf{a} \} | i \rangle,$$

the factors  $\exp[-i\mathbf{k}\cdot\mathbf{r}_j]$  and  $\exp[-i\mathbf{k}\cdot\mathbf{R}_a]$  can be expanded as:

$$\exp[-i\mathbf{k}\cdot\mathbf{r}_j] = 1 + (-i\mathbf{k}\cdot\mathbf{r}_j) + 1/2 (-i\mathbf{k}\cdot\mathbf{r}_j)^2 + \dots$$

$$\exp[-i\mathbf{k}\cdot\mathbf{R}_a] = 1 + (-i\mathbf{k}\cdot\mathbf{R}_a) + 1/2 (-i\mathbf{k}\cdot\mathbf{R}_a)^2 + \dots$$

Because  $|\mathbf{k}| = 2\pi/\lambda$ , and the scales of  $\mathbf{r}_j$  and  $\mathbf{R}_a$  are of the dimension of the molecule,  $\mathbf{k}\cdot\mathbf{r}_j$  and  $\mathbf{k}\cdot\mathbf{R}_a$  are less than unity in magnitude, within this so-called "long-wavelength" approximation.

### 1. Electric Dipole Transitions

Introducing these expansions into the expression for  $f_{f,i}$  gives rise to terms of various powers in  $1/\lambda$ . The lowest order terms are:

$$f_{f,i}(E1) = \langle f | \sum_j (e/m_e c) \mathbf{A}_0 \cdot \mathbf{r}_j + \sum_a (Z_a e/m_a c) \mathbf{A}_0 \cdot \mathbf{R}_a | i \rangle$$

and are called "electric dipole" terms, and are denoted E1. To see why these matrix elements are termed E1, we use the following identity (see Chapter 1) between the momentum operator  $-i\hbar\nabla$  and the corresponding position operator  $\mathbf{r}$ :

$$\mathbf{r}_j = - (m_e/\hbar^2) [H, \mathbf{r}_j]$$

$$\mathbf{R}_a = - (m_a/\hbar^2) [H, \mathbf{R}_a].$$

This derives from the fact that  $H$  contains  $\sum_j p_j^2$  and  $\sum_a p_a^2$  in its kinetic energy operators (as  $\sum_j p_j^2$  and  $\sum_a p_a^2$ ).

Substituting these expressions into the above  $f_{f,i}(E1)$  equation and using  $H|i\rangle = E_i|i\rangle$  or  $H|f\rangle = E_f|f\rangle$ , one obtains:

$$\begin{aligned} f_{f,i}(E1) &= (E_f - E_i) \mathbf{A}_0 \cdot \langle f | \sum_j (e/\hbar^2 c) \mathbf{r}_j + \sum_a (Z_a e/\hbar^2 c) \mathbf{R}_a | i \rangle \\ &= \sum_j f_{f,i} \mathbf{A}_0 \cdot \langle f | \sum_j (e/\hbar c) \mathbf{r}_j + \sum_a (Z_a e/\hbar c) \mathbf{R}_a | i \rangle \\ &= \langle f | \boldsymbol{\mu} | i \rangle, \end{aligned}$$

where  $\mu$  is the electric dipole moment operator for the electrons and nuclei:

$$\mu = \sum_j e \mathbf{r}_j + \sum_a Z_a e \mathbf{R}_a.$$

The fact that the E1 approximation to  $\Gamma_{f,i}$  contains matrix elements of the electric dipole operator between the initial and final states makes it clear why this is called the electric dipole contribution to  $\Gamma_{f,i}$ ; within the E1 notation, the E stands for electric moment and the 1 stands for the first such moment (i.e., the dipole moment).

Within this approximation, the overall rate of transitions is given by:

$$\begin{aligned} R_{i,f} &= 2 \sum_{\mathbf{k}} g(\mathbf{k}) |\langle f | \mu | i \rangle|^2 \\ &= 2 \sum_{\mathbf{k}} g(\mathbf{k}) \left( \frac{1}{\hbar c} \right)^2 |\mathbf{A}_0 \cdot \langle f | \mu | i \rangle|^2. \end{aligned}$$

Recalling that  $\mathbf{E}(\mathbf{r},t) = -1/c \nabla \phi = -1/c \mathbf{A}_0 \sin(\omega t - \mathbf{k} \cdot \mathbf{r})$ , the magnitude of  $\mathbf{A}_0$  can be replaced by that of  $\mathbf{E}$ , and this rate expression becomes

$$R_{i,f} = (2/\hbar^2) \sum_{\mathbf{k}} g(\mathbf{k}) |\mathbf{E}_0 \cdot \langle f | \mu | i \rangle|^2.$$

This expresses the widely used E1 approximation to the Fermi-Wentzel golden rule.

## 2. Magnetic Dipole and Electric Quadrupole Transitions

When E1 predictions for the rates of transitions between states vanish (e.g., for symmetry reasons as discussed below), it is essential to examine higher order contributions to  $\Gamma_{f,i}$ . The next terms in the above long-wavelength expansion vary as  $1/\omega^2$  and have the form:

$$\begin{aligned} \Gamma_{f,i}(E2+M1) &= \sum_j \langle f | \sum_j (e/m_e c) [-i\mathbf{k} \cdot \mathbf{r}_j] \mathbf{A}_0 \cdot \mathbf{p}_j \\ &+ \sum_a (Z_a e/m_a c) [-i\mathbf{k} \cdot \mathbf{R}_a] \mathbf{A}_0 \cdot \mathbf{p}_a | i \rangle. \end{aligned}$$

For reasons soon to be shown, they are called electric quadrupole (E2) and magnetic dipole (M1) terms. Clearly, higher and higher order terms can be so generated. Within the long-wavelength regime, however, successive terms should decrease in magnitude because of the successively higher powers of  $1/\omega^2$  that they contain.

To further analyze the above E2 + M1 factors, let us label the propagation direction of the light as the z-axis (the axis along which  $\mathbf{k}$  lies) and the direction of  $\mathbf{A}_0$  as the x-axis. These axes are so-called "lab-fixed" axes because their orientation is determined by the direction of the light source and the direction of polarization of the light source's  $\mathbf{E}$  field, both of which are specified by laboratory conditions. The molecule being subjected to this light can be oriented at arbitrary angles relative to these lab axes.

With the x, y, and z axes so defined, the above expression for  $f_{f,i}(\text{E2+M1})$  becomes

$$f_{f,i}(\text{E2+M1}) = -i (A_0^2 / c) \langle f | j (e / m_e c) z_j / x_j + a (Z_a e / m_a c) z_a / x_a | i \rangle.$$

Now writing (for both  $z_j$  and  $z_a$ )

$$z / x = 1/2 (z / x - x / z + z / x + x / z),$$

and using

$$j = - (m_e / \hbar^2) [ H, \mathbf{r}_j ]$$

$$a = - (m_a / \hbar^2) [ H, \mathbf{R}_a ],$$

the contributions of  $1/2 (z / x + x / z)$  to  $f_{f,i}(\text{E2+M1})$  can be rewritten as

$$f_{f,i}(\text{E2}) = -i \frac{(A_0 e^2 / c \hbar) f_{f,i}}{c \hbar} \langle f | j z_j x_j + a Z_a z_a x_a | i \rangle.$$

The operator  $i z_j x_j + a Z_a z_a x_a$  that appears above is the z,x element of the electric quadrupole moment operator  $Q_{z,x}$ ; it is for this reason that this particular component is labeled E2 and denoted the electric quadrupole contribution.

The remaining  $1/2 (z / x - x / z)$  contribution to  $f_{f,i}(\text{E2+M1})$  can be rewritten in a form that makes its content more clear by first noting that

$$1/2 (z / x - x / z) = (i/2\hbar) (z p_x - x p_z) = (i/2\hbar) L_y$$

contains the y-component of the angular momentum operator. Hence, the following contribution to  $\rho_{f,i}(E_2+M_1)$  arises:

$$\rho_{f,i}(M_1) = \frac{A_0^2 e}{2 c \hbar} \langle f | \sum_j L_{y_j} / m_e + \sum_a Z_a L_{y_a} / m_a | i \rangle.$$

The magnetic dipole moment of the electrons about the y axis is

$$\mu_{y,\text{electrons}} = \sum_j (e/2m_e c) L_{y_j};$$

that of the nuclei is

$$\mu_{y,\text{nuclei}} = \sum_a (Z_a e / 2m_a c) L_{y_a}.$$

The  $\rho_{f,i}(M_1)$  term thus describes the interaction of the magnetic dipole moments of the electrons and nuclei with the magnetic field (of strength  $|H| = A_0 k$ ) of the light (which lies along the y axis):

$$\rho_{f,i}(M_1) = \frac{|H|}{\hbar} \langle f | \mu_{y,\text{electrons}} + \mu_{y,\text{nuclei}} | i \rangle.$$

The total rate of transitions from  $|i\rangle$  to  $|f\rangle$  is given, through first-order in perturbation theory, by

$$R_{i,f} = 2 \pi \rho_{f,i} | \rho_{f,i} |^2,$$

where  $\rho_{f,i}$  is a sum of its E1, E2, M1, etc. pieces. In the next chapter, molecular symmetry will be shown to be of use in analyzing these various pieces. It should be kept in mind that the contributions caused by E1 terms will dominate, within the long-wavelength approximation, unless symmetry causes these terms to vanish. It is primarily under such circumstances that consideration of M1 and E2 transitions is needed.

### III. The Kinetics of Photon Absorption and Emission

#### A. The Phenomenological Rate Laws

Before closing this chapter, it is important to emphasize the context in which the transition rate expressions obtained here are most commonly used. The perturbative approach used in the above development gives rise to various contributions to the overall rate coefficient for transitions from an initial state  $i$  to a final state  $f$ ; these contributions include the electric dipole, magnetic dipole, and electric quadrupole first order terms as well contributions arising from second (and higher) order terms in the perturbation solution.

In principle, once the rate expression

$$R_{i,f} = 2 \operatorname{Re} \langle f, i | \hat{V} | i, i \rangle^2$$

has been evaluated through some order in perturbation theory and including the dominant electromagnetic interactions, one can make use of these state-to-state rates, which are computed on a per-molecule basis, to describe the time evolution of the populations of the various energy levels of the molecule under the influence of the light source's electromagnetic fields.

For example, given two states, denoted  $i$  and  $f$ , between which transitions can be induced by photons of frequency  $\nu_{f,i}$ , the following kinetic model is often used to describe the time evolution of the numbers of molecules  $n_i$  and  $n_f$  in the respective states:

$$\frac{dn_i}{dt} = -R_{i,f} n_i + R_{f,i} n_f$$

$$\frac{dn_f}{dt} = -R_{f,i} n_f + R_{i,f} n_i$$

Here,  $R_{i,f}$  and  $R_{f,i}$  are the rates (per molecule) of transitions for the  $i \Rightarrow f$  and  $f \Rightarrow i$  transitions respectively. As noted above, these rates are proportional to the intensity of the light source (i.e., the photon intensity) at the resonant frequency and to the square of a matrix element connecting the respective states. This matrix element square is  $|\langle f, i | \hat{V} | i, i \rangle|^2$  in the former case and  $|\langle f, i | \hat{V} | i, i \rangle|^2$  in the latter. Because the perturbation operator whose matrix elements are  $\langle f, i | \hat{V} | i, i \rangle$  and  $\langle f, i | \hat{V} | i, i \rangle$  is Hermitian (this is true through all orders of perturbation theory and for all terms in the long-wavelength expansion), these two quantities are complex conjugates of one another, and, hence  $|\langle f, i | \hat{V} | i, i \rangle|^2 = |\langle f, i | \hat{V} | i, i \rangle|^2$ , from which it follows that  $R_{i,f} = R_{f,i}$ . This means that the state-to-state absorption and stimulated emission rate coefficients (i.e., the rate per molecule undergoing the transition) are identical. This result is referred to as the principle of **microscopic reversibility**.

Quite often, the states between which transitions occur are members of levels that contain more than a single state. For example, in rotational spectroscopy a transition between a state in the  $J = 3$  level of a diatomic molecule and a state in the  $J = 4$  level involve such states; the respective levels are  $2J+1 = 7$  and  $2J+1 = 9$  fold degenerate, respectively.

To extend the above kinetic model to this more general case in which degenerate levels occur, one uses the number of molecules in each **level** ( $N_i$  and  $N_f$  for the two levels in the above example) as the time dependent variables. The kinetic equations then governing their time evolution can be obtained by summing the state-to-state equations over all states in each level

$$i \text{ in level I } \left( \frac{dn_i}{dt} \right) = \frac{dN_I}{dt}$$

$$f \text{ in level F } \left( \frac{dn_f}{dt} \right) = \frac{dN_F}{dt}$$

and realizing that each state within a given level can undergo transitions to all states within the other level (hence the total rates of production and consumption must be summed over all states to or from which transitions can occur). This generalization results in a set of rate laws for the populations of the respective levels:

$$\frac{dN_i}{dt} = -g_f R_{i,f} N_i + g_i R_{f,i} N_f$$

$$\frac{dN_f}{dt} = -g_i R_{f,i} N_f + g_f R_{i,f} N_i .$$

Here,  $g_i$  and  $g_f$  are the degeneracies of the two levels (i.e., the number of states in each level) and the  $R_{i,f}$  and  $R_{f,i}$ , which are equal as described above, are the state-to-state rate coefficients introduced earlier.

## B. Spontaneous and Stimulated Emission

It turns out (the development of this concept is beyond the scope of this text) that the rate at which an excited level can emit photons and decay to a lower energy level is dependent on two factors: (i) the rate of **stimulated** photon emission as covered above, and (ii) the rate of **spontaneous** photon emission. The former rate  $g_f R_{i,f}$  (per molecule) is proportional to the light intensity  $g(f_i)$  at the resonance frequency. It is conventional to

separate out this intensity factor by defining an intensity independent rate coefficient  $B_{i,f}$  for this process as:

$$g_f R_{i,f} = g(f,i) B_{i,f}.$$

Clearly,  $B_{i,f}$  embodies the final-level degeneracy factor  $g_f$ , the perturbation matrix elements, and the  $2$  factor in the earlier expression for  $R_{i,f}$ . The spontaneous rate of transition from the excited to the lower level is found to be independent of photon intensity, because it deals with a process that does not require collision with a photon to occur, and is usually denoted  $A_{i,f}$ . The rate of photon-stimulated upward transitions from state  $f$  to state  $i$  ( $g_i R_{f,i} = g_i R_{i,f}$  in the present case) is also proportional to  $g(f,i)$ , so it is written by convention as:

$$g_i R_{f,i} = g(f,i) B_{f,i}.$$

An important relation between the  $B_{i,f}$  and  $B_{f,i}$  parameters exists and is based on the identity  $R_{i,f} = R_{f,i}$  that connects the state-to-state rate coefficients:

$$\frac{(B_{i,f})}{(B_{f,i})} = \frac{(g_f R_{i,f})}{(g_i R_{f,i})} = \frac{g_f}{g_i}.$$

This relationship will prove useful in the following sections.

### C. Saturated Transitions and Transparency

Returning to the kinetic equations that govern the time evolution of the populations of two levels connected by photon absorption and emission, and adding in the term needed for spontaneous emission, one finds (with the initial level being of the lower energy):

$$\frac{dN_i}{dt} = -gB_{i,f} N_i + (A_{f,i} + gB_{f,i})N_f$$

$$\frac{dN_f}{dt} = - (A_{f,i} + gB_{f,i})N_f + gB_{i,f} N_i$$

where  $g = g(\nu)$  denotes the light intensity at the resonance frequency.

At steady state, the populations of these two levels are given by setting  $\frac{dN_i}{dt} = \frac{dN_f}{dt} = 0$ :

$$\frac{N_f}{N_i} = \frac{(gB_{i,f})}{(A_{f,i} + gB_{f,i})} .$$

When the light source's intensity is so large as to render  $gB_{f,i} \gg A_{f,i}$  (i.e., when the rate of spontaneous emission is small compared to the stimulated rate), this population ratio reaches  $(B_{i,f}/B_{f,i})$ , which was shown earlier to equal  $(g_f/g_i)$ . In this case, one says that the populations have been **saturated** by the intense light source. Any further increase in light intensity will result in zero increase in the rate at which photons are being absorbed. Transitions that have had their populations saturated by the application of intense light sources are said to display optical **transparency** because they are unable to absorb (or emit) any further photons because of their state of saturation.

#### D. Equilibrium and Relations Between A and B Coefficients

When the molecules in the two levels being discussed reach equilibrium (at which time the  $\frac{dN_i}{dt} = \frac{dN_f}{dt} = 0$  also holds) with a photon source that itself is in equilibrium characterized by a temperature T, we must have:

$$\frac{N_f}{N_i} = \frac{g_f}{g_i} \exp(-(E_f - E_i)/kT) = \frac{g_f}{g_i} \exp(-h\nu/kT)$$

where  $g_f$  and  $g_i$  are the degeneracies of the states labeled f and i. The photon source that is characterized by an equilibrium temperature T is known as a **black body** radiator, whose intensity profile  $g(\nu)$  (in  $\text{erg cm}^{-3} \text{ sec}^{-1}$ ) is known to be of the form:

$$g(\nu) = \frac{2(h\nu)^3}{c^3 h^2} (\exp(h\nu/kT) - 1)^{-1} .$$

Equating the kinetic result that must hold at equilibrium:

$$\frac{N_f}{N_i} = \frac{(gB_{i,f})}{(A_{f,i} + gB_{f,i})}$$

to the thermodynamic result:

$$\frac{N_f}{N_i} = \frac{g_f}{g_i} \exp(-h\nu/kT),$$

and using the above black body  $g(\nu)$  expression and the identity

$$\frac{(B_{i,f})}{(B_{f,i})} = \frac{g_f}{g_i},$$

one can solve for the  $A_{f,i}$  rate coefficient in terms of the  $B_{f,i}$  coefficient. Doing so yields:

$$A_{f,i} = B_{f,i} \frac{2(h\nu)^3}{c^3 h^2}.$$

## E. Summary

In summary, the so-called **Einstein A and B rate coefficients** connecting a lower-energy initial state  $i$  and a final state  $f$  are related by the following conditions:

$$B_{i,f} = \frac{g_f}{g_i} B_{f,i}$$

and

$$A_{f,i} = \frac{2(h\nu)^3}{c^3 h^2} B_{f,i}.$$

These phenomenological level-to-level rate coefficients are related to the state-to-state  $R_{i,f}$  coefficients derived by applying perturbation theory to the electromagnetic perturbation through

$$g_f R_{i,f} = g(\nu_{f,i}) B_{i,f}.$$

The A and B coefficients can be used in a kinetic equation model to follow the time evolution of the populations of the corresponding levels:

$$\frac{dN_i}{dt} = -g_{B_{i,f}} N_i + (A_{f,i} + g_{B_{f,i}}) N_f$$

$$\frac{dN_f}{dt} = - (A_{f,i} + g_{B_{f,i}}) N_f + g_{B_{i,f}} N_i .$$

These equations possess steady state solutions

$$\frac{N_f}{N_i} = \frac{(g_{B_{i,f}})}{(A_{f,i} + g_{B_{f,i}})}$$

which, for large  $g$ ( ), produce saturation conditions:

$$\frac{N_f}{N_i} = \frac{(B_{i,f})}{(B_{f,i})} = \frac{g_f}{g_i} .$$

## Chapter 15

*The tools of time-dependent perturbation theory can be applied to transitions among electronic, vibrational, and rotational states of molecules.*

### I. Rotational Transitions

Within the approximation that the electronic, vibrational, and rotational states of a molecule can be treated as independent, the total molecular wavefunction of the "initial" state is a product

$$\psi_i = \psi_{ei} \psi_{vi} \psi_{ri}$$

of an electronic function  $\psi_{ei}$ , a vibrational function  $\psi_{vi}$ , and a rotational function  $\psi_{ri}$ . A similar product expression holds for the "final" wavefunction  $\psi_f$ .

In microwave spectroscopy, the energy of the radiation lies in the range of fractions of a  $\text{cm}^{-1}$  through several  $\text{cm}^{-1}$ ; such energies are adequate to excite rotational motions of molecules but are not high enough to excite any but the weakest vibrations (e.g., those of weakly bound Van der Waals complexes). In rotational transitions, the electronic and vibrational states are thus left unchanged by the excitation process; hence  $\psi_{ei} = \psi_{ef}$  and  $\psi_{vi} = \psi_{vf}$ .

Applying the first-order electric dipole transition rate expressions

$$R_{i,f} = 2 \pi g(\nu_{f,i}) |\langle \psi_f | \mu | \psi_i \rangle|^2$$

obtained in Chapter 14 to this case requires that the E1 approximation

$$R_{i,f} = (2 \pi / \hbar^2) g(\nu_{f,i}) |\mathbf{E}_0 \cdot \langle \psi_f | \boldsymbol{\mu} | \psi_i \rangle|^2$$

be examined in further detail. Specifically, the electric dipole matrix elements  $\langle \psi_f | \boldsymbol{\mu} | \psi_i \rangle$  with  $\boldsymbol{\mu} = \sum_j e \mathbf{r}_j + \sum_a Z_a e \mathbf{R}_a$  must be analyzed for  $\psi_i$  and  $\psi_f$  being of the product form shown above.

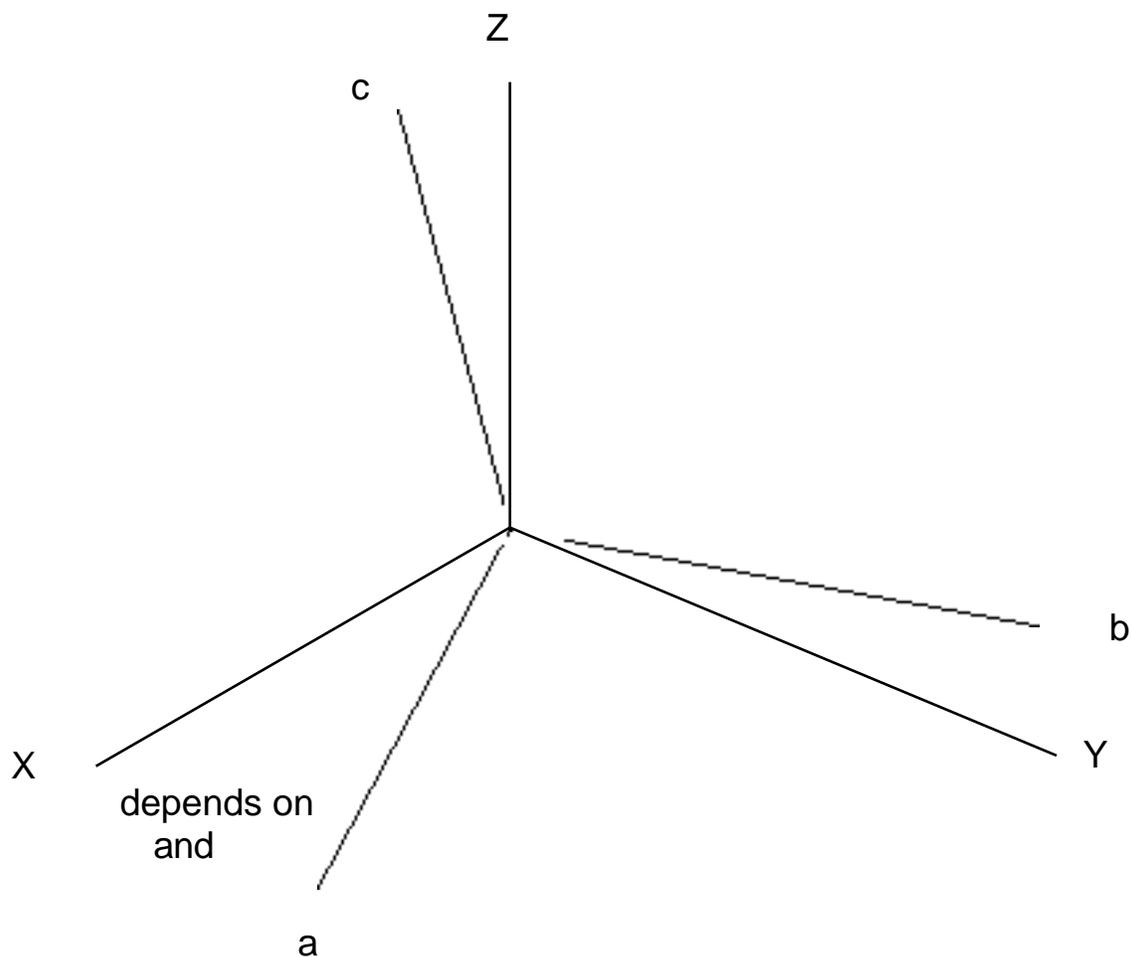
The integrations over the electronic coordinates contained in  $\langle \psi_f | \boldsymbol{\mu} | \psi_i \rangle$ , as well as the integrations over vibrational degrees of freedom yield "expectation values" of the electric dipole moment operator because the electronic and vibrational components of  $\psi_i$  and  $\psi_f$  are identical:

$$\langle e_i | \boldsymbol{\mu} | e_i \rangle = \boldsymbol{\mu}(\mathbf{R})$$

is the dipole moment of the initial electronic state (which is a function of the internal geometrical degrees of freedom of the molecule, denoted  $\mathbf{R}$ ); and

$$\langle v_i | \boldsymbol{\mu}(\mathbf{R}) | v_i \rangle = \boldsymbol{\mu}_{\text{ave}}$$

is the vibrationally averaged dipole moment for the particular vibrational state labeled  $v_i$ . The vector  $\boldsymbol{\mu}_{\text{ave}}$  has components along various directions and can be viewed as a vector "locked" to the molecule's internal coordinate axis (labeled a, b, c as below).



The rotational part of the  $\langle \psi_f | \mu | \psi_i \rangle$  integral is not of the expectation value form because the initial rotational function  $\psi_{ir}$  is not the same as the final  $\psi_{fr}$ . This integral has the form:

$$\langle \psi_{ir} | \mu_{ave} | \psi_{fr} \rangle = \int (Y_{L,M}^*(\theta, \phi) \mu_{ave} Y_{L',M'}(\theta, \phi) \sin \theta d\theta d\phi)$$

for linear molecules whose initial and final rotational wavefunctions are  $Y_{L,M}$  and  $Y_{L',M'}$ , respectively, and

$$\langle \psi_{ir} | \mu_{ave} | \psi_{fr} \rangle = \sqrt{\frac{2L+1}{8\pi}} \sqrt{\frac{2L'+1}{8\pi}}$$

$$\int (D_{L,M,K}(\theta, \phi, \chi) \mu_{ave} D_{L',M',K'}^*(\theta, \phi, \chi) \sin \theta d\theta d\phi d\chi)$$

for spherical or symmetric top molecules (here,  $\sqrt{\frac{2L+1}{8\pi}} D_{L,M,K}(\theta, \phi, \chi)$  are the normalized rotational wavefunctions described in Chapter 13 and in Appendix G). The angles  $\theta$ ,  $\phi$ , and  $\chi$  refer to how the molecule-fixed coordinate system is oriented with respect to the space-fixed X, Y, Z axis system.

### A. Linear Molecules

For linear molecules, the vibrationally averaged dipole moment  $\mu_{ave}$  lies along the molecular axis; hence its orientation in the lab-fixed coordinate system can be specified in terms of the same angles ( $\theta$  and  $\phi$ ) that are used to describe the rotational functions  $Y_{L,M}(\theta, \phi)$ . Therefore, the three components of the  $\langle \psi_{ir} | \mu_{ave} | \psi_{fr} \rangle$  integral can be written as:

$$\langle \psi_{ir} | \mu_{ave} | \psi_{fr} \rangle_x = \mu \int (Y_{L,M}^*(\theta, \phi) \sin \theta \cos \phi Y_{L',M'}(\theta, \phi) \sin \theta d\theta d\phi)$$

$$\langle \psi_{ir} | \mu_{ave} | \psi_{fr} \rangle_y = \mu \int (Y_{L,M}^*(\theta, \phi) \sin \theta \sin \phi Y_{L',M'}(\theta, \phi) \sin \theta d\theta d\phi)$$

$$\langle ir | \mu_{ave} | fr \rangle_z = \mu (Y_{L,M}^*(\theta, \phi) \cos \theta - Y_{L',M'}(\theta, \phi) \sin \theta) d\theta d\phi,$$

where  $\mu$  is the magnitude of the averaged dipole moment. **If the molecule has no dipole moment**, all of the above electric dipole integrals vanish and the **intensity** of E1 rotational transitions **is zero**.

The three E1 integrals can be further analyzed by noting that  $\cos \theta = Y_{1,0}$ ;  $\sin \theta \cos \theta = Y_{1,1} + Y_{1,-1}$ ; and  $\sin \theta \sin \theta = Y_{1,1} - Y_{1,-1}$  and using the angular momentum coupling methods illustrated in Appendix G. In particular, the result given in that appendix:

$$\begin{aligned} & D_{j, m, m'} D_{l, n, n'} \\ &= \sum_{J, M, M'} \langle J, M | j, m; l, n \rangle \langle j, m'; l, n' | J, M' \rangle D_{J, M, M'} \end{aligned}$$

when multiplied by  $D_{J, M, M'}^*$  and integrated over  $\sin \theta d\theta d\phi$ , yields:

$$\begin{aligned} & (D_{J, M, M'}^* D_{j, m, m'} D_{l, n, n'} \sin \theta d\theta d\phi) \\ &= \frac{8}{2J+1} \langle J, M | j, m; l, n \rangle \langle j, m'; l, n' | J, M' \rangle \\ &= 8 \frac{j-l-J}{m-n-M} \frac{j-l-J}{m'-n'-M'} (-1)^{M+M'}. \end{aligned}$$

To use this result in the present linear-molecule case, we note that the  $D_{J, M, K}$  functions and the  $Y_{J, M}$  functions are related by:

$$Y_{J, M}(\theta, \phi) = \sqrt{(2J+1)/4} D_{J, M, 0}(\theta, \phi).$$

The normalization factor is now  $\sqrt{(2J+1)/4}$  rather than  $\sqrt{(2J+1)/8}$  because the  $Y_{J, M}$  are no longer functions of  $\theta$ , and thus the need to integrate over  $\theta$  disappears.

Likewise, the  $\theta$ -dependence of  $D_{J, M, K}$  disappears for  $K = 0$ .

We now use these identities in the three E1 integrals of the form

$$\mu (Y_{L,M}^*(\theta, \phi) Y_{1,m}(\theta, \phi) Y_{L',M'}(\theta, \phi) \sin \theta d\theta d\phi),$$

with  $m = 0$  being the Z- axis integral, and the Y- and X- axis integrals being combinations of the  $m = 1$  and  $m = -1$  results. Doing so yields:

$$\begin{aligned} & \mu \int (Y_{L,M}^*(\theta, \phi) Y_{1,m}(\theta, \phi) Y_{L',M'}(\theta, \phi) \sin \theta \, d\theta \, d\phi) \\ &= \mu \sqrt{\frac{2L+1}{4} \frac{2L'+1}{4} \frac{3}{4}} (D_{L,M,0} D_{1,m,0}^* D_{L',M',0}^* \sin \theta \, d\theta \, d\phi / 2). \end{aligned}$$

The last factor of  $1/2$  is inserted to cancel out the integration over  $\theta$  that, because all K-factors in the rotation matrices equal zero, trivially yields  $2$ . Now, using the result shown above expressing the integral over three rotation matrices, these E1 integrals for the linear-molecule case reduce to:

$$\begin{aligned} & \mu \int (Y_{L,M}^*(\theta, \phi) Y_{1,m}(\theta, \phi) Y_{L',M'}(\theta, \phi) \sin \theta \, d\theta \, d\phi) \\ &= \mu \sqrt{\frac{2L+1}{4} \frac{2L'+1}{4} \frac{3}{4} \frac{8}{2}} \frac{L' - 1}{M' - m - M} \frac{L}{0 \ 0 \ -0} (-1)^M \\ &= \mu \sqrt{(2L+1)(2L'+1) \frac{3}{4}} \frac{L' - 1}{M' - m - M} \frac{L}{0 \ 0 \ -0} (-1)^M. \end{aligned}$$

Applied to the z-axis integral (identifying  $m = 0$ ), this result therefore vanishes unless:

$$M = M'$$

and

$$L = L' + 1 \text{ or } L' - 1.$$

Even though angular momentum coupling considerations would allow  $L = L'$  (because coupling two angular momenta with  $j = 1$  and  $j = L'$  should give  $L'+1$ ,  $L'$ , and  $L'-1$ ), the 3-j symbol  $\frac{L' - 1}{0 \ 0 \ -0} \frac{L}{0 \ 0 \ -0}$  vanishes for the  $L = L'$  case since 3-j symbols have the following symmetry

$$\frac{L' \quad 1 \quad L}{M' \quad m \quad -M} = (-1)^{L+L'+1} \frac{L' \quad 1 \quad L}{-M' \quad -m \quad M}$$

with respect to the M, M', and m indices. Applied to the  $\frac{L' \quad 1 \quad L}{0 \quad 0 \quad -0}$  3-j symbol, this means that this particular 3-j element vanishes for L = L' since L + L' + 1 is odd and hence (-1)<sup>L + L' + 1</sup> is -1.

Applied to the x- and y- axis integrals, which contain m = ± 1 components, this same analysis yields:

$$\mu \sqrt{(2L+1)(2L'+1)} \frac{3}{4} \frac{L' \quad 1 \quad L}{M' \quad \pm 1 \quad -M} \frac{L' \quad 1 \quad L}{0 \quad 0 \quad -0} (-1)^M$$

which then requires that

$$M = M' \pm 1$$

and

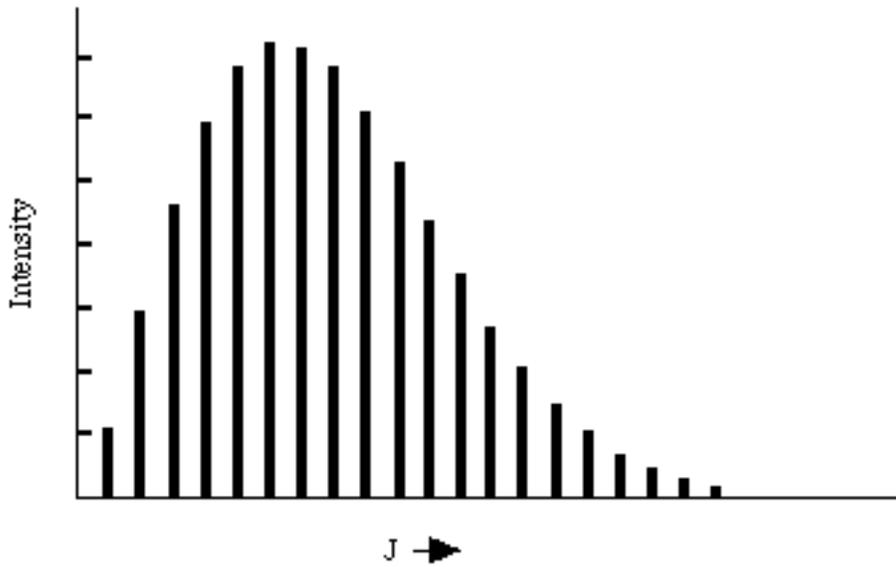
$$L = L' + 1, L' - 1,$$

with L = L' again being forbidden because of the second 3-j symbol.

These results provide so-called "**selection rules**" because they limit the L and M values of the final rotational state, given the L', M' values of the initial rotational state. In the figure shown below, the L = L' + 1 absorption spectrum of NO at 120 °K is given. The intensities of the various peaks are related to the populations of the lower-energy rotational states which are, in turn, proportional to (2L' + 1) exp(- L'(L'+1) ħ<sup>2</sup>/8<sup>2</sup>IkT). Also included in the intensities are so-called **line strength factors** that are proportional to the squares of the quantities:

$$\mu \sqrt{(2L+1)(2L'+1)} \frac{3}{4} \frac{L' \quad 1 \quad L}{M' \quad m \quad -M} \frac{L' \quad 1 \quad L}{0 \quad 0 \quad -0} (-1)^M$$

which appear in the E1 integrals analyzed above (recall that the rate of photon absorption  $R_{i,f} = (2/\hbar^2) g_{f,i} |\mathbf{E}_0 \cdot \langle f | \mu | i \rangle|^2$  involves the squares of these matrix elements). The book by Zare gives an excellent treatment of line strength factors' contributions to rotation, vibration, and electronic line intensities.



## B. Non-Linear Molecules

For molecules that are non-linear and whose rotational wavefunctions are given in terms of the spherical or symmetric top functions  $D^*_{L,M,K}$ , the dipole moment  $\mu_{ave}$  can have components along any or all three of the molecule's internal coordinates (e.g., the three molecule-fixed coordinates that describe the orientation of the principal axes of the moment of inertia tensor). For a spherical top molecule,  $|\mu_{ave}|$  vanishes, so E1 transitions do not occur.

For symmetric top species,  $\mu_{ave}$  lies along the symmetry axis of the molecule, so the orientation of  $\mu_{ave}$  can again be described in terms of  $\theta$  and  $\phi$ , the angles used to locate the orientation of the molecule's symmetry axis relative to the lab-fixed coordinate system. As a result, the E1 integral again can be decomposed into three pieces:

$$\langle ir | \mu_{ave} | fr \rangle_x = \mu \int (D_{L,M,K}(\theta, \phi) \cos \theta \cos \phi D^*_{L',M',K'}(\theta, \phi) \sin \theta d\theta d\phi d\phi)$$

$$\langle ir | \mu_{ave} | fr \rangle_y = \mu \int (D_{L,M,K}(\theta, \phi) \cos \theta \sin \phi D^*_{L',M',K'}(\theta, \phi) \sin \theta d\theta d\phi d\phi)$$

$$\langle ir | \mu_{ave} | fr \rangle_z = \mu \int (D_{L,M,K}(\theta, \phi) \sin \theta D^*_{L',M',K'}(\theta, \phi) \sin \theta d\theta d\phi d\phi).$$

Using the fact that  $\cos \theta = D^*_{1,0,0}$ ;  $\sin \theta \cos \theta = D^*_{1,1,0} + D^*_{1,-1,0}$ ; and  $\sin \theta \sin \theta = D^*_{1,1,0} - D^*_{1,-1,0}$ , and the tools of angular momentum coupling allows these integrals to be expressed, as above, in terms of products of the following 3-j symbols:

$$\frac{L' - 1 \quad L}{M' - m \quad -M} \quad \frac{L' - 1 \quad L}{K' \quad 0 \quad -K} ,$$

from which the following selection rules are derived:

$$L = L' + 1, L', L' - 1 \quad (\text{but not } L = L' = 0),$$

$$K = K',$$

$$M = M' + m,$$

with  $m = 0$  for the Z-axis integral and  $m = \pm 1$  for the X- and Y- axis integrals. In addition, if  $K = K' = 0$ , the  $L = L'$  transitions are also forbidden by the second 3-j symbol vanishing.

## II. Vibration-Rotation Transitions

When the initial and final electronic states are identical but the respective vibrational and rotational states are not, one is dealing with transitions between vibration-rotation states of the molecule. These transitions are studied in infrared (IR) spectroscopy using light of energy in the  $30 \text{ cm}^{-1}$  (far IR) to  $5000 \text{ cm}^{-1}$  range. The electric dipole matrix element analysis still begins with the electronic dipole moment integral  $\langle e_i | \mu | e_i \rangle = \mu(\mathbf{R})$ , but the integration over internal vibrational coordinates no longer produces the vibrationally averaged dipole moment. Instead one forms the vibrational **transition dipole** integral:

$$\langle v_f | \mu(\mathbf{R}) | v_i \rangle = \mu_{f,i}$$

between the initial  $v_i$  and final  $v_f$  vibrational states.

### A. The Dipole Moment Derivatives

Expressing  $\mu(\mathbf{R})$  in a power series expansion about the equilibrium bond length position (denoted  $\mathbf{R}_e$  collectively and  $R_{a,e}$  individually):

$$\mu(\mathbf{R}) = \mu(\mathbf{R}_e) + \sum_a \left( \frac{\partial \mu}{\partial R_a} \right) (R_a - R_{a,e}) + \dots,$$

substituting into the  $\langle \psi_f | \mu(\mathbf{R}) | \psi_i \rangle$  integral, and using the fact that  $\psi_i$  and  $\psi_f$  are orthogonal (because they are eigenfunctions of vibrational motion on the same electronic surface and hence of the same vibrational Hamiltonian), one obtains:

$$\begin{aligned} \langle \psi_f | \mu(\mathbf{R}) | \psi_i \rangle &= \mu(\mathbf{R}_e) \langle \psi_f | \psi_i \rangle + \sum_a \left( \frac{\partial \mu}{\partial R_a} \right) \langle \psi_f | (R_a - R_{a,e}) | \psi_i \rangle + \dots \\ &= \sum_a \left( \frac{\partial \mu}{\partial R_a} \right) \langle \psi_f | (R_a - R_{a,e}) | \psi_i \rangle + \dots \end{aligned}$$

This result can be interpreted as follows:

- i. Each independent vibrational mode of the molecule contributes to the  $\mu_{f,i}$  vector an amount equal to  $\left( \frac{\partial \mu}{\partial R_a} \right) \langle \psi_f | (R_a - R_{a,e}) | \psi_i \rangle + \dots$ .
- ii. Each such contribution contains one part  $\left( \frac{\partial \mu}{\partial R_a} \right)$  that depends on how the molecule's dipole moment function varies with vibration along that particular mode (labeled a),
- iii. and a second part  $\langle \psi_f | (R_a - R_{a,e}) | \psi_i \rangle$  that depends on the character of the initial and final vibrational wavefunctions.

If the vibration does not produce a **modulation of the dipole moment** (e.g., as with the symmetric stretch vibration of the CO<sub>2</sub> molecule), its infrared intensity vanishes because  $\left( \frac{\partial \mu}{\partial R_a} \right) = 0$ . One says that such transitions are infrared "inactive".

## B. Selection Rules on the Vibrational Quantum Number in the Harmonic Approximation

If the vibrational functions are described within the harmonic oscillator approximation, it can be shown that the  $\langle \psi_f | (R_a - R_{a,e}) | \psi_i \rangle$  integrals vanish unless  $v_f = v_i + 1$ ,  $v_f = v_i - 1$  (and that these integrals are proportional to  $(v_i + 1)^{1/2}$  and  $(v_i)^{1/2}$  in the respective cases). Even when  $\psi_f$  and  $\psi_i$  are rather non-harmonic, it turns out that such  $v_f = v_i \pm 1$  transitions have the largest  $\langle \psi_f | (R_a - R_{a,e}) | \psi_i \rangle$  integrals and therefore the highest infrared intensities. For these reasons, transitions that correspond to  $v_f = v_i \pm 1$  are called "**fundamental**"; those with  $v_f = v_i \pm 2$  are called "first **overtone**" transitions.

In summary then, vibrations for which the molecule's dipole moment is modulated as the vibration occurs (i.e., for which  $(\partial \mu / \partial R_a)$  is non-zero) and for which  $\nu = \pm 1$  tend to have large infrared intensities; overtones of such vibrations tend to have smaller intensities, and those for which  $(\partial \mu / \partial R_a) = 0$  have no intensity.

### C. Rotational Selection Rules for Vibrational Transitions

The result of all of the vibrational modes' contributions to  $\langle \psi_f | \partial \mu / \partial R_a | \psi_i \rangle$  is a vector  $\mu_{\text{trans}}$  that is termed the vibrational "transition dipole" moment. This is a vector with components along, in principle, all three of the internal axes of the molecule. For each particular vibrational transition (i.e., each particular  $\psi_i$  and  $\psi_f$ ) its orientation in space depends only on the orientation of the molecule; it is thus said to be locked to the molecule's coordinate frame. As such, its orientation relative to the lab-fixed coordinates (which is needed to effect a derivation of rotational selection rules as was done earlier in this Chapter) can be described much as was done above for the vibrationally averaged dipole moment that arises in purely rotational transitions. There are, however, important differences in detail. In particular,

- i. For a linear molecule  $\mu_{\text{trans}}$  can have components either along (e.g., when stretching vibrations are excited; these cases are denoted  $\sigma$ -cases) or perpendicular to (e.g., when bending vibrations are excited; they are denoted  $\pi$  cases) the molecule's axis.
- ii. For symmetric top species,  $\mu_{\text{trans}}$  need not lie along the molecule's symmetry axis; it can have components either along or perpendicular to this axis.
- iii. For spherical tops,  $\mu_{\text{trans}}$  will vanish whenever the vibration does not induce a dipole moment in the molecule. Vibrations such as the totally symmetric  $a_1$  C-H stretching motion in  $\text{CH}_4$  do not induce a dipole moment, and are thus infrared inactive; non-totally-symmetric vibrations can also be inactive if they induce no dipole moment.

As a result of the above considerations, the angular integrals

$$\langle \psi_{i,r} | \mu_{\text{trans}} | \psi_{f,r} \rangle = \int Y_{L,M}^*(\theta, \phi) \mu_{\text{trans}} Y_{L',M'}(\theta, \phi) \sin \theta \, d\theta \, d\phi$$

and

$$\langle ir | \mu_{\text{trans}} | fr \rangle = (D_{L,M,K}(\alpha, \beta, \gamma)) \mu_{\text{trans}} D_{L',M',K'}^*(\alpha, \beta, \gamma) \sin \alpha \, d \, d \, d$$

that determine the rotational selection rules appropriate to vibrational transitions produce similar, but not identical, results as in the purely rotational transition case.

The derivation of these selection rules proceeds as before, with the following additional considerations. The transition dipole moment's  $\mu_{\text{trans}}$  components along the lab-fixed axes must be related to its molecule-fixed coordinates (that are determined by the nature of the vibrational transition as discussed above). This transformation, as given in Zare's text, reads as follows:

$$(\mu_{\text{trans}})_m = \sum_k D_{1,m,k}^*(\alpha, \beta, \gamma) (\mu_{\text{trans}})_k$$

where  $(\mu_{\text{trans}})_m$  with  $m = 1, 0, -1$  refer to the components along the lab-fixed (X, Y, Z) axes and  $(\mu_{\text{trans}})_k$  with  $k = 1, 0, -1$  refer to the components along the molecule-fixed (a, b, c) axes.

This relationship, when used, for example, in the symmetric or spherical top E1 integral:

$$\langle ir | \mu_{\text{trans}} | fr \rangle = (D_{L,M,K}(\alpha, \beta, \gamma)) \mu_{\text{trans}} D_{L',M',K'}^*(\alpha, \beta, \gamma) \sin \alpha \, d \, d \, d$$

gives rise to products of 3-j symbols of the form:

$$\frac{L' - 1 \quad L}{M' \quad m \quad -M} \quad \frac{L' - 1 \quad L}{K' \quad k \quad -K} \quad .$$

The product of these 3-j symbols is nonvanishing only under certain conditions that provide the rotational selection rules applicable to vibrational lines of symmetric and spherical top molecules.

Both 3-j symbols will vanish unless

$$L = L' + 1, L', \text{ or } L' - 1.$$

In the special case in which  $L = L' = 0$  (and hence with  $M = M' = 0 = K = K'$ , which means that  $m = 0 = k$ ), these 3-j symbols again vanish. Therefore, transitions with

$$L = L' = 0$$

are again **forbidden**. As usual, the fact that the lab-fixed quantum number  $m$  can range over  $m = 1, 0, -1$ , requires that

$$M = M' + 1, M', M'-1.$$

The selection rules for  $K$  depend on the nature of the vibrational transition, in particular, on the component of  $\mu_{\text{trans}}$  along the molecule-fixed axes. For the second 3-j symbol to not vanish, one must have

$$K = K' + k,$$

where  $k = 0, 1$ , and  $-1$  refer to these molecule-fixed components of the transition dipole. Depending on the nature of the transition, various  $k$  values contribute.

### 1. Symmetric Tops

In a symmetric top molecule such as  $\text{NH}_3$ , if the transition dipole lies along the molecule's symmetry axis, only  $k = 0$  contributes. Such vibrations preserve the molecule's symmetry relative to this symmetry axis (e.g. the totally symmetric N-H stretching mode in  $\text{NH}_3$ ). The additional selection rule  $K = 0$  is thus obtained. Moreover, for  $K = K' = 0$ , all transitions with  $L = 0$  vanish because the second 3-j symbol vanishes. In summary, one has:

$$K = 0; \quad M = \pm 1, 0; \quad L = \pm 1, 0 \text{ (but } L = L' = 0 \text{ is forbidden and all } L = 0 \text{ are forbidden for } K = K' = 0)$$

for symmetric tops with vibrations whose transition dipole lies along the symmetry axis.

If the transition dipole lies perpendicular to the symmetry axis, only  $k = \pm 1$  contribute. In this case, one finds

$K = \pm 1; \quad M = \pm 1, 0; \quad L = \pm 1, 0$  (neither  $L = L' = 0$  nor  $K = K' = 0$  can occur for such transitions, so there are no additional constraints).

## 2. Linear Molecules

When the above analysis is applied to a diatomic species such as HCl, only  $k = 0$  is present since the only vibration present in such a molecule is the bond stretching vibration, which has  $\sigma$  symmetry. Moreover, the rotational functions are spherical harmonics (which can be viewed as  $D^*_{L',M',K'}$  (  $l, m, k$  ) functions with  $K' = 0$ ), so the  $K$  and  $K'$  quantum numbers are identically zero. As a result, the product of 3-j symbols

$$\frac{L' \quad 1 \quad L}{M' \quad m \quad -M} \quad \frac{L' \quad 1 \quad L}{K' \quad k \quad -K}$$

reduces to

$$\frac{L' \quad 1 \quad L}{M' \quad m \quad -M} \quad \frac{L' \quad 1 \quad L}{0 \quad 0 \quad 0} \quad ,$$

which will vanish unless

$$L = L' + 1, L' - 1,$$

but not  $L = L'$  (since parity then causes the second 3-j symbol to vanish), and

$$M = M' + 1, M', M' - 1.$$

The  $L = L' + 1$  transitions are termed **R-branch** absorptions and those obeying  $L = L' - 1$  are called **P-branch** transitions. Hence, the selection rules

$$M = \pm 1, 0; \quad L = \pm 1$$

are identical to those for purely rotational transitions.

When applied to linear polyatomic molecules, these same selection rules result if the vibration is of  $\sigma$  symmetry (i.e., has  $k = 0$ ). If, on the other hand, the transition is of  $\pi$  symmetry (i.e., has  $k = \pm 1$ ), so the transition dipole lies perpendicular to the molecule's axis, one obtains:

$$M = \pm 1, 0; \quad L = \pm 1, 0.$$

These selection rules are derived by realizing that in addition to  $k = \pm 1$ , one has:

(i) a linear-molecule rotational wavefunction that in the  $v = 0$  vibrational level is described in terms of a rotation matrix  $D_{L',M',0}(\alpha, \beta, \gamma)$  with no angular momentum along the molecular axis,  $K' = 0$ ; (ii) a  $v = 1$  molecule whose rotational wavefunction must be given by a rotation matrix  $D_{L,M,1}(\alpha, \beta, \gamma)$  with one unit of angular momentum about the molecule's axis,  $K = 1$ . In the latter case, the angular momentum is produced by the degenerate vibration itself. As a result, the selection rules above derive from the following product of 3-j symbols:

$$\frac{L' - 1 \quad L}{M' \quad m \quad -M} \frac{L' - 1 \quad L}{0 \quad 1 \quad -1} .$$

Because  $L = 0$  transitions are allowed for vibrations, one says that vibrations possess **Q-branches** in addition to their R- and P- branches (with  $L = 1$  and  $-1$ , respectively).

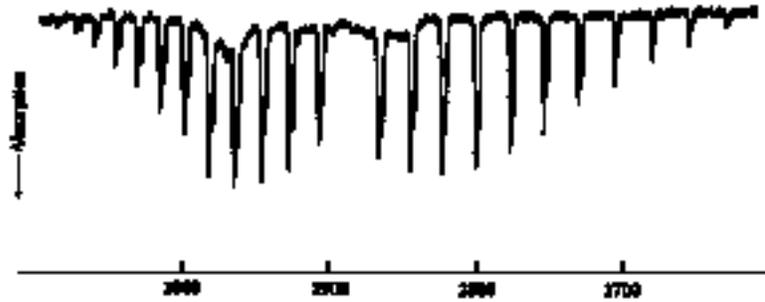
In the figure shown below, the  $v = 0 \Rightarrow v = 1$  (fundamental) vibrational absorption spectrum of HCl is shown. Here the peaks at lower energy (to the right of the figure) belong to P-branch transitions and occur at energies given approximately by:

$$\begin{aligned} E &= \hbar \nu_{\text{stretch}} + (h^2/8 \quad 2I) ((L-1)L - L(L+1)) \\ &= \hbar \nu_{\text{stretch}} - 2 (h^2/8 \quad 2I) L. \end{aligned}$$

The R-branch transitions occur at higher energies given approximately by:

$$\begin{aligned} E &= \hbar \nu_{\text{stretch}} + (h^2/8 \quad 2I) ((L+1)(L+2) - L(L+1)) \\ &= \hbar \nu_{\text{stretch}} + 2 (h^2/8 \quad 2I) (L+1). \end{aligned}$$

The absorption that is "missing" from the figure below lying slightly below  $2900 \text{ cm}^{-1}$  is the **Q-branch** transition for which  $L = L'$ ; it is absent because the selection rules forbid it.



It should be noted that the spacings between the experimentally observed peaks in HCl are not constant as would be expected based on the above P- and R- branch formulas. This is because the moment of inertia appropriate for the  $v = 1$  vibrational level is different than that of the  $v = 0$  level. These effects of vibration-rotation coupling can be modeled by allowing the  $v = 0$  and  $v = 1$  levels to have rotational energies written as

$$E = \hbar \omega_{\text{stretch}} (v + 1/2) + (h^2/8 I_v) (L(L+1))$$

where  $v$  and  $L$  are the vibrational and rotational quantum numbers. The P- and R- branch transition energies that pertain to these energy levels can then be written as:

$$E_P = \hbar \omega_{\text{stretch}} - [(h^2/8 I_1) + (h^2/8 I_0)] L + [(h^2/8 I_1) - (h^2/8 I_0)] L^2$$

$$E_R = \hbar \omega_{\text{stretch}} + 2 (h^2/8 I_1)$$

$$+ [(3h^2/8 I_1) - (h^2/8 I_0)] L + [(h^2/8 I_1) - (h^2/8 I_0)] L^2 .$$

Clearly, these formulas reduce to those shown earlier in the  $I_1 = I_0$  limit.

If the vibrationally averaged bond length is longer in the  $v = 1$  state than in the  $v = 0$  state, which is to be expected,  $I_1$  will be larger than  $I_0$ , and therefore  $[(h^2/8 I_1) - (h^2/8 I_0)]$  will be negative. In this case, the spacing between neighboring P-branch lines will increase as shown above for HCl. In contrast, the fact that  $[(h^2/8 I_1) - (h^2/8 I_0)]$  is negative causes the spacing between neighboring R- branch lines to decrease, again as shown for HCl.

### III. Electronic-Vibration-Rotation Transitions

When electronic transitions are involved, the initial and final states generally differ in their electronic, vibrational, and rotational energies. Electronic transitions usually require light in the  $5000 \text{ cm}^{-1}$  to  $100,000 \text{ cm}^{-1}$  regime, so their study lies within the domain of visible and ultraviolet spectroscopy. Excitations of inner-shell and core orbital electrons may require even higher energy photons, and under these conditions, E2 and M1 transitions may become more important because of the short wavelength of the light involved.

#### A. The Electronic Transition Dipole and Use of Point Group Symmetry

Returning to the expression

$$R_{i,f} = (2/\hbar^2) g(\nu_{f,i}) |\mathbf{E}_0 \cdot \langle f | \boldsymbol{\mu} | i \rangle|^2$$

for the rate of photon absorption, we realize that the electronic integral now involves

$$\langle e_f | \boldsymbol{\mu} | e_i \rangle = \boldsymbol{\mu}_{f,i}(\mathbf{R}),$$

a transition dipole matrix element between the initial  $e_i$  and final  $e_f$  electronic wavefunctions. This element is a function of the internal vibrational coordinates of the molecule, and again is a vector locked to the molecule's internal axis frame.

Molecular point-group symmetry can often be used to determine whether a particular transition's dipole matrix element will vanish and, as a result, the electronic transition will be "forbidden" and thus predicted to have zero intensity. If the direct product of the symmetries of the initial and final electronic states  $e_i$  and  $e_f$  do not match the symmetry of the electric dipole operator (which has the symmetry of its x, y, and z components; these symmetries can be read off the right most column of the character tables given in Appendix E), the matrix element will vanish.

For example, the formaldehyde molecule  $\text{H}_2\text{CO}$  has a ground electronic state (see Chapter 11) that has  $^1A_1$  symmetry in the  $C_{2v}$  point group. Its  $n \Rightarrow \pi^*$  singlet excited state also has  $^1A_1$  symmetry because both the  $n$  and  $\pi^*$  orbitals are of  $b_1$  symmetry. In contrast, the lowest  $n \Rightarrow \pi^*$  singlet excited state is of  $^1A_2$  symmetry because the highest energy oxygen centered  $n$  orbital is of  $b_2$  symmetry and the  $\pi^*$  orbital is of  $b_1$  symmetry, so the Slater determinant in which both the  $n$  and  $\pi^*$  orbitals are singly occupied has its symmetry dictated by the  $b_2 \times b_1$  direct product, which is  $A_2$ .

The  $\Rightarrow$  \* transition thus involves ground ( ${}^1A_1$ ) and excited ( ${}^1A_1$ ) states whose direct product ( $A_1 \times A_1$ ) is of  $A_1$  symmetry. This transition thus requires that the electric dipole operator possess a component of  $A_1$  symmetry. A glance at the  $C_{2v}$  point group's character table shows that the molecular z-axis is of  $A_1$  symmetry. Thus, if the light's electric field has a non-zero component along the  $C_2$  symmetry axis (the molecule's z-axis), the  $\Rightarrow$  \* transition is predicted to be allowed. Light polarized along either of the molecule's other two axes cannot induce this transition.

In contrast, the  $n \Rightarrow$  \* transition has a ground-excited state direct product of  $B_2 \times B_1 = A_2$  symmetry. The  $C_{2v}$ 's point group character table clearly shows that the electric dipole operator (i.e., its x, y, and z components in the molecule-fixed frame) has no component of  $A_2$  symmetry; thus, light of no electric field orientation can induce this  $n \Rightarrow$  \* transition. We thus say that the  $n \Rightarrow$  \* transition is E1 forbidden (although it is M1 allowed).

Beyond such electronic symmetry analysis, it is also possible to derive vibrational and rotational selection rules for electronic transitions that are E1 allowed. As was done in the vibrational spectroscopy case, it is conventional to expand  $\mu_{f,i}(\mathbf{R})$  in a power series about the equilibrium geometry of the initial electronic state (since this geometry is more characteristic of the molecular structure prior to photon absorption):

$$\mu_{f,i}(\mathbf{R}) = \mu_{f,i}(\mathbf{R}_e) + \sum_a \mu_{f,i}^{(a)} / R_a (R_a - R_{a,e}) + \dots$$

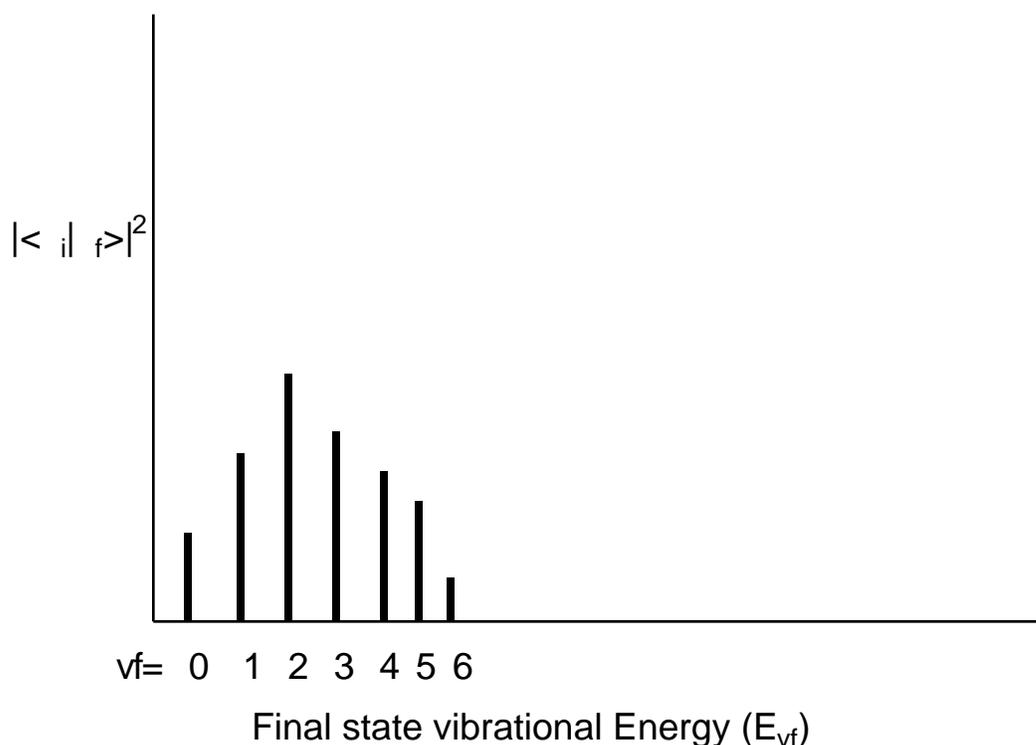
## B. The Franck-Condon Factors

The first term in this expansion, when substituted into the integral over the vibrational coordinates, gives  $\mu_{f,i}(\mathbf{R}_e) \langle \psi_f | \psi_i \rangle$ , which has the form of the electronic transition dipole multiplied by the "overlap integral" between the initial and final vibrational wavefunctions. The  $\mu_{f,i}(\mathbf{R}_e)$  factor was discussed above; it is the electronic E1 transition integral evaluated at the equilibrium geometry of the absorbing state. Symmetry can often be used to determine whether this integral vanishes, as a result of which the E1 transition will be "forbidden".

Unlike the vibration-rotation case, the vibrational overlap integrals  $\langle \psi_f | \psi_i \rangle$  do not necessarily vanish because  $\psi_f$  and  $\psi_i$  are no longer eigenfunctions of the same vibrational Hamiltonian.  $\psi_f$  is an eigenfunction whose potential energy is the final electronic state's energy surface;  $\psi_i$  has the initial electronic state's energy surface as its potential. The squares of these  $\langle \psi_f | \psi_i \rangle$  integrals, which are what eventually enter into the transition rate expression  $R_{i,f} = (2/\hbar^2) g(\psi_f, \psi_i) |\mathbf{E}_0 \cdot \langle \psi_f | \boldsymbol{\mu} | \psi_i \rangle|^2$ , are called

**"Franck-Condon factors"**. Their relative magnitudes play strong roles in determining the relative intensities of various vibrational "bands" (i.e., peaks) within a particular electronic transition's spectrum.

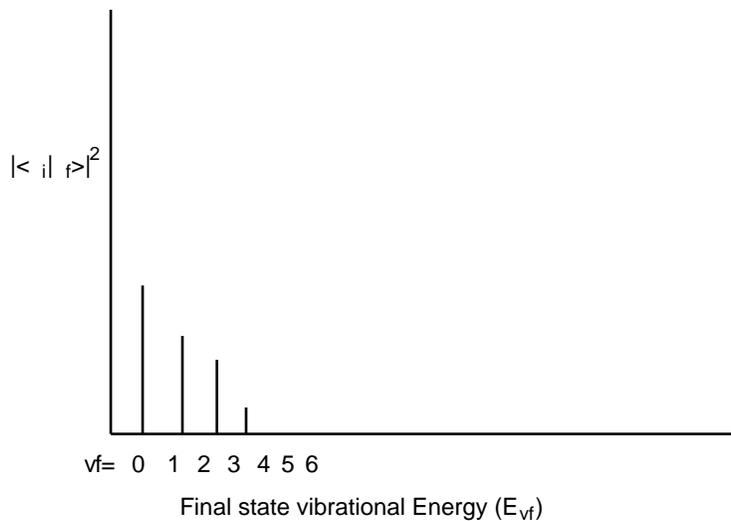
Whenever an electronic transition causes a large change in the geometry (bond lengths or angles) of the molecule, the Franck-Condon factors tend to display the characteristic "broad progression" shown below when considered for one initial-state vibrational level  $v_i$  and various final-state vibrational levels  $v_f$ :



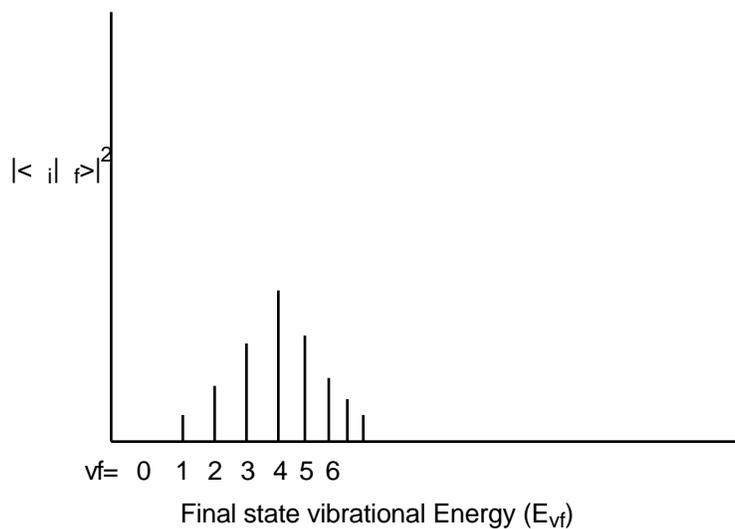
Notice that as one moves to higher  $v_f$  values, the energy spacing between the states ( $E_{v_f} - E_{v_f-1}$ ) decreases; this, of course, reflects the anharmonicity in the excited state vibrational potential. For the above example, the transition to the  $v_f = 2$  state has the largest Franck-Condon factor. This means that the overlap of the initial state's vibrational wavefunction  $v_i$  is largest for the final state's  $v_f$  function with  $v_f = 2$ .

As a qualitative rule of thumb, the larger the geometry difference between the initial and final state potentials, the broader will be the Franck-Condon profile (as shown above) and the larger the  $v_f$  value for which this profile peaks. Differences in harmonic frequencies between the two states can also broaden the Franck-Condon profile, although not as significantly as do geometry differences.

For example, if the initial and final states have very similar geometries and frequencies along the mode that is excited when the particular electronic excitation is realized, the following type of Franck-Condon profile may result:



In contrast, if the initial and final electronic states have very different geometries and/or vibrational frequencies along some mode, a very broad Franck-Condon envelope peaked at high- $vf$  will result as shown below:



### C. Vibronic Effects

The second term in the above expansion of the transition dipole matrix element  $\mu_{fi} / R_a (R_a - R_{a,e})$  can become important to analyze when the first term  $\mu_{fi}(\mathbf{R}_e)$  vanishes (e.g., for reasons of symmetry). This dipole derivative term, when substituted into the integral over vibrational coordinates gives

$\mu_{fi} / R_a \langle \nu_f | (R_a - R_{a,e}) | \nu_i \rangle$ . Transitions for which  $\mu_{fi}(\mathbf{R}_e)$  vanishes but for which  $\mu_{fi} / R_a$  does not for the  $a^{\text{th}}$  vibrational mode are said to derive intensity through "vibronic coupling" with that mode. The intensities of such modes are dependent on how strongly the electronic dipole integral varies along the mode (i.e., on  $\mu_{fi} / R_a$ ) as well as on the magnitude of the vibrational integral  $\langle \nu_f | (R_a - R_{a,e}) | \nu_i \rangle$ .

An example of an E1 forbidden but "vibronically allowed" transition is provided by the singlet  $n \Rightarrow \pi^*$  transition of  $\text{H}_2\text{CO}$  that was discussed earlier in this section. As detailed there, the ground electronic state has  $^1A_1$  symmetry, and the  $n \Rightarrow \pi^*$  state is of  $^1A_2$  symmetry, so the E1 transition integral

$\langle e_f | \mu | e_i \rangle$  vanishes for all three (x, y, z) components of the electric dipole operator  $\mu$ . However, vibrations that are of  $b_2$  symmetry (e.g., the H-C-H asymmetric stretch vibration) can induce intensity in the  $n \Rightarrow \pi^*$  transition as follows:

(i) For such vibrations, the  $b_2$  mode's  $\nu_i = 0$  to  $\nu_f = 1$  vibronic integral

$\langle \nu_f | (R_a - R_{a,e}) | \nu_i \rangle$  will be non-zero and probably quite substantial (because, for harmonic oscillator functions these "fundamental" transition integrals are dominant- see earlier);

(ii) Along these same  $b_2$  modes, the electronic transition dipole integral derivative  $\mu_{fi} / R_a$  will be non-zero, even though the integral itself  $\mu_{fi}(\mathbf{R}_e)$  vanishes when evaluated at the initial state's equilibrium geometry.

To understand why the derivative  $\mu_{fi} / R_a$  can be non-zero for distortions (denoted  $R_a$ ) of  $b_2$  symmetry, consider this quantity in greater detail:

$$\begin{aligned} \mu_{fi} / R_a &= \langle e_f | \mu | e_i \rangle / R_a \\ &= \langle e_f / R_a | \mu | e_i \rangle + \langle e_f | \mu | e_i / R_a \rangle + \langle e_f | \mu / R_a | e_i \rangle. \end{aligned}$$

The third integral vanishes because the derivative of the dipole operator itself  $\mu = \sum_i e_i \mathbf{r}_i + \sum_a Z_a e_a \mathbf{R}_a$  with respect to the coordinates of atomic centers, yields an operator that contains only a sum of scalar quantities (the elementary charge  $e$  and the

magnitudes of various atomic charges  $Z_a$ ); as a result and because the integral over the electronic wavefunctions  $\langle \psi_{ef} | \psi_{ei} \rangle$  vanishes, this contribution yields zero. The first and second integrals need not vanish by symmetry because the wavefunction derivatives  $\partial \psi_{ef} / \partial R_a$  and  $\partial \psi_{ei} / \partial R_a$  do not possess the same symmetry as their respective wavefunctions  $\psi_{ef}$  and  $\psi_{ei}$ . In fact, it can be shown that the symmetry of such a derivative is given by the direct product of the symmetries of its wavefunction and the symmetry of the vibrational mode that gives rise to the  $\partial / \partial R_a$ . For the H<sub>2</sub>CO case at hand, the b<sub>2</sub> mode vibration can induce in the excited <sup>1</sup>A<sub>2</sub> state a derivative component (i.e.,  $\partial \psi_{ef} / \partial R_a$ ) that is of <sup>1</sup>B<sub>1</sub> symmetry) and this same vibration can induce in the <sup>1</sup>A<sub>1</sub> ground state a derivative component of <sup>1</sup>B<sub>2</sub> symmetry.

As a result, the contribution  $\langle \partial \psi_{ef} / \partial R_a | \mu | \psi_{ei} \rangle$  to  $\mu_{f,i} / R_a$  arising from vibronic coupling within the excited electronic state can be expected to be non-zero for components of the dipole operator  $\mu$  that are of  $(\partial \psi_{ef} / \partial R_a \times \psi_{ei}) = (B_1 \times A_1) = B_1$  symmetry. Light polarized along the molecule's x-axis gives such a b<sub>1</sub> component to  $\mu$  (see the C<sub>2v</sub> character table in Appendix E). The second contribution  $\langle \psi_{ef} | \mu | \partial \psi_{ei} / \partial R_a \rangle$  can be non-zero for components of  $\mu$  that are of  $(\psi_{ef} \times \partial \psi_{ei} / \partial R_a) = (A_2 \times B_2) = B_1$  symmetry; again, light of x-axis polarization can induce such a transition.

In summary, electronic transitions that are E1 forbidden by symmetry can derive significant (e.g., in H<sub>2</sub>CO the singlet n  $\Rightarrow$  \* transition is rather intense) intensity through vibronic coupling. In such coupling, one or more vibrations (either in the initial or the final state) cause the respective electronic wavefunction to acquire (through  $\partial / \partial R_a$ ) a symmetry component that is different than that of itself. The symmetry of  $\partial / \partial R_a$ , which is given as the direct product of the symmetry of  $\psi$  and that of the vibration, can then cause the electric dipole integral  $\langle \mu | \partial \psi / \partial R_a \rangle$  to be non-zero even when  $\langle \mu | \psi \rangle$  is zero. Such vibronically allowed transitions are said to derive their intensity through vibronic borrowing.

#### D. Rotational Selection Rules for Electronic Transitions

Each vibrational peak within an electronic transition can also display rotational structure (depending on the spacing of the rotational lines, the resolution of the spectrometer, and the presence or absence of substantial line broadening effects such as

those discussed later in this Chapter). The selection rules for such transitions are derived in a fashion that parallels that given above for the vibration-rotation case. The major difference between this electronic case and the earlier situation is that the vibrational transition dipole moment  $\mu_{\text{trans}}$  appropriate to the former is replaced by  $\mu_{f,i}(\mathbf{R}_e)$  for conventional (i.e., non-vibronic) transitions or  $\mu_{f,i}/R_a$  (for vibronic transitions).

As before, when  $\mu_{f,i}(\mathbf{R}_e)$  (or  $\mu_{f,i}/R_a$ ) lies along the molecular axis of a linear molecule, the transition is denoted  $\sigma$  and  $k = 0$  applies; when this vector lies perpendicular to the axis it is called  $\pi$  and  $k = \pm 1$  pertains. The resultant **linear-molecule** rotational selection rules are the same as in the vibration-rotation case:

$$L = \pm 1, \text{ and } M = \pm 1, 0 \text{ (for } \sigma \text{ transitions).}$$

$$L = \pm 1, 0 \text{ and } M = \pm 1, 0 \text{ (for } \pi \text{ transitions).}$$

In the latter case, the  $L = L' = 0$  situation does not arise because a  $\pi$  transition has one unit of angular momentum along the molecular axis which would preclude both  $L$  and  $L'$  vanishing.

For **non-linear molecules** of the spherical or symmetric top variety,  $\mu_{f,i}(\mathbf{R}_e)$  (or  $\mu_{f,i}/R_a$ ) may be aligned along or perpendicular to a symmetry axis of the molecule. The selection rules that result are

$$L = \pm 1, 0; \quad M = \pm 1, 0; \text{ and } K = 0 \text{ (} L = L' = 0 \text{ is not allowed and all } L = 0 \text{ are forbidden when } K = K' = 0 \text{)}$$

which applies when  $\mu_{f,i}(\mathbf{R}_e)$  or  $\mu_{f,i}/R_a$  lies along the symmetry axis, and

$$L = \pm 1, 0; \quad M = \pm 1, 0; \text{ and } K = \pm 1 \text{ (} L = L' = 0 \text{ is not allowed)}$$

which applies when  $\mu_{f,i}(\mathbf{R}_e)$  or  $\mu_{f,i}/R_a$  lies perpendicular to the symmetry axis.

#### IV. Time Correlation Function Expressions for Transition Rates

The first-order E1 "golden-rule" expression for the rates of photon-induced transitions can be recast into a form in which certain specific physical models are easily introduced and insights are easily gained. Moreover, by using so-called equilibrium averaged time correlation functions, it is possible to obtain rate expressions appropriate to a

large number of molecules that exist in a distribution of initial states (e.g., for molecules that occupy many possible rotational and perhaps several vibrational levels at room temperature).

#### A. State-to-State Rate of Energy Absorption or Emission

To begin, the expression obtained earlier

$$R_{i,f} = (2/\hbar^2) g(\nu_{f,i}) |\mathbf{E}_0 \cdot \langle f | \boldsymbol{\mu} | i \rangle|^2,$$

that is appropriate to transitions between a particular initial state  $i$  and a specific final state  $f$ , is rewritten as

$$R_{i,f} = (2/\hbar^2) g(\nu) |\mathbf{E}_0 \cdot \langle f | \boldsymbol{\mu} | i \rangle|^2 \delta(\nu_{f,i} - \nu) d\nu.$$

Here, the  $\delta(\nu_{f,i} - \nu)$  function is used to specifically enforce the "resonance condition" that resulted in the time-dependent perturbation treatment given in Chapter 14; it states that the photons' frequency  $\nu$  must be resonant with the transition frequency  $\nu_{f,i}$ . It should be noted that by allowing  $\nu$  to run over positive and negative values, the photon absorption (with  $\nu_{f,i}$  positive and hence  $\nu$  positive) and the stimulated emission case (with  $\nu_{f,i}$  negative and hence  $\nu$  negative) are both included in this expression (as long as  $g(\nu)$  is defined as  $g(|\nu|)$  so that the negative- $\nu$  contributions are multiplied by the light source intensity at the corresponding positive  $\nu$  value).

The following integral identity can be used to replace the  $\delta$ -function:

$$\delta(\nu_{f,i} - \nu) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp[i(\nu_{f,i} - \nu)t] dt$$

by a form that is more amenable to further development. Then, the state-to-state rate of transition becomes:

$$R_{i,f} = (1/h^2) \int g(\omega) |\mathbf{E}_0 \cdot \langle f | \mu | i \rangle|^2 \exp[i(\omega_{fi} - \omega)t] dt d\omega$$

## B. Averaging Over Equilibrium Boltzmann Population of Initial States

If this expression is then multiplied by the **equilibrium probability**  $p_i$  that the molecule is found in the state  $i$  and summed over all such initial states and summed over all final states  $f$  that can be reached from  $i$  with photons of energy  $h\omega$ , the equilibrium averaged rate of photon absorption by the molecular sample is obtained:

$$R_{\text{eq.ave.}} = (1/h^2) \sum_{i,f} p_i$$

$$\int g(\omega) |\mathbf{E}_0 \cdot \langle f | \mu | i \rangle|^2 \exp[i(\omega_{fi} - \omega)t] dt d\omega$$

This expression is appropriate for an ensemble of molecules that can be in various initial states  $i$  with probabilities  $p_i$ . The corresponding result for transitions that originate in a particular state ( $i$ ) but end up in any of the "allowed" (by energy and selection rules) final states reads:

$$R_{\text{state } i} = (1/h^2) \sum_f g(\omega) |\mathbf{E}_0 \cdot \langle f | \mu | i \rangle|^2$$

$$\int \exp[i(\omega_{fi} - \omega)t] dt d\omega$$

For a canonical ensemble, in which the number of molecules, the temperature, and the system volume are specified,  $p_i$  takes the form:

$$p_i = \frac{g_i \exp(-E_i^0/kT)}{Q}$$

where  $Q$  is the canonical partition function of the molecules and  $g_i$  is the degeneracy of the state  $i$  whose energy is  $E_i^0$ .

In the above expression for  $R_{\text{eq.ave.}}$ , a double sum occurs. Writing out the elements that appear in this sum in detail, one finds:

$$\sum_{i,f} \langle i | \mathbf{E}_0 \cdot \boldsymbol{\mu} | f \rangle \langle f | \boldsymbol{\mu} | i \rangle \exp(i \omega_{fi} t).$$

In situations in which one is interested in developing an expression for the intensity arising from transitions to all allowed final states, the sum over these final states can be carried out explicitly by first writing

$$\langle f | \boldsymbol{\mu} | i \rangle \exp(i \omega_{fi} t) = \langle f | \exp(iHt/\hbar) \boldsymbol{\mu} \exp(-iHt/\hbar) | i \rangle$$

and then using the fact that the set of states  $\{ |k\rangle \}$  are complete and hence obey

$$\sum_k \langle k | k \rangle = 1.$$

The result of using these identities as well as the **Heisenberg definition** of the time-dependence of the dipole operator

$$\boldsymbol{\mu}(t) = \exp(iHt/\hbar) \boldsymbol{\mu} \exp(-iHt/\hbar),$$

is:

$$\sum_i \langle i | \mathbf{E}_0 \cdot \boldsymbol{\mu} | i \rangle \langle i | \mathbf{E}_0 \cdot \boldsymbol{\mu}(t) | i \rangle.$$

In this form, one says that the time dependence has been reduced to that of an equilibrium averaged (n.b., the  $\langle i | \mathbf{E}_0 \cdot \boldsymbol{\mu} | i \rangle$ ) **time correlation function** involving the component of the dipole operator along the external electric field at  $t = 0$  ( $\mathbf{E}_0 \cdot \boldsymbol{\mu}$ ) and this component at a different time  $t$  ( $\mathbf{E}_0 \cdot \boldsymbol{\mu}(t)$ ).

### C. Photon Emission and Absorption

If  $f_{i,j}$  is positive (i.e., in the photon absorption case), the above expression will yield a non-zero contribution when multiplied by  $\exp(-i\omega t)$  and integrated over positive  $\omega$ -values. If  $f_{i,j}$  is negative (as for stimulated photon emission), this expression will contribute, again when multiplied by  $\exp(-i\omega t)$ , for negative  $\omega$ -values. In the latter situation,  $f_i$  is the equilibrium probability of finding the molecule in the (excited) state from which emission will occur; this probability can be related to that of the lower state  $f_l$  by

$$f_{\text{excited}} = f_{\text{lower}} \exp[ -(E_{\text{excited}}^0 - E_{\text{lower}}^0)/kT ]$$

$$= f_{\text{lower}} \exp[ -\hbar\omega/kT ].$$

In this form, it is important to realize that the excited and lower states are treated as individual states, not as levels that might contain a degenerate set of states.

The absorption and emission cases can be combined into a single net expression for the rate of photon absorption by recognizing that the latter process leads to photon production, and thus must be entered with a negative sign. The resultant expression for the net rate of decrease of photons is:

$$R_{\text{eq,ave.net}} = (1/\hbar^2) \sum_i \sum_l f_l (1 - \exp(-\hbar\omega/kT))$$

$$g(\omega) < i | (\mathbf{E}_0 \cdot \boldsymbol{\mu}) | l \rangle \langle l | (\mathbf{E}_0 \cdot \boldsymbol{\mu}) | i \rangle \exp(-i\omega t) d\omega dt.$$

#### D. The Line Shape and Time Correlation Functions

Now, it is convention to introduce the so-called "line shape" function  $I(\omega)$ :

$$I(\omega) = \sum_i \sum_l < i | (\mathbf{E}_0 \cdot \boldsymbol{\mu}) | l \rangle \langle l | (\mathbf{E}_0 \cdot \boldsymbol{\mu}) | i \rangle \exp(-i\omega t) dt$$

in terms of which the net photon absorption rate is

$$R_{\text{eq.ave.net}} = (1/h^2) (1 - \exp(-h\nu/kT)) \int g(\nu) I(\nu) d\nu.$$

As stated above, the function

$$C(t) = \sum_i \sum_i \langle i | (\mathbf{E}_0 \cdot \boldsymbol{\mu}) \mathbf{E}_0 \cdot \boldsymbol{\mu}(t) | i \rangle$$

is called the equilibrium averaged **time correlation function** of the component of the electric dipole operator along the direction of the external electric field  $\mathbf{E}_0$ . Its Fourier transform is  $I(\nu)$ , the **spectral line shape** function. The convolution of  $I(\nu)$  with the light source's  $g(\nu)$  function, multiplied by  $(1 - \exp(-h\nu/kT))$ , the correction for stimulated photon emission, gives the net rate of photon absorption.

#### E. Rotational, Translational, and Vibrational Contributions to the Correlation Function

To apply the time correlation function machinery to each particular kind of spectroscopic transition, one proceeds as follows:

1. For purely **rotational transitions**, the initial and final electronic and vibrational states are the same. Moreover, the electronic and vibrational states are not summed over in the analog of the above development because one is interested in obtaining an expression for a particular  $i_v = i_e \implies f_v = f_e$  electronic-vibrational transition's lineshape. As a result, the sum over final states contained in the expression (see earlier)  $\sum_{i,f} \langle i | \boldsymbol{\mu} | f \rangle \mathbf{E}_0 \cdot \langle f | \boldsymbol{\mu}(t) | i \rangle \exp(i\omega_{fi}t)$  applies only to summing over final rotational states. In more detail, this can be shown as follows:

$$\begin{aligned} & \sum_{i,f} \langle i | \boldsymbol{\mu} | f \rangle \mathbf{E}_0 \cdot \langle f | \boldsymbol{\mu}(t) | i \rangle \\ &= \sum_{i,f} \langle i | \boldsymbol{\mu} | f \rangle \mathbf{E}_0 \cdot \langle f | \boldsymbol{\mu}(t) | i \rangle \\ &= \sum_{i,f} \langle i | \boldsymbol{\mu} | f \rangle \mathbf{E}_0 \cdot \langle f | \boldsymbol{\mu}(t) | i \rangle \\ &= \sum_{i,f} \langle i | \boldsymbol{\mu} | f \rangle \mathbf{E}_0 \cdot \langle f | \boldsymbol{\mu}(t) | i \rangle \end{aligned}$$

$$= \langle i_{ir} | i_{iv} | i_{ie} \mathbf{E}_0 \cdot \langle i_{ir} | \mu_{ave.iv} \mathbf{E}_0 \cdot \mu_{ave.iv}(t) | i_{ir} \rangle.$$

In moving from the second to the third lines of this derivation, the following identity was used:

$$\begin{aligned} \langle i_{fr} | i_{iv} | i_{ie} | \mu(t) | i_{ir} | i_{iv} | i_{ie} \rangle &= \langle i_{fr} | i_{iv} | i_{ie} | \exp(iHt/\hbar) \\ &\quad \mu \exp(-iHt/\hbar) | i_{ir} | i_{iv} | i_{ie} \rangle \\ &= \langle i_{fr} | i_{iv} | i_{ie} | \exp(iH_{v,r}t/\hbar) \mu(\mathbf{R}) \exp(-iH_{v,r}t/\hbar) | i_{ir} | i_{iv} | i_{ie} \rangle, \end{aligned}$$

where  $H$  is the full (electronic plus vibrational plus rotational) Hamiltonian and  $H_{v,r}$  is the vibrational and rotational Hamiltonian for motion on the electronic surface of the state  $i_{ie}$  whose dipole moment is  $\mu(\mathbf{R})$ . From the third line to the fourth, the (approximate) separation of rotational and vibrational motions in  $H_{v,r}$

$$H_{v,r} = H_v + H_r$$

has been used along with the fact that  $i_{iv}$  is an eigenfunction of  $H_v$ :

$$H_v | i_{iv} \rangle = E_{iv} | i_{iv} \rangle$$

to write

$$\begin{aligned} \langle i_{iv} | \mu(\mathbf{R},t) | i_{iv} \rangle &= \exp(i H_r t/\hbar) \langle i_{iv} | \exp(i H_v t/\hbar) \\ &\quad \mu(\mathbf{R}) \exp(-i H_v t/\hbar) | i_{iv} \rangle \exp(-i H_r t/\hbar) \\ &= \exp(i H_r t/\hbar) \langle i_{iv} | \exp(i E_{iv} t/\hbar) \\ &\quad \mu(\mathbf{R}) \exp(-i E_{iv} t/\hbar) | i_{iv} \rangle \exp(-i H_r t/\hbar) \\ &= \exp(i H_r t/\hbar) \langle i_{iv} | \mu(\mathbf{R}) | i_{iv} \rangle \exp(-i H_r t/\hbar) \end{aligned}$$

$$= \mu_{\text{ave,iv}}(t).$$

In effect,  $\mu$  is replaced by the vibrationally averaged electronic dipole moment  $\mu_{\text{ave,iv}}$  for each initial vibrational state that can be involved, and the time correlation function thus becomes:

$$C(t) = \sum_{i, ir, iv, ie} \langle i | \mathbf{E}_0 \cdot \mu_{\text{ave,iv}} | ir \rangle \langle ir | \mathbf{E}_0 \cdot \mu_{\text{ave,iv}}(t) | iv \rangle \langle iv | ie \rangle,$$

where  $\mu_{\text{ave,iv}}(t)$  is the averaged dipole moment for the vibrational state  $iv$  at time  $t$ , given that it was  $\mu_{\text{ave,iv}}$  at time  $t = 0$ . The time dependence of  $\mu_{\text{ave,iv}}(t)$  is induced by the rotational Hamiltonian  $H_r$ , as shown clearly in the steps detailed above:

$$\mu_{\text{ave,iv}}(t) = \exp(i H_r t/\hbar) \langle iv | \mu(\mathbf{R}) | iv \rangle \exp(-i H_r t/\hbar).$$

In this particular case, the equilibrium average is taken over the initial rotational states whose probabilities are denoted  $ir$ , any initial vibrational states that may be populated, with probabilities  $iv$ , and any populated electronic states, with probabilities  $ie$ .

2. For **vibration-rotation transitions** within a single electronic state, the initial and final electronic states are the same, but the initial and final vibrational and rotational states differ. As a result, the sum over final states contained in the expression  $\sum_{i, f} \langle i | \mathbf{E}_0 \cdot \mu | f \rangle \langle f | \mathbf{E}_0 \cdot \mu | i \rangle \exp(i \omega_{fi} t)$  applies only to summing over final vibrational and rotational states. Paralleling the development made in the pure rotation case given above, this can be shown as follows:

$$\begin{aligned} & \sum_{i, f} \langle i | \mathbf{E}_0 \cdot \mu | f \rangle \langle f | \mathbf{E}_0 \cdot \mu | i \rangle \exp(i \omega_{fi} t) \\ &= \sum_{i, f} \langle i | \mathbf{E}_0 \cdot \mu | fr \rangle \langle fr | \mathbf{E}_0 \cdot \mu | iv \rangle \langle iv | ie \rangle \langle ie | \mu(t) | ir \rangle \langle ir | iv \rangle \\ &= \sum_{i, f} \langle i | \mathbf{E}_0 \cdot \mu | fr \rangle \langle fr | \mathbf{E}_0 \cdot \mu | iv \rangle \langle iv | ie \rangle \langle ie | \mu(\mathbf{R}, t) | ir \rangle \langle ir | iv \rangle \\ &= \sum_{i, f} \langle i | \mathbf{E}_0 \cdot \mu | fr \rangle \langle fr | \mathbf{E}_0 \cdot \mu | iv \rangle \langle iv | ie \rangle \langle ie | \mu(\mathbf{R}_e) + \mu_a (\mathbf{R}_a - \mathbf{R}_{a,eq}) / R_a | fr \rangle \langle fr | iv \rangle \end{aligned}$$

$$\begin{aligned}
& \mathbf{E}_0 \cdot \langle i_r | \exp(iH_r t/\hbar) (\mu(\mathbf{R}_e) + \mu_a (\mathbf{R}_a - \mathbf{R}_{a,eq}) / R_a) \\
& \exp(-iH_r t/\hbar) | i_r \rangle \exp(i \mathbf{E}_0 \cdot \mathbf{r}_{fv,iv} t) \\
& = \langle i_r, i_v, i_e | \exp(iH_r t/\hbar) (\mu(\mathbf{R}_e) + \mu_a (\mathbf{R}_a - \mathbf{R}_{a,eq}) / R_a) \\
& \exp(-iH_r t/\hbar) | i_r \rangle \exp(i \mathbf{E}_0 \cdot \mathbf{r}_{fv,iv} t) \\
& = \langle i_r | \mu_a (\mathbf{R}_a - \mathbf{R}_{a,eq}) / R_a \mathbf{E}_0 \cdot \exp(iH_r t/\hbar) \mu_a (\mathbf{R}_a - \mathbf{R}_{a,eq}) / R_a \exp(-iH_r t/\hbar) | i_r \rangle \\
& = \langle i_r | \mu_a (\mathbf{R}_a - \mathbf{R}_{a,eq}) / R_a \mathbf{E}_0 \cdot \exp(iH_r t/\hbar) \mu_a (\mathbf{R}_a - \mathbf{R}_{a,eq}) / R_a \exp(-iH_r t/\hbar) | i_r \rangle \\
& = \langle i_r | (\mathbf{E}_0 \cdot \mu_{trans}) \mathbf{E}_0 \cdot \mu_{trans} \exp(iH_r t/\hbar) \exp(-iH_r t/\hbar) | i_r \rangle,
\end{aligned}$$

where the vibrational transition dipole matrix element is defined as before

$$\mu_{trans} = \langle i_r | \mu_a (\mathbf{R}_a - \mathbf{R}_{a,eq}) / R_a | i_r \rangle$$

and derives its time dependence above from the rotational Hamiltonian:

$$\mu_{trans}(t) = \exp(iH_r t/\hbar) \mu_{trans} \exp(-iH_r t/\hbar).$$

The corresponding final expression for the time correlation function C(t) becomes:

$$C(t) = \langle i_r | (\mathbf{E}_0 \cdot \mu_{trans}) \mathbf{E}_0 \cdot \mu_{trans}(t) | i_r \rangle \exp(i \mathbf{E}_0 \cdot \mathbf{r}_{fv,iv} t).$$

The net rate of photon absorption remains:

$$R_{eq,ave.net} = (1/\hbar^2) (1 - \exp(-\hbar \omega)) g(\omega) I(\omega) d\omega,$$

where I(ω) is the Fourier transform of C(t).

The expression for C(t) clearly contains two types of time dependences: (i) the exp(i E<sub>0</sub> · r<sub>fv,iv</sub> t), upon Fourier transforming to obtain I(ω), produces δ-function "spikes" at

frequencies  $\omega_{i,f} = \omega_{v,i,v}$  equal to the spacings between the initial and final vibrational states, and (ii) rotational motion time dependence that causes  $\mu_{\text{trans}}(t)$  to change with time. The latter appears in the form of a correlation function for the component of  $\mu_{\text{trans}}$  along  $\mathbf{E}_0$  at time  $t = 0$  and this component at another time  $t$ . The convolution of both these time dependences determines the form of  $I(\omega)$ .

3. For **electronic-vibration-rotation transitions**, the initial and final electronic states are different as are the initial and final vibrational and rotational states. As a result, the sum over final states contained in the expression  $\sum_{i,f} \langle i | \mu | f \rangle \mathbf{E}_0 \cdot \langle f | \mu | i \rangle \exp(i\omega_{f,i}t)$  applies to summing over final electronic, vibrational, and rotational states. Paralleling the development made in the pure rotation case given above, this can be shown as follows:

$$\begin{aligned}
 & \sum_{i,f} \langle i | \mu | f \rangle \mathbf{E}_0 \cdot \langle f | \mu(t) | i \rangle \\
 &= \sum_{i,f} \langle i | \mu | f \rangle \mathbf{E}_0 \cdot \langle i_r | \mu | f_r \rangle \langle i_v | \mu | f_v \rangle \langle i_e | \mu | f_e \rangle \mathbf{E}_0 \cdot \langle f_r | \mu(t) | i_r \rangle \langle f_v | \mu(t) | i_v \rangle \langle f_e | \mu(t) | i_e \rangle \\
 &= \sum_{i,f} \langle i_r | \mu | f_r \rangle \langle i_v | \mu | f_v \rangle \langle i_e | \mu | f_e \rangle \mathbf{E}_0 \cdot \langle i_r | \mu_{i,f}(\mathbf{R}) | f_r \rangle \langle f_r | \mu_{i,f}(\mathbf{R},t) | i_r \rangle \langle i_v | \mu_{i,f}(\mathbf{R}) | f_v \rangle \langle f_v | \mu_{i,f}(\mathbf{R},t) | i_v \rangle \langle i_e | \mu_{i,f}(\mathbf{R}) | f_e \rangle \langle f_e | \mu_{i,f}(\mathbf{R},t) | i_e \rangle \\
 &= \sum_{i,f} \langle i_r | \mu_{i,f}(\mathbf{R}_e) | f_r \rangle \langle i_v | \mu_{i,f}(\mathbf{R}_e) | f_v \rangle \langle i_e | \mu_{i,f}(\mathbf{R}_e) | f_e \rangle \mathbf{E}_0 \cdot \langle f_r | \exp(iH_r t/\hbar) \mu_{i,f}(\mathbf{R}_e) \exp(-iH_r t/\hbar) | i_r \rangle \exp(i\omega_{f_v,i_v}t + i E_{i,f}t/\hbar) \\
 &= \sum_{i,f} \langle i_r | \mu_{i,f}(\mathbf{R}_e) | f_r \rangle \langle i_v | \mu_{i,f}(\mathbf{R}_e) | f_v \rangle \langle i_e | \mu_{i,f}(\mathbf{R}_e) | f_e \rangle \mathbf{E}_0 \cdot \langle f_r | \mu_{i,f}(\mathbf{R}_e,t) | i_r \rangle \langle f_v | \mu_{i,f}(\mathbf{R}_e,t) | i_v \rangle \langle f_e | \mu_{i,f}(\mathbf{R}_e,t) | i_e \rangle \exp(i\omega_{f_v,i_v}t + i E_{i,f}t/\hbar),
 \end{aligned}$$

where

$$\mu_{i,f}(\mathbf{R}_e,t) = \exp(iH_r t/\hbar) \mu_{i,f}(\mathbf{R}_e) \exp(-iH_r t/\hbar)$$

is the electronic transition dipole matrix element, evaluated at the equilibrium geometry of the absorbing state, that derives its time dependence from the rotational Hamiltonian  $H_r$  as in the time correlation functions treated earlier.

This development thus leads to the following definition of  $C(t)$  for the electronic, vibration, and rotation case:

$$C(t) = \langle i, f | \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e) \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e, t) | i, r \rangle \langle i, v | \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e) \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e, t) | i, v \rangle^2 \exp(i \omega_{fv,iv} t + i E_{i,f} t / \hbar)$$

but the net rate of photon absorption remains:

$$R_{\text{eq.ave.net}} = (1/\hbar^2) (1 - \exp(-\hbar \omega_{fv,iv} / kT)) \int g(\omega) I(\omega) d\omega$$

Here,  $I(\omega)$  is the Fourier transform of the above  $C(t)$  and  $E_{i,f}$  is the adiabatic electronic energy difference (i.e., the energy difference between the  $v = 0$  level in the final electronic state and the  $v = 0$  level in the initial electronic state) for the electronic transition of interest. The above  $C(t)$  clearly contains Franck-Condon factors as well as time dependence  $\exp(i \omega_{fv,iv} t + i E_{i,f} t / \hbar)$  that produces  $\delta$ -function spikes at each electronic-vibrational transition frequency and rotational time dependence contained in the time correlation function quantity  $\langle i, r | \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e) \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e, t) | i, r \rangle$ .

To summarize, the line shape function  $I(\omega)$  produces the net rate of photon absorption

$$R_{\text{eq.ave.net}} = (1/\hbar^2) (1 - \exp(-\hbar \omega_{fv,iv} / kT)) \int g(\omega) I(\omega) d\omega$$

in all of the above cases, and  $I(\omega)$  is the Fourier transform of a corresponding time-dependent  $C(t)$  function in all cases. However, the pure rotation, vibration-rotation, and electronic-vibration-rotation cases differ in the form of their respective  $C(t)$ 's. Specifically,

$$C(t) = \langle i | \mathbf{E}_0 \cdot \boldsymbol{\mu}_{\text{ave},iv} | i, r \rangle \langle i, v | \mathbf{E}_0 \cdot \boldsymbol{\mu}_{\text{ave},iv}(t) | i, v \rangle$$

in the pure rotational case,

$$C(t) = \langle i_{ir} | \mathbf{E}_0 \cdot \mu_{\text{trans}} | i_{iv} \rangle \exp(i \omega_{iv} t)$$

in the vibration-rotation case, and

$$C(t) = \langle i_{ir} | \mathbf{E}_0 \cdot \mu_{i,f}(\mathbf{R}_e) | i_{iv} \rangle \langle i_{iv} | \mathbf{E}_0 \cdot \mu_{i,f}(\mathbf{R}_e, t) | i_{fv} \rangle^2 \exp(i \omega_{iv} t + i E_{i,f} t / \hbar)$$

in the electronic-vibration-rotation case.

All of these time correlation functions contain time dependences that arise from rotational motion of a dipole-related vector (i.e., the vibrationally averaged dipole  $\mu_{\text{ave},iv}(t)$ , the vibrational transition dipole  $\mu_{\text{trans}}(t)$ , or the electronic transition dipole  $\mu_{i,f}(\mathbf{R}_e, t)$ ) and the latter two also contain oscillatory time dependences (i.e.,  $\exp(i \omega_{iv} t)$  or  $\exp(i \omega_{iv} t + i E_{i,f} t / \hbar)$ ) that arise from vibrational or electronic-vibrational energy level differences. In the treatments of the following sections, consideration is given to the rotational contributions under circumstances that characterize, for example, dilute gaseous samples where the collision frequency is low and liquid-phase samples where rotational motion is better described in terms of diffusional motion.

## F. Line Broadening Mechanisms

If the rotational motion of the molecules is assumed to be entirely unhindered (e.g., by any environment or by collisions with other molecules), it is appropriate to express the time dependence of each of the dipole time correlation functions listed above in terms of a "free rotation" model. For example, when dealing with diatomic molecules, the electronic-vibrational-rotational  $C(t)$  appropriate to a specific electronic-vibrational transition becomes:

$$C(t) = (q_r q_v q_e q_t)^{-1} \frac{J}{2J+1} \exp(-h^2 J(J+1) / (8 I k T)) \exp(-h \nu_{\text{vib}} t / k T) \langle i_{ie} < J | \mathbf{E}_0 \cdot \mu_{i,f}(\mathbf{R}_e) | J \rangle \langle J | \mathbf{E}_0 \cdot \mu_{i,f}(\mathbf{R}_e, t) | J \rangle \langle i_{iv} | \mathbf{E}_0 \cdot \mu_{i,f}(\mathbf{R}_e, t) | i_{fv} \rangle^2 \exp(i [\omega_{\text{vib}}] t + i E_{i,f} t / \hbar).$$

Here,

$$q_r = (8 \pi^2 I k T / h^2)$$

is the rotational partition function ( $I$  being the molecule's moment of inertia  $I = \mu R_e^2$ , and  $h^2 J(J+1)/(8 \pi^2 I)$  the molecule's rotational energy for the state with quantum number  $J$  and degeneracy  $2J+1$ )

$$q_v = \exp(-h \nu_{vib}/2kT) (1 - \exp(-h \nu_{vib}/kT))^{-1}$$

is the vibrational partition function ( $\nu_{vib}$  being the vibrational frequency),  $g_{ie}$  is the degeneracy of the initial electronic state,

$$q_t = (2 \pi m k T / h^2)^{3/2} V$$

is the translational partition function for the molecules of mass  $m$  moving in volume  $V$ , and  $E_{i,f}$  is the adiabatic electronic energy spacing.

The functions  $\langle J | \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e) \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e, t) | J \rangle$  describe the time evolution of the dipole-related vector (the electronic transition dipole in this case) for the rotational state  $J$ . In a "free-rotation" model, this function is taken to be of the form:

$$\begin{aligned} & \langle J | \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e) \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e, t) | J \rangle \\ & = \langle J | \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e) \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e, 0) | J \rangle \cos \frac{h J(J+1) t}{4 I} \end{aligned}$$

where

$$\frac{h J(J+1)}{4 I} = \nu_J$$

is the rotational frequency (in cycles per second) for rotation of the molecule in the state labeled by  $J$ . This oscillatory time dependence, combined with the  $\exp(i \nu_{v,iv} t + i E_{i,f} t / \hbar)$  time dependence arising from the electronic and vibrational factors, produce, when this  $C(t)$  function is Fourier transformed to generate  $I(\omega)$  a series of  $\delta$ -function "peaks" whenever

$$\omega = \nu_{v,iv} + E_{i,f}/\hbar \pm \nu_J$$

The intensities of these peaks are governed by the

$$(q_r q_v q_e q_t)^{-1} \frac{J+1}{2J+1} \exp(-h^2 J(J+1)/(8 I k T)) \exp(-h \nu_{\text{vib}}/kT) g_e$$

Boltzmann population factors as well as by the  $|\langle i_v | f_v \rangle|^2$  Franck-Condon factors and the  $\langle J | \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e) \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e, 0) | J \rangle$  terms.

This same analysis can be applied to the pure rotation and vibration-rotation  $C(t)$  time dependences with analogous results. In the former,  $\delta$ -function peaks are predicted to occur at

$$\omega = \pm J$$

and in the latter at

$$\omega = \nu_{v,i} \pm J;$$

with the intensities governed by the time independent factors in the corresponding expressions for  $C(t)$ .

In experimental measurements, such sharp  $\delta$ -function peaks are, of course, not observed. Even when very narrow band width laser light sources are used (i.e., for which  $g(\omega)$  is an extremely narrowly peaked function), spectral lines are found to possess finite widths. Let us now discuss several sources of line broadening, some of which will relate to deviations from the "unhindered" rotational motion model introduced above.

### 1. Doppler Broadening

In the above expressions for  $C(t)$ , the averaging over initial rotational, vibrational, and electronic states is explicitly shown. There is also an average over the translational motion implicit in all of these expressions. Its role has not (yet) been emphasized because the molecular energy levels, whose spacings yield the characteristic frequencies at which light can be absorbed or emitted, do not depend on translational motion. However, the frequency of the electromagnetic field experienced by moving molecules does depend on the velocities of the molecules, so this issue must now be addressed.

Elementary physics classes express the so-called **Doppler shift** of a wave's frequency induced by movement either of the light source or of the molecule (Einstein tells us these two points of view must give identical results) as follows:

$$\text{observed} = \text{nominal} (1 + v_z/c)^{-1} \text{nominal} (1 - v_z/c + \dots).$$

Here,  $\nu_{\text{nominal}}$  is the frequency of the unmoving light source seen by unmoving molecules,  $v_z$  is the velocity of relative motion of the light source and molecules,  $c$  is the speed of light, and  $\nu_{\text{observed}}$  is the Doppler shifted frequency (i.e., the frequency seen by the molecules). The second identity is obtained by expanding, in a power series, the  $(1 + v_z/c)^{-1}$  factor, and is valid in truncated form when the molecules are moving with speeds significantly below the speed of light.

For all of the cases considered earlier, a  $C(t)$  function is subjected to Fourier transformation to obtain a spectral lineshape function  $I(\omega)$ , which then provides the essential ingredient for computing the net rate of photon absorption. In this Fourier transform process, the variable  $\omega$  is assumed to be the frequency of the electromagnetic field experienced by the molecules. The above considerations of Doppler shifting then leads one to realize that the correct functional form to use in converting  $C(t)$  to  $I(\omega)$  is:

$$I(\omega) = \int C(t) \exp(-it(\omega - \nu_{\text{nominal}}(1 - v_z/c))) dt,$$

where  $\nu_{\text{nominal}}$  is the nominal frequency of the light source.

As stated earlier, within  $C(t)$  there is also an equilibrium average over translational motion of the molecules. For a gas-phase sample undergoing random collisions and at thermal equilibrium, this average is characterized by the well known Maxwell-Boltzmann velocity distribution:

$$\left(\frac{m}{2\pi kT}\right)^{3/2} \exp(-m(v_x^2 + v_y^2 + v_z^2)/2kT) dv_x dv_y dv_z.$$

Here  $m$  is the mass of the molecules and  $v_x$ ,  $v_y$ , and  $v_z$  label the velocities along the lab-fixed cartesian coordinates.

Defining the  $z$ -axis as the direction of propagation of the light's photons and carrying out the averaging of the Doppler factor over such a velocity distribution, one obtains:

$$\int \exp(-it(\omega - \nu_{\text{nominal}}(1 - v_z/c))) \left(\frac{m}{2\pi kT}\right)^{3/2} \exp(-m(v_x^2 + v_y^2 + v_z^2)/2kT) dv_x dv_y dv_z$$

$$= \exp(-i \omega t) (m/2kT)^{1/2} \exp(i v_z/c) \exp(-mv_z^2/2kT) dv_z$$

$$= \exp(-i \omega t) \exp(-\frac{1}{2}t^2kT/(2mc^2)).$$

This result, when substituted into the expressions for  $C(t)$ , yields expressions identical to those given for the three cases treated above but with one modification. The translational motion average need no longer be considered in each  $C(t)$ ; instead, the earlier expressions for  $C(t)$  must each be multiplied by a factor  $\exp(-\frac{1}{2}t^2kT/(2mc^2))$  that embodies the translationally averaged Doppler shift. The spectral line shape function  $I(\omega)$  can then be obtained for each  $C(t)$  by simply Fourier transforming:

$$I(\omega) = \int_{-\infty}^{\infty} \exp(-i \omega t) C(t) dt .$$

When applied to the rotation, vibration-rotation, or electronic-vibration-rotation cases within the "unhindered" rotation model treated earlier, the Fourier transform involves integrals of the form:

$$I(\omega) = \int_{-\infty}^{\infty} \exp(-i \omega t) \exp(-\frac{1}{2}t^2kT/(2mc^2)) \exp(i(\nu_{v,iv} + E_{i,f}/\hbar \pm J)t) dt .$$

This integral would arise in the electronic-vibration-rotation case; the other two cases would involve integrals of the same form but with the  $E_{i,f}/\hbar$  absent in the vibration-rotation situation and with  $\nu_{v,iv} + E_{i,f}/\hbar$  missing for pure rotation transitions. All such integrals can be carried out analytically and yield:

$$I(\omega) = \sqrt{\frac{2mc^2}{2kT}} \exp[-(\omega - \nu_{v,iv} - E_{i,f}/\hbar \pm J)^2 mc^2/(2kT)].$$

The result is a series of **Gaussian** "peaks" in  $\omega$ -space, centered at:

$$= \nu_{i,v} + E_{i,f}/h \pm J$$

with widths ( ) determined by

$$\Delta\nu = \sqrt{2kT/(mc^2)},$$

given the temperature T and the mass of the molecules m. The hotter the sample, the faster the molecules are moving on average, and the broader is the distribution of Doppler shifted frequencies experienced by these molecules. The net result then of the Doppler effect is to produce a line shape function that is similar to the "unhindered" rotation model's series of  $J$ -functions but with each  $J$ -function peak broadened into a Gaussian shape.

## 2. Pressure Broadening

To include the effects of collisions on the rotational motion part of any of the above  $C(t)$  functions, one must introduce a model for how such collisions change the dipole-related vectors that enter into  $C(t)$ . The most elementary model used to address collisions applies to gaseous samples which are assumed to undergo unhindered rotational motion until struck by another molecule at which time a randomizing "kick" is applied to the dipole vector and after which the molecule returns to its unhindered rotational movement.

The effects of such collisionally induced kicks are treated within the so-called **pressure broadening** (sometimes called collisional broadening) model by modifying the free-rotation correlation function through the introduction of an exponential damping factor  $\exp(-|t|/\tau)$ :

$$\langle J | \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e) \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e, 0) | J \rangle \cos \frac{\hbar J(J+1) t}{4 I}$$

$$\langle J | \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e) \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e, 0) | J \rangle \cos \frac{\hbar J(J+1) t}{4 I} \exp(-|t|/\tau).$$

This damping function's time scale parameter  $\tau$  is assumed to characterize the average time between collisions and thus should be inversely proportional to the collision frequency. Its magnitude is also related to the effectiveness with which collisions cause the dipole function to deviate from its unhindered rotational motion (i.e., related to the collision strength). In effect, the exponential damping causes the time correlation function  $\langle J | \mathbf{E}_0 \cdot$

$\mu_{i,f}(\mathbf{R}_e) \mathbf{E}_0 \cdot \mu_{i,f}(\mathbf{R}_e, t) | J \rangle$  to "lose its memory" and to decay to zero; this "memory" point of view is based on viewing  $\langle J | \mathbf{E}_0 \cdot \mu_{i,f}(\mathbf{R}_e) \mathbf{E}_0 \cdot \mu_{i,f}(\mathbf{R}_e, t) | J \rangle$  as the projection of  $\mathbf{E}_0 \cdot \mu_{i,f}(\mathbf{R}_e, t)$  along its  $t = 0$  value  $\mathbf{E}_0 \cdot \mu_{i,f}(\mathbf{R}_e, 0)$  as a function of time  $t$ .

Introducing this additional  $\exp(-|t|/\tau)$  time dependence into  $C(t)$  produces, when  $C(t)$  is Fourier transformed to generate  $I(\omega)$ ,

$$I(\omega) = \int_{-\infty}^{\infty} \exp(-i\omega t) \exp(-|t|/\tau) \exp(-\frac{1}{2} t^2 kT / (2mc^2)) \exp(i(\omega_{v,iv} + E_{i,f}/\hbar \pm J)t) dt.$$

In the limit of very small Doppler broadening, the  $(\frac{1}{2} t^2 kT / (2mc^2))$  factor can be ignored (i.e.,  $\exp(-\frac{1}{2} t^2 kT / (2mc^2))$  set equal to unity), and

$$I(\omega) = \int_{-\infty}^{\infty} \exp(-i\omega t) \exp(-|t|/\tau) \exp(i(\omega_{v,iv} + E_{i,f}/\hbar \pm J)t) dt$$

results. This integral can be performed analytically and generates:

$$I(\omega) = \frac{1}{4} \left\{ \frac{1/\tau}{(1/\tau)^2 + (\omega - \omega_{v,iv} - E_{i,f}/\hbar \pm J)^2} + \frac{1/\tau}{(1/\tau)^2 + (\omega + \omega_{v,iv} + E_{i,f}/\hbar \pm J)^2} \right\},$$

a pair of **Lorentzian** peaks in  $\omega$ -space centered again at

$$\omega = \pm [\omega_{v,iv} + E_{i,f}/\hbar \pm J].$$

The full width at half height of these Lorentzian peaks is  $2/\tau$ . One says that the individual peaks have been pressure or collisionally broadened.

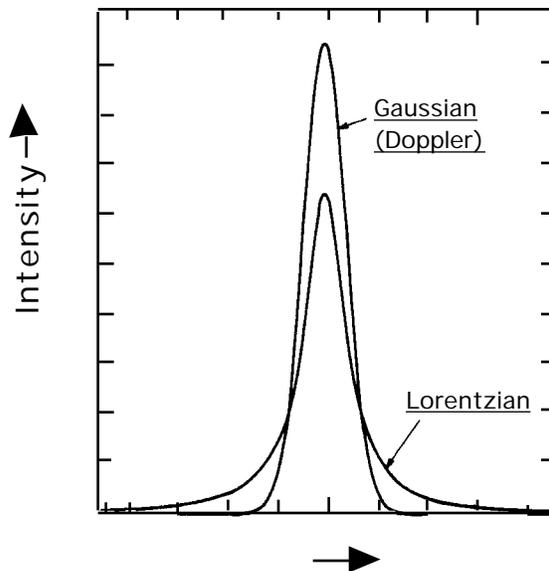
When the Doppler broadening can not be neglected relative to the collisional broadening, the above integral

$$I(\omega) = \int_{-\infty}^{\infty} \exp(-i\omega t) \exp(-|t|/\tau) \exp(-\frac{1}{2} t^2 kT / (2mc^2)) \exp(i(\omega_{v,iv} + E_{i,f}/\hbar \pm J)t) dt$$

is more difficult to perform. Nevertheless, it can be carried out and again produces a series of peaks centered at

$$= \nu_{i,i'} + E_{i,f}/h \pm J$$

but whose widths are determined both by Doppler and pressure broadening effects. The resultant line shapes are thus no longer purely Lorentzian nor Gaussian (which are compared in the figure below for both functions having the same full width at half height and the same integrated area), but have a shape that is called a **Voigt** shape.



### 3. Rotational Diffusion Broadening

Molecules in liquids and very dense gases undergo frequent collisions with the other molecules; that is, the mean time between collisions is short compared to the rotational period for their unhindered rotation. As a result, the time dependence of the dipole related correlation function can no longer be modeled in terms of free rotation that is interrupted by (infrequent) collisions and Doppler shifted. Instead, a model that describes the incessant buffeting of the molecule's dipole by surrounding molecules becomes appropriate. For liquid samples in which these frequent collisions cause the molecule's dipole to undergo angular motions that cover all angles (i.e., in contrast to a frozen glass or

solid in which the molecule's dipole would undergo strongly perturbed pendular motion about some favored orientation), the so-called **rotational diffusion** model is often used.

In this picture, the rotation-dependent part of  $C(t)$  is expressed as:

$$\begin{aligned} & \langle J | \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e) \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e, t) | J \rangle \\ & = \langle J | \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e) \mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e, 0) | J \rangle \exp(-2D_{\text{rot}}|t|), \end{aligned}$$

where  $D_{\text{rot}}$  is the rotational diffusion constant whose magnitude details the time decay in the averaged value of  $\mathbf{E}_0 \cdot \boldsymbol{\mu}_{i,f}(\mathbf{R}_e, t)$  at time  $t$  with respect to its value at time  $t = 0$ ; the larger  $D_{\text{rot}}$ , the faster is this decay.

As with pressure broadening, this exponential time dependence, when subjected to Fourier transformation, yields:

$$I(\omega) = \int_{-\infty}^{\infty} \exp(-i\omega t) \exp(-2D_{\text{rot}}|t|) \exp(-\frac{1}{2}t^2kT/(2mc^2)) \exp(i(\nu_{i,v} + E_{i,f}/\hbar \pm J)t) dt.$$

Again, in the limit of very small Doppler broadening, the  $(\frac{1}{2}t^2kT/(2mc^2))$  factor can be ignored (i.e.,  $\exp(-\frac{1}{2}t^2kT/(2mc^2))$  set equal to unity), and

$$I(\omega) = \int_{-\infty}^{\infty} \exp(-i\omega t) \exp(-2D_{\text{rot}}|t|) \exp(i(\nu_{i,v} + E_{i,f}/\hbar \pm J)t) dt$$

results. This integral can be evaluated analytically and generates:

$$\begin{aligned} I(\omega) = \frac{1}{4} \left\{ \frac{2D_{\text{rot}}}{(2D_{\text{rot}})^2 + (\omega - \nu_{i,v} - E_{i,f}/\hbar \pm J)^2} \right. \\ \left. + \frac{2D_{\text{rot}}}{(2D_{\text{rot}})^2 + (\omega + \nu_{i,v} + E_{i,f}/\hbar \pm J)^2} \right\}, \end{aligned}$$

a pair of **Lorentzian** peaks in  $\omega$ -space centered again at

$$\omega = \pm [\nu_{i,v} + E_{i,f}/\hbar \pm J].$$

The full width at half height of these Lorentzian peaks is  $4D_{\text{rot}}$ . In this case, one says that the individual peaks have been broadened via rotational diffusion. When the Doppler broadening can not be neglected relative to the collisional broadening, the above integral

$$I(\nu) = \int_{-\infty}^{\infty} \exp(-i(\nu - \nu_0)t) \exp(-2D_{\text{rot}}|t|) \exp(-\frac{1}{2}t^2kT/(2mc^2)) \exp(i(\nu_0 - \nu_{\text{iv}} + E_{i,f}/\hbar \pm J)t) dt .$$

is more difficult to perform. Nevertheless, it can be carried out and again produces a series of peaks centered at

$$\nu = \nu_0 \pm [\nu_0 - \nu_{\text{iv}} + E_{i,f}/\hbar \pm J]$$

but whose widths are determined both by Doppler and rotational diffusion effects.

#### 4. Lifetime or Heisenberg Homogeneous Broadening

Whenever the absorbing species undergoes one or more processes that depletes its numbers, we say that it has a finite lifetime. For example, a species that undergoes unimolecular dissociation has a finite lifetime, as does an excited state of a molecule that decays by spontaneous emission of a photon. Any process that depletes the absorbing species contributes another source of time dependence for the dipole time correlation functions  $C(t)$  discussed above. This time dependence is usually modeled by appending, in a multiplicative manner, a factor  $\exp(-|t|/\tau)$ . This, in turn modifies the line shape function  $I(\nu)$  in a manner much like that discussed when treating the rotational diffusion case:

$$I(\nu) = \int_{-\infty}^{\infty} \exp(-i(\nu - \nu_0)t) \exp(-|t|/\tau) \exp(-\frac{1}{2}t^2kT/(2mc^2)) \exp(i(\nu_0 - \nu_{\text{iv}} + E_{i,f}/\hbar \pm J)t) dt .$$

Not surprisingly, when the Doppler contribution is small, one obtains:

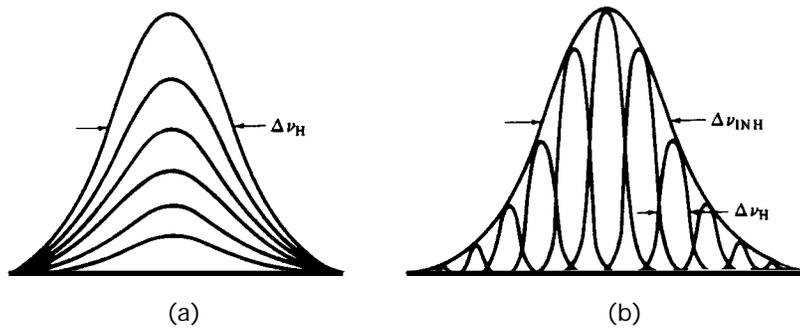
$$I(\nu) = \frac{1}{4} \left\{ \frac{1/\tau}{(1/\tau)^2 + (\nu_0 - \nu_{\text{iv}} - E_{i,f}/\hbar \pm J)^2} + \frac{1/\tau}{(1/\tau)^2 + (\nu_0 + \nu_{\text{iv}} + E_{i,f}/\hbar \pm J)^2} \right\} .$$

In these Lorentzian lines, the parameter  $\tau$  describes the kinetic decay lifetime of the molecule. One says that the spectral lines have been **lifetime or Heisenberg broadened** by an amount proportional to  $1/\tau$ . The latter terminology arises because the finite lifetime of the molecular states can be viewed as producing, via the Heisenberg uncertainty relation  $\Delta E \tau > \hbar$ , states whose energy is "uncertain" to within an amount  $\Delta E$ .

## 5. Site Inhomogeneous Broadening

Among the above line broadening mechanisms, the pressure, rotational diffusion, and lifetime broadenings are all of the **homogeneous** variety. This means that each molecule in the sample is affected in exactly the same manner by the broadening process. For example, one does not find some molecules with short lifetimes and others with long lifetimes, in the Heisenberg case; the entire ensemble of molecules is characterized by a single lifetime.

In contrast, Doppler broadening is **inhomogeneous** in nature because each molecule experiences a broadening that is characteristic of its particular nature (velocity  $v_z$  in this case). That is, the fast molecules have their lines broadened more than do the slower molecules. Another important example of inhomogeneous broadening is provided by so-called **site broadening**. Molecules imbedded in a liquid, solid, or glass do not, at the instant of photon absorption, all experience exactly the same interactions with their surroundings. The distribution of instantaneous "solvation" environments may be rather "narrow" (e.g., in a highly ordered solid matrix) or quite "broad" (e.g., in a liquid at high temperature). Different environments produce different energy level splittings  $\Delta E_{i,f} = E_{f,i} + \epsilon_{i,f}$  (because the initial and final states are "solvated" differently by the surroundings) and thus different frequencies at which photon absorption can occur. The distribution of energy level splittings causes the sample to absorb at a range of frequencies as illustrated in the figure below where homogeneous and inhomogeneous line shapes are compared.



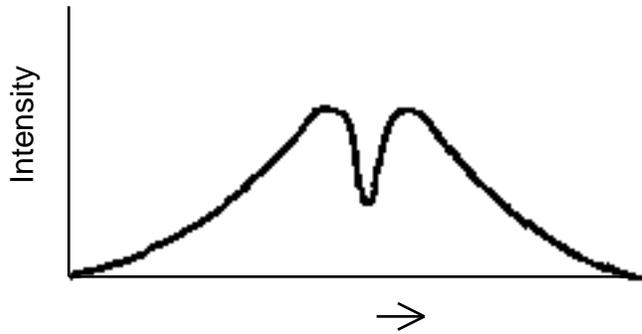
Homogeneous (a) and inhomogeneous (b) band shapes having inhomogeneous width  $\Delta\nu_{INH}$  and homogeneous width  $\Delta\nu_H$ .

The spectral line shape function  $I(\nu)$  is further broadened when site inhomogeneity is present and significant. These effects can be modeled by convolving the kind of  $I^0(\nu; E)$  function that results from Doppler, lifetime, rotational diffusion, and pressure broadening with a Gaussian distribution  $P(E)$  that describes the inhomogeneous distribution of energy level splittings:

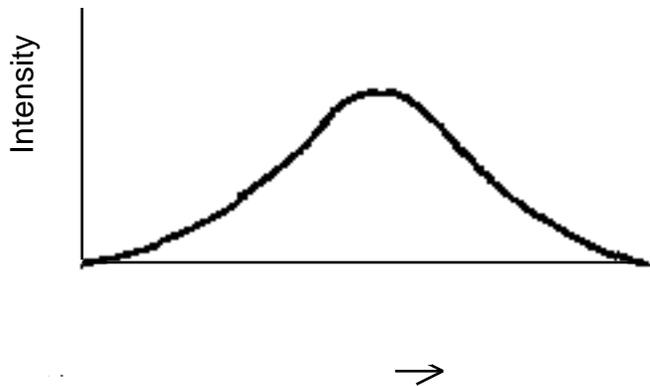
$$I(\nu) = \int I^0(\nu; E) P(E) dE.$$

Here  $I^0(\nu; E)$  is a line shape function such as those described earlier each of which contains a set of frequencies (e.g.,  $\nu = \nu_{i,v} + E_{i,f}/h \pm J = \nu_0 + E/h$ ) at which absorption or emission occurs.

A common experimental test for inhomogeneous broadening involves **hole burning**. In such experiments, an intense light source (often a laser) is tuned to a frequency  $\nu_{burn}$  that lies within the spectral line being probed for inhomogeneous broadening. Then, a second tunable light source is used to scan through the profile of the spectral line, and, for example, an absorption spectrum is recorded. Given an absorption profile as shown below in the absence of the intense burning light source:



one expects to see a profile such as that shown below:



if inhomogeneous broadening is operative.

The interpretation of the change in the absorption profile caused by the bright light source proceeds as follows:

- (i) In the ensemble of molecules contained in the sample, some molecules will absorb at or near the frequency of the bright light source  $\nu_{\text{burn}}$ ; other molecules (those whose environments do not produce energy level splittings that match  $\nu_{\text{burn}}$ ) will not absorb at this frequency.
- (ii) Those molecules that do absorb at  $\nu_{\text{burn}}$  will have their transition saturated by the intense light source, thereby rendering this frequency region of the line profile transparent to further absorption.
- (iii) When the "probe" light source is scanned over the line profile, it will induce absorptions for those molecules whose local environments did not allow them to be saturated by the  $\nu_{\text{burn}}$  light. The absorption profile recorded by this probe light source's detector thus will match that of the original line profile, until

(iv) the probe light source's frequency matches  $\nu_{\text{burn}}$ , upon which no absorption of the probe source's photons will be recorded because molecules that absorb in this frequency regime have had their transition saturated.

(v) Hence, a "**hole**" will appear in the spectrum recorded by the probe light source's detector in the region of  $\nu_{\text{burn}}$ .

Unfortunately, the technique of hole burning does not provide a fully reliable method for identifying inhomogeneously broadened lines. If a hole is observed in such a burning experiment, this provides ample evidence, but if one is not seen, the result is not definitive. In the latter case, the transition may not be strong enough (i.e., may not have a large enough "rate of photon absorption" ) for the intense light source to saturate the transition to the extent needed to form a hole.

*This completes our introduction to the subject of molecular spectroscopy. More advanced treatments of many of the subjects treated here as well as many aspects of modern experimental spectroscopy can be found in the text by Zare on angular momentum as well as in Steinfeld's text Molecules and Radiation, 2<sup>nd</sup> Edition, by J. I. Steinfeld, MIT Press (1985).*

## Chapter 18

*The single Slater determinant wavefunction (properly spin and symmetry adapted) is the starting point of the most common mean field potential. It is also the origin of the molecular orbital concept.*

### I. Optimization of the Energy for a Multiconfiguration Wavefunction

#### A. The Energy Expression

The most straightforward way to introduce the concept of optimal molecular orbitals is to consider a trial wavefunction of the form which was introduced earlier in Chapter 9.II. The expectation value of the Hamiltonian for a wavefunction of the multiconfigurational form

$$= \sum_I C_I \Phi_I,$$

where  $\Phi_I$  is a space- and spin-adapted CSF which consists of determinantal wavefunctions  $|\Phi_I\rangle = |\Phi_{I1} \Phi_{I2} \Phi_{I3} \dots \Phi_{IN}|$ , can be written as:

$$E = \sum_{I,J=1}^M C_I C_J \langle \Phi_I | H | \Phi_J \rangle.$$

The spin- and space-symmetry of the  $\Phi_I$  determine the symmetry of the state whose energy is to be optimized.

In this form, it is clear that  $E$  is a quadratic function of the CI amplitudes  $C_J$ ; it is a quartic functional of the spin-orbitals because the Slater-Condon rules express each  $\langle \Phi_I | H | \Phi_J \rangle$  CI matrix element in terms of one- and two-electron integrals  $\langle i | f | j \rangle$  and  $\langle i j | g | k l \rangle$  over these spin-orbitals.

#### B. Application of the Variational Method

The variational method can be used to optimize the above expectation value expression for the electronic energy (i.e., to make the functional stationary) as a function of the CI coefficients  $C_J$  and the LCAO-MO coefficients  $\{C_{\mu,i}\}$  that characterize the spin-orbitals. However, in doing so the set of  $\{C_{\mu,i}\}$  can not be treated as entirely independent variables. The fact that the spin-orbitals  $\{\phi_i\}$  are assumed to be orthonormal imposes a set of constraints on the  $\{C_{\mu,i}\}$ :

$$\langle i | j \rangle = \delta_{ij} = \mu, \quad C_{\mu,i}^* \langle \mu | \rangle > C_{\mu,j}$$

These constraints can be enforced within the variational optimization of the energy function mentioned above by introducing a set of Lagrange multipliers  $\{ \lambda_{ij} \}$ , one for each constraint condition, and subsequently differentiating

$$E - \sum_{i,j} \lambda_{ij} [ \delta_{ij} - \mu, \quad C_{\mu,i}^* \langle \mu | \rangle > C_{\mu,j} ]$$

with respect to each of the  $C_{\mu,i}$  variables.

### C. The Fock and Secular Equations

Upon doing so, the following set of equations is obtained (early references to the derivation of such equations include A. C. Wahl, J. Chem. Phys. 41, 2600 (1964) and F. Grein and T. C. Chang, Chem. Phys. Lett. 12, 44 (1971); a more recent overview is presented in R. Shepard, p 63, in Adv. in Chem. Phys. LXIX, K. P. Lawley, Ed., Wiley-Interscience, New York (1987); the subject is also treated in the textbook Second Quantization Based Methods in Quantum Chemistry, P. Jørgensen and J. Simons, Academic Press, New York (1981)) :

$$\sum_{J=1, M} H_{I,J} C_J = E C_I, \quad I = 1, 2, \dots M, \text{ and}$$

$$F_{ij} = \lambda_{ij} \delta_{ij},$$

where the  $\lambda_{ij}$  are Lagrange multipliers.

The first set of equations govern the  $\{C_J\}$  amplitudes and are called the CI- secular equations. The second set determine the LCAO-MO coefficients of the spin-orbitals  $\{ \phi_j \}$  and are called the Fock equations. The Fock operator  $F$  is given in terms of the one- and two-electron operators in  $H$  itself as well as the so-called one- and two-electron density matrices  $\rho_{ij}$  and  $\rho_{i,j,k,l}$  which are defined below. These density matrices reflect the averaged occupancies of the various spin orbitals in the CSFs of  $\Psi$ . The resultant expression for  $F$  is:

$$F_{ij} = \sum_j \lambda_{ij} h_{ij} + \sum_{j,k,l} \lambda_{j,k,l} \rho_{i,j,k,l} J_{j,l} \delta_{ik}$$

where  $h$  is the one-electron component of the Hamiltonian (i.e., the kinetic energy operator and the sum of coulombic attractions to the nuclei). The operator  $J_{j,l}$  is defined by:

$$J_{j,l} \chi_k(\mathbf{r}) = \int \chi_j^*(\mathbf{r}') \chi_l(\mathbf{r}') / |\mathbf{r} - \mathbf{r}'| d\mathbf{r}' \chi_k(\mathbf{r}),$$

where the integration denoted  $d\mathbf{r}'$  is over the spatial and spin coordinates. The so-called spin integration simply means that the  $\alpha$  or  $\beta$  spin function associated with  $\chi_l$  must be the same as the  $\alpha$  or  $\beta$  spin function associated with  $\chi_j$  or the integral will vanish. This is a consequence of the orthonormality conditions  $\langle \chi_i | \chi_j \rangle = \langle \chi_i | \chi_j \rangle = 1$ ,  $\langle \chi_i | \chi_j \rangle = \langle \chi_j | \chi_i \rangle = 0$ .

#### D. One- and Two- Electron Density Matrices

The density matrices introduced above can most straightforwardly be expressed in terms of the CI amplitudes and the nature of the orbital occupancies in the CSFs of  $\Psi$  as follows:

1.  $\rho_{i,i}$  is the sum over all CSFs, in which  $\chi_i$  is occupied, of the square of the  $C_I$  coefficient of that CSF:

$$\rho_{i,i} = \sum_I (\text{with } \chi_i \text{ occupied}) C_I^2.$$

2.  $\rho_{i,j}$  is the sum over pairs of CSFs which differ by a single spin-orbital occupancy (i.e., one having  $\chi_i$  occupied where the other has  $\chi_j$  occupied after the two are placed into maximal coincidence—the sign factor (sign) arising from bringing the two to maximal coincidence is attached to the final density matrix element):

$$\rho_{i,j} = \sum_{I,J} (\text{sign}) (\text{with } \chi_i \text{ occupied in } I \text{ where } \chi_j \text{ is in } J) C_I C_J.$$

The two-electron density matrix elements are given in similar fashion:

3.  $\rho_{i,j,i,j} = \sum_I (\text{with both } \chi_i \text{ and } \chi_j \text{ occupied}) C_I^2$ ;

4.  $\rho_{i,j,j,i} = - \sum_I (\text{with both } \chi_i \text{ and } \chi_j \text{ occupied}) C_I^2 = - \rho_{i,j,i,j}$

(it can be shown, in general that  $i, j, k, l$  is odd under exchange of  $i$  and  $j$ , odd under exchange of  $k$  and  $l$  and even under  $(i, j) \leftrightarrow (k, l)$  exchange; this implies that  $i, j, k, l$  vanishes if  $i = j$  or  $k = l$ );

$$5. \quad i, j, k, j = \sum_{I, J} (\text{sign}) (\text{with } j \text{ in both } I \text{ and } J \\ \text{and } i \text{ in } I \text{ where } k \text{ is in } J) C_I C_J \\ = j, i, j, k = - i, j, j, k = - j, i, k, j;$$

$$6. \quad i, j, k, l = \sum_{I, J} (\text{sign}) (\text{with } i \text{ in } I \text{ where } k \text{ is in } J \text{ and } j \text{ in } I \text{ where } l \text{ is in } J) C_I \\ C_J \\ = j, i, l, k = - j, i, k, l = - i, j, l, k = j, i, l, k.$$

These density matrices are themselves quadratic functions of the CI coefficients and they reflect all of the permutational symmetry of the determinantal functions used in constructing  $\rho$ ; they are a compact representation of all of the Slater-Condon rules as applied to the particular CSFs which appear in  $\rho$ . They contain all information about the spin-orbital occupancy of the CSFs in  $\rho$ . The one- and two- electron integrals  $\langle i | f | j \rangle$  and  $\langle i j | g | k l \rangle$  contain all of the information about the magnitudes of the kinetic and Coulombic interaction energies.

## II. The Single-Determinant Wavefunction

The simplest trial function of the form given above is the single Slater determinant function:

$$= | 1 \ 2 \ 3 \ \dots \ N |.$$

For such a function, the CI part of the energy minimization is absent (the classic papers in which the SCF equations for closed- and open-shell systems are treated are C. C. J. Roothaan, Rev. Mod. Phys. 23, 69 (1951); 32, 179 (1960)) and the density matrices simplify greatly because only one spin-orbital occupancy is operative. In this case, the orbital optimization conditions reduce to:

$$F_i = \sum_j i_j j_j,$$

where the so-called Fock operator  $F$  is given by

$$F \phi_i = h \phi_i + \sum_{j(\text{occupied})} [J_j - K_j] \phi_i.$$

The coulomb ( $J_j$ ) and exchange ( $K_j$ ) operators are defined by the relations:

$$J_j \phi_i = \int \phi_j^*(r') \phi_j(r') \frac{1}{|r-r'|} d^3r' \phi_i(r), \text{ and}$$

$$K_j \phi_i = \int \phi_j^*(r') \phi_i(r') \frac{1}{|r-r'|} d^3r' \phi_j(r).$$

Again, the integration implies integration over the spin variables associated with the  $\phi_j$  (and, for the exchange operator,  $\phi_i$ ), as a result of which the exchange integral vanishes unless the spin function of  $\phi_j$  is the same as that of  $\phi_i$ ; the coulomb integral is non-vanishing no matter what the spin functions of  $\phi_j$  and  $\phi_i$ .

The sum over coulomb and exchange interactions in the Fock operator runs only over those spin-orbitals that are occupied in the trial  $\psi$ . Because a unitary transformation among the orbitals that appear in  $\psi$  leaves the determinant unchanged (this is a property of determinants-  $\det(UA) = \det(U) \det(A) = 1 \det(A)$ , if  $U$  is a unitary matrix), it is possible to choose such a unitary transformation to make the  $i,j$  matrix diagonal. Upon so doing, one is left with the so-called canonical Hartree-Fock equations:

$$F \phi_i = \epsilon_i \phi_i,$$

where  $\epsilon_i$  is the diagonal value of the  $i,j$  matrix after the unitary transformation has been applied; that is,  $\epsilon_i$  is an eigenvalue of the  $i,j$  matrix. These equations are of the eigenvalue-eigenfunction form with the Fock operator playing the role of an effective one-electron Hamiltonian and the  $\phi_i$  playing the role of the one-electron eigenfunctions.

It should be noted that the Hartree-Fock equations  $F \phi_i = \epsilon_i \phi_i$  possess solutions for the spin-orbitals which appear in  $\psi$  (the so-called occupied spin-orbitals) as well as for orbitals which are not occupied in  $\psi$  (the so-called virtual spin-orbitals). In fact, the  $F$  operator is hermitian, so it possesses a complete set of orthonormal eigenfunctions; only those which appear in  $\psi$  appear in the coulomb and exchange potentials of the Fock operator. The physical meaning of the occupied and virtual orbitals will be clarified later in this Chapter (Section VII.A)

### III. The Unrestricted Hartree-Fock Spin Impurity Problem

As formulated above in terms of spin-orbitals, the Hartree-Fock (HF) equations yield orbitals that do not guarantee that  $\Psi$  possesses proper spin symmetry. To illustrate the point, consider the form of the equations for an open-shell system such as the Lithium atom Li. If  $1s_\alpha$ ,  $1s_\beta$ , and  $2s$  spin-orbitals are chosen to appear in the trial function  $\Psi$ , then the Fock operator will contain the following terms:

$$F = h + J_{1s_\alpha} + J_{1s_\beta} + J_{2s} - [K_{1s_\alpha} + K_{1s_\beta} + K_{2s}].$$

Acting on an  $1s_\alpha$  spin-orbital  $\phi_k$  with F and carrying out the spin integrations, one obtains

$$F \phi_k = h \phi_k + (2J_{1s_\alpha} + J_{2s}) \phi_k - (K_{1s_\alpha} + K_{2s}) \phi_k.$$

In contrast, when acting on a  $1s_\beta$  spin-orbital, one obtains

$$F \phi_k = h \phi_k + (2J_{1s_\beta} + J_{2s}) \phi_k - (K_{1s_\beta}) \phi_k.$$

Spin-orbitals of  $1s_\alpha$  and  $1s_\beta$  type do not experience the same exchange potential in this model, which is clearly due to the fact that  $\Psi$  contains two  $1s$  spin-orbitals and only one  $2s$  spin-orbital.

One consequence of the spin-polarized nature of the effective potential in F is that the optimal  $1s_\alpha$  and  $1s_\beta$  spin-orbitals, which are themselves solutions of  $F \phi_i = \epsilon_i \phi_i$ , do not have identical orbital energies (i.e.,  $\epsilon_{1s_\alpha} \neq \epsilon_{1s_\beta}$ ) and are not spatially identical to one another (i.e.,  $1s_\alpha$  and  $1s_\beta$  do not have identical LCAO-MO expansion coefficients). This resultant spin polarization of the orbitals in  $\Psi$  gives rise to spin impurities in  $\Psi$ . That is, the determinant  $|1s_\alpha 1s_\beta 2s|$  is not a pure doublet spin eigenfunction although it is an  $S_z$  eigenfunction with  $M_s = 1/2$ ; it contains both  $S = 1/2$  and  $S = 3/2$  components. If the  $1s_\alpha$  and  $1s_\beta$  spin-orbitals were spatially identical, then  $|1s_\alpha 1s_\beta 2s|$  would be a pure spin eigenfunction with  $S = 1/2$ .

The above single-determinant wavefunction is commonly referred to as being of the **unrestricted Hartree-Fock** (UHF) type because no restrictions are placed on the spatial nature of the orbitals which appear in  $\Psi$ . In general, UHF wavefunctions are not of pure spin symmetry for any open-shell system. Such a UHF treatment forms the starting point of early versions of the widely used and highly successful Gaussian 70 through Gaussian-

8X series of electronic structure computer codes which derive from J. A. Pople and co-workers (see, for example, M. J. Frisch, J. S. Binkley, H. B. Schlegel, K Raghavachari, C. F. Melius, R. L. Martin, J. J. P. Stewart, F. W. Bobrowicz, C. M. Rohling, L. R. Kahn, D. J. Defrees, R. Seeger, R. A. Whitehead, D. J. Fox, E. M. Fleuder, and J. A. Pople, Gaussian 86, Carnegie-Mellon Quantum Chemistry Publishing Unit, Pittsburgh, PA (1984)).

The inherent spin-impurity problem is sometimes 'fixed' by using the orbitals which are obtained in the UHF calculation to subsequently form a properly spin-adapted wavefunction. For the above Li atom example, this amounts to forming a new wavefunction (after the orbitals are obtained via the UHF process) using the techniques detailed in Section 3 and Appendix G:

$$= 1/2 [ |1s \ 1s' \ 2s \ | - |1s \ 1s' \ 2s \ | ] .$$

This wavefunction is a pure  $S = 1/2$  state. This prescription for avoiding spin contamination (i.e., carrying out the UHF calculation and then forming a new spin-pure ) is referred to as spin-projection.

It is, of course, possible to first form the above spin-pure as a trial wavefunction and to then determine the orbitals  $1s$   $1s'$  and  $2s$  which minimize its energy; in so doing, one is dealing with a spin-pure function from the start. The problem with carrying out this process, which is referred to as a spin-adapted Hartree-Fock calculation, is that the resultant  $1s$  and  $1s'$  orbitals still do not have identical spatial attributes. Having a set of orbitals ( $1s$ ,  $1s'$ ,  $2s$ , and the virtual orbitals) that form a non-orthogonal set ( $1s$  and  $1s'$  are neither identical nor orthogonal) makes it difficult to progress beyond the single-configuration wavefunction as one often wishes to do. That is, it is difficult to use a spin-adapted wavefunction as a starting point for a correlated-level treatment of electronic motions.

Before addressing head-on the problem of how to best treat orbital optimization for open-shell species, it is useful to examine how the HF equations are solved in practice in terms of the LCAO-MO process.

#### IV. The LCAO-MO Expansion

The HF equations  $F_i = \epsilon_i \psi_i$  comprise a set of integro-differential equations; their differential nature arises from the kinetic energy operator in  $h$ , and the coulomb and exchange operators provide their integral nature. The solutions of these equations must be

achieved iteratively because the  $J_i$  and  $K_i$  operators in  $F$  depend on the orbitals  $\psi_i$  which are to be solved for. Typical iterative schemes begin with a 'guess' for those  $\psi_i$  which appear in  $\psi$ , which then allows  $F$  to be formed. Solutions to  $F \psi_i = \epsilon_i \psi_i$  are then found, and those  $\psi_i$  which possess the space and spin symmetry of the occupied orbitals of  $\psi$  and which have the proper energies and nodal character are used to generate a new  $F$  operator (i.e., new  $J_i$  and  $K_i$  operators). The new  $F$  operator then gives new  $\psi_i$  and  $\epsilon_i$  via solution of the new  $F \psi_i = \epsilon_i \psi_i$  equations. This iterative process is continued until the  $\psi_i$  and  $\epsilon_i$  do not vary significantly from one iteration to the next, at which time one says that the process has converged. This iterative procedure is referred to as the Hartree-Fock self-consistent field (SCF) procedure because iteration eventually leads to coulomb and exchange potential fields that are consistent from iteration to iteration.

In practice, solution of  $F \psi_i = \epsilon_i \psi_i$  as an integro-differential equation can be carried out only for atoms (C. Froese-Fischer, *Comp. Phys. Commun.* **1**, 152 (1970)) and linear molecules (P. A. Christiansen and E. A. McCullough, *J. Chem. Phys.* **67**, 1877 (1977)) for which the angular parts of the  $\psi_i$  can be exactly separated from the radial because of the axial- or full- rotation group symmetry (e.g.,  $\psi_i = Y_{l,m} R_{n,l}(r)$  for an atom and  $\psi_i = \exp(im\phi) R_{n,l,m}(r, \theta)$  for a linear molecule). In such special cases,  $F \psi_i = \epsilon_i \psi_i$  gives rise to a set of coupled equations for the  $R_{n,l}(r)$  or  $R_{n,l,m}(r, \theta)$  which can and have been solved. However, for non-linear molecules, the HF equations have not yet been solved in such a manner because of the three-dimensional nature of the  $\psi_i$  and of the potential terms in  $F$ .

In the most commonly employed procedures used to solve the HF equations for non-linear molecules, the  $\psi_i$  are expanded in a basis of functions  $\chi_\mu$  according to the LCAO-MO procedure:

$$\psi_i = \sum_{\mu} C_{\mu,i} \chi_{\mu}$$

Doing so then reduces  $F \psi_i = \epsilon_i \psi_i$  to a matrix eigenvalue-type equation of the form:

$$F_{\mu\nu} C_{\nu,i} = \epsilon_i S_{\mu\nu} C_{\nu,i}$$

where  $S_{\mu\nu} = \langle \chi_{\mu} | \chi_{\nu} \rangle$  is the overlap matrix among the atomic orbitals (aos) and

$$F_{\mu\nu} = \langle \chi_{\mu} | h | \chi_{\nu} \rangle + \sum_j [ \langle \chi_{\mu} | g_j | \chi_{\nu} \rangle - \delta_{\mu\nu} \sum_j \langle \chi_{\mu} | g_j | \chi_{\mu} \rangle ]$$

is the matrix representation of the Fock operator in the ao basis. The coulomb and exchange- density matrix elements in the ao basis are:

$$C_{\mu} = \sum_{i(\text{occupied})} C_{\mu,i} C_{\mu,i}, \text{ and}$$

$$C_{\mu}^{\text{ex}} = \sum_{i(\text{occ., and same spin})} C_{\mu,i} C_{\mu,i},$$

where the sum in  $C_{\mu}^{\text{ex}}$  runs over those occupied spin-orbitals whose  $m_s$  value is equal to that for which the Fock matrix is being formed (for a closed-shell species,  $C_{\mu}^{\text{ex}} = 1/2$ ).

It should be noted that by moving to a matrix problem, one does not remove the need for an iterative solution; the  $F_{\mu}$  matrix elements depend on the  $C_{\mu,i}$  LCAO-MO coefficients which are, in turn, solutions of the so-called Roothaan matrix Hartree-Fock equations-  $F_{\mu} C_{\mu,i} = \epsilon_i S_{\mu} C_{\mu,i}$ . One should also note that, just as  $F_{\mu}$  possesses a complete set of eigenfunctions, the matrix  $F_{\mu}$ , whose dimension  $M$  is equal to the number of atomic basis orbitals used in the LCAO-MO expansion, has  $M$  eigenvalues  $\epsilon_i$  and  $M$  eigenvectors whose elements are the  $C_{\mu,i}$ . Thus, there are occupied and virtual molecular orbitals (mos) each of which is described in the LCAO-MO form with  $C_{\mu,i}$  coefficients obtained via solution of

$$F_{\mu} C_{\mu,i} = \epsilon_i S_{\mu} C_{\mu,i}.$$

## V. Atomic Orbital Basis Sets

### A. STOs and GTOs

The basis orbitals commonly used in the LCAO-MO-SCF process fall into two classes:

#### 1. Slater-type orbitals

$$\psi_{n,l,m}(r, \theta, \phi) = N_{n,l,m} Y_{l,m}(\theta, \phi) r^{n-1} e^{-r},$$

which are characterized by quantum numbers  $n$ ,  $l$ , and  $m$  and exponents (which characterize the 'size' of the basis function). The symbol  $N_{n,l,m}$  denotes the normalization constant.

## 2. Cartesian Gaussian-type orbitals

$$\chi_{a,b,c}(r, \theta, \phi) = N'_{a,b,c} x^a y^b z^c \exp(-\alpha r^2),$$

characterized by quantum numbers  $a$ ,  $b$ , and  $c$  which detail the angular shape and direction of the orbital and exponents  $\alpha$  which govern the radial 'size' of the basis function. For example, orbitals with  $a$ ,  $b$ , and  $c$  values of 1,0,0 or 0,1,0 or 0,0,1 are  $p_x$ ,  $p_y$ , and  $p_z$  orbitals; those with  $a,b,c$  values of 2,0,0 or 0,2,0 or 0,0,2 and 1,1,0 or 0,1,1 or 1,0,1 span the space of five  $d$  orbitals and one  $s$  orbital (the sum of the 2,0,0 and 0,2,0 and 0,0,2 orbitals is an  $s$  orbital because  $x^2 + y^2 + z^2 = r^2$  is independent of  $\theta$  and  $\phi$ ).

For both types of orbitals, the coordinates  $r$ ,  $\theta$ , and  $\phi$  refer to the position of the electron relative to a set of axes attached to the center on which the basis orbital is located. Although Slater-type orbitals (STOs) are preferred on fundamental grounds (e.g., as demonstrated in Appendices A and B, the hydrogen atom orbitals are of this form and the exact solution of the many-electron Schrödinger equation can be shown to be of this form (in each of its coordinates) near the nuclear centers), STOs are used primarily for atomic and linear-molecule calculations because the multi-center integrals  $\langle \chi_a | g | \chi_c \rangle$  (each basis orbital can be on a separate atomic center) which arise in polyatomic-molecule calculations can not efficiently be performed when STOs are employed. In contrast, such integrals can routinely be done when Gaussian-type orbitals (GTOs) are used. This fundamental advantage of GTOs has led to the dominance of these functions in molecular quantum chemistry.

To understand why integrals over GTOs can be carried out when analogous STO-based integrals are much more difficult, one must only consider the orbital products  $(\chi_a(r_1) \chi_c(r_2))$  which arise in such integrals. For orbitals of the GTO form, such products involve  $\exp(-\alpha(\mathbf{r}-\mathbf{R}_a)^2) \exp(-\beta(\mathbf{r}-\mathbf{R}_c)^2)$ . By completing the square in the exponent, this product can be rewritten as follows:

$$\begin{aligned} & \exp(-\alpha(\mathbf{r}-\mathbf{R}_a)^2) \exp(-\beta(\mathbf{r}-\mathbf{R}_c)^2) \\ &= \exp(-(\alpha+\beta)(\mathbf{r}-\mathbf{R}')^2) \exp(-\beta'(\mathbf{R}_a-\mathbf{R}_c)^2), \end{aligned}$$

where

$$\mathbf{R}' = [\alpha \mathbf{R}_a + \beta \mathbf{R}_c] / (\alpha + \beta) \text{ and}$$

$$\psi = \frac{1}{\sqrt{2}} (\psi_a + \psi_c).$$

Thus, the product of two GTOs on different centers is equal to a single other GTO at a center  $\mathbf{R}'$  between the two original centers. As a result, even a four-center two-electron integral over GTOs can be written as, at most, a two-center two-electron integral; it turns out that this reduction in centers is enough to allow all such integrals to be carried out. A similar reduction does not arise for STOs because the product of two STOs can not be rewritten as a new STO at a new center.

To overcome the primary weakness of GTO functions, that they have incorrect behavior near the nuclear centers (i.e., their radial derivatives vanish at the nucleus whereas the derivatives of STOs are non-zero), it is common to combine two, three, or more GTOs, with combination coefficients which are fixed and not treated as LCAO-MO parameters, into new functions called contracted GTOs or CGTOs. Typically, a series of tight, medium, and loose GTOs (i.e., GTOs with large, medium, and small  $\alpha$  values, respectively) are multiplied by so-called contraction coefficients and summed to produce a CGTO which appears to possess the proper 'cusp' (i.e., non-zero slope) at the nuclear center (although even such a combination can not because each GTO has zero slope at the nucleus).

## B. Basis Set Libraries

Much effort has been devoted to developing sets of STO or GTO basis orbitals for main-group elements and the lighter transition metals. This ongoing effort is aimed at providing standard basis set libraries which:

1. Yield reasonable chemical accuracy in the resultant wavefunctions and energies.
2. Are cost effective in that their use in practical calculations is feasible.
3. Are relatively transferrable in the sense that the basis for a given atom is flexible enough to be used for that atom in a variety of bonding environments (where the atom's hybridization and local polarity may vary).

## C. The Fundamental Core and Valence Basis

In constructing an atomic orbital basis to use in a particular calculation, one must choose from among several classes of functions. First, the size and nature of the primary

core and valence basis must be specified. Within this category, the following choices are common:

1. A minimal basis in which the number of STO or CGTO orbitals is equal to the number of core and valence atomic orbitals in the atom.
2. A double-zeta (DZ) basis in which twice as many STOs or CGTOs are used as there are core and valence atomic orbitals. The use of more basis functions is motivated by a desire to provide additional variational flexibility to the LCAO-MO process. This flexibility allows the LCAO-MO process to generate molecular orbitals of variable diffuseness as the local electronegativity of the atom varies. Typically, double-zeta bases include pairs of functions with one member of each pair having a smaller exponent (or  $\alpha$  value) than in the minimal basis and the other member having a larger exponent.
3. A triple-zeta (TZ) basis in which three times as many STOs or CGTOs are used as the number of core and valence atomic orbitals.
4. Dunning has developed CGTO bases which range from approximately DZ to substantially beyond TZ quality (T. H. Dunning, *J. Chem. Phys.* **53**, 2823 (1970); T. H. Dunning and P. J. Hay in Methods of Electronic Structure Theory, H. F. Schaefer, III Ed., Plenum Press, New York (1977)). These bases involve contractions of primitive GTO bases which Huzinaga had earlier optimized (S. Huzinaga, *J. Chem. Phys.* **42**, 1293 (1965)) for use as uncontracted functions (i.e., for which Huzinaga varied the  $\alpha$  values to minimize the energies of several electronic states of the corresponding atom). These Dunning bases are commonly denoted, for example, as follows for first-row atoms: (10s,6p/5s,4p), which means that 10 s-type primitive GTOs have been contracted to produce 5 separate s-type CGTOs and that 6 primitive p-type GTOs were contracted to generate 4 separate p-type CGTOs. More recent basis sets from the Dunning group are given in T. Dunning, *J. Chem. Phys.* **90**, 1007 (1990).
5. Even-tempered basis sets (M. W. Schmidt and K. Ruedenberg, *J. Chem. Phys.* **71**, 3961 (1979)) consist of GTOs in which the orbital exponents  $\alpha_k$  belonging to series of orbitals consist of geometrical progressions:  $\alpha_k = a^{-k}$ , where  $a$  and  $\alpha_0$  characterize the particular set of GTOs.
6. STO-3G bases were employed some years ago (W. J. Hehre, R. F. Stewart, and J. A. Pople, *J. Chem. Phys.* **51**, 2657 (1969)) but are less popular recently. These bases are constructed by least squares fitting GTOs to STOs which have been optimized for various electronic states of the atom. When three GTOs are employed to fit each STO, a STO-3G basis is formed.
7. 4-31G, 5-31G, and 6-31G bases (R. Ditchfield, W. J. Hehre, and J. A. Pople, *J. Chem. Phys.* **54**, 724 (1971); W. J. Hehre, R. Ditchfield, and J. A. Pople, *J. Chem.*

Phys. 56, 2257 (1972); P. C. Hariharan and J. A. Pople, Theoret. Chim. Acta. (Berl.) 28, 213 (1973); R. Krishnan, J. S. Binkley, R. Seeger, and J. A. Pople, J. Chem. Phys. 72, 650 (1980)) employ a single CGTO of contraction length 4, 5, or 6 to describe the core orbital. The valence space is described at the DZ level with the first CGTO constructed from 3 primitive GTOs and the second CGTO built from a single primitive GTO.

The values of the orbital exponents (  $s$  or  $s$  ) and the GTO-to-CGTO contraction coefficients needed to implement a particular basis of the kind described above have been tabulated in several journal articles and in computer data bases (in particular, in the data base contained in the book Handbook of Gaussian Basis Sets: A Compendium for Ab initio Molecular Orbital Calculations, R. Poirer, R. Kari, and I. G. Csizmadia, Elsevier Science Publishing Co., Inc., New York, New York (1985)).

Several other sources of basis sets for particular atoms are listed in the Table shown below (here JCP and JACS are abbreviations for the Journal of Chemical Physics and the Journal of The American Chemical Society, respectively).

<u>Literature Reference</u>	<u>Basis Type</u>	<u>Atoms</u>
Hehre, W.J.; Stewart, R.F.; Pople, J.A. JCP <u>51</u> , 2657 (1969). Hehre, W.J.; Ditchfield, R.; Stewart, R.F.; Pople, J.A. JCP <u>52</u> , 2769 (1970).	STO-3G	H-Ar
Binkley, J.S.; Pople, J.A.; Hehre, W.J. JACS <u>102</u> , 939 (1980).	3-21G	H-Ne
Gordon, M.S.; Binkley, J.S.; Pople, J.A.; Pietro, W.J.; Hehre, W.J. JACS <u>104</u> , 2797 (1982).	3-21G	Na-Ar
Dobbs, K.D.; Hehre, W.J. J. Comput. Chem. <u>7</u> , 359 (1986).	3-21G	K,Ca,Ga
Dobbs, K.D.; Hehre, W.J. J. Comput. Chem. <u>8</u> , 880 (1987).	3-21G	Sc-Zn
Ditchfield, R.; Hehre, W.J.; Pople, J.A. JCP <u>54</u> , 724 (1971).	6-31G	H

Dill, J.D.; Pople, J.A. JCP <u>62</u> , 2921 (1975).	6-31G	Li,B
Binkley, J.S.; Pople, J.A. JCP <u>66</u> , 879 (1977).	6-31G	Be
Hehre, W.J.; Ditchfield, R.; Pople, J.A. JCP <u>56</u> , 2257 (1972).	6-31G	C-F
Francl, M.M.; Pietro, W.J.; Hehre, W.J.; Binkley, J.S.; Gordon, M.S.; DeFrees, D.J.; Pople, J.A. JCP <u>77</u> , 3654 (1982).	6-31G	Na-Ar
Dunning, T. JCP <u>53</u> , 2823 (1970).	(4s/2s)	H
	(4s/3s)	H
	(9s5p/3s2p)	B-F
	(9s5p/4s2p)	B-F
	(9s5p/5s3p)	B-F
Dunning, T. JCP <u>55</u> , 716 (1971).	(5s/3s)	H
	(10s/4s)	Li
	(10s/5s)	Be
	(10s6p/5s3p)	B-Ne
	(10s6p/5s4p)	B-Ne
Krishnan, R.; Binkley, J.S.; Seeger, R.; Pople, J.A. JCP <u>72</u> , 650 (1980).	6-311G	H-Ne
Dunning, unpublished VDZ.	(4s/2s)	H
	(9s5p/3s2)	Li,Be,C-Ne
	(12s8p/4s3p)	Na-Ar
Dunning, unpublished VTZ.	(5s/3s)	H
	(6s/3s)	H

(12s6p/4s3p) Li,Be,C-Ne  
(17s10p/5s4p) Mg-Ar

Dunning, unpublished VQZ. (7s/4s) H  
(8s/4s) H  
(16s7p/5s4p) B-Ne

Dunning, T. JCP 90, 1007 (1989). (4s1p/2s1p) H  
(pVDZ,pVTZ,pVQZ correlation-consistent) (5s2p1d/3s2p1d) H  
(6s3p1d1f/4s3p2d1f) H  
(9s4p1d/3s2p1d) B-Ne  
(10s5p2d1f/4s3p2d1f) B-Ne  
(12s6p3d2f1g/5s4p3d2f1g) B-Ne

Huzinaga, S.; Klobukowski, M.; Tatewaki, H. (14s/2s) Li,Be  
Can. J. Chem. 63, 1812 (1985). (14s9p/2s1p) B-Ne  
(16s9p/3s1p) Na-Mg  
(16s11p/3s2p) Al-Ar

Huzinaga, S.; Klobukowski, M. (14s10p/2s1p) B-Ne  
THEOCHEM. 44, 1 (1988). (17s10p/3s1p) Na-Mg  
(17s13p/3s2p) Al-Ar  
(20s13p/4s2p) K-Ca  
(20s13p10d/4s2p1d) Sc-Zn  
(20s14p9d/4s3d1d) Ga

McLean, A.D.; Chandler, G.S. (12s8p/4s2p) Na-Ar, P<sup>-</sup>, S<sup>-</sup>, Cl<sup>-</sup>  
JCP 72, 5639 (1980). (12s8p/5s3p) Na-Ar, P<sup>-</sup>, S<sup>-</sup>, Cl<sup>-</sup>  
(12s8p/6s4p) Na-Ar, P<sup>-</sup>, S<sup>-</sup>, Cl<sup>-</sup>  
(12s9p/6s4p) Na-Ar, P<sup>-</sup>, S<sup>-</sup>, Cl<sup>-</sup>  
(12s9p/6s5p) Na-Ar, P<sup>-</sup>, S<sup>-</sup>, Cl<sup>-</sup>

Dunning, T.H.Jr.; Hay, P.J. Chapter 1 in (11s7p/6s4p) Al-Cl  
'Methods of Electronic Structure Theory',  
Schaefer, H.F.III, Ed., Plenum Press,

N.Y., 1977.

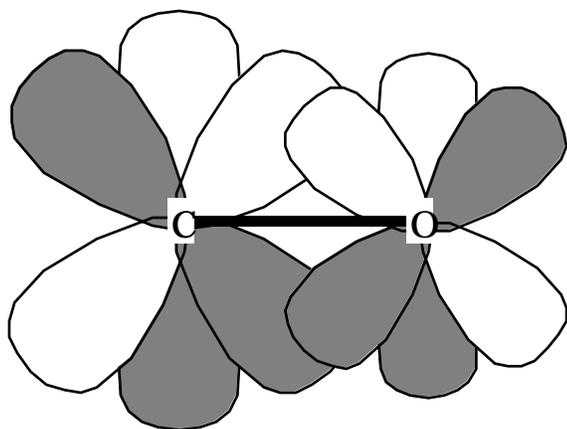
Hood, D.M.; Pitzer, R.M.; Schaefer, H.F.III (14s11p6d/10s8p3d) Sc-Zn  
JCP 71, 705 (1979).

Schmidt, M.W.; Ruedenberg, K.	([N]s), N=3-10	H
JCP <u>71</u> , 3951 (1979).	([2N]s), N=3-10	He
(regular even-tempered)	([2N]s), N=3-14	Li,Be
	([2N]s[N]p),N=3-11	B,N-Ne
	([2N]s[N]p),N=3-13	C
	([2N]s[N]p),N=4-12	Na,Mg
	([2N-6]s[N]p),N=7-15	Al-Ar

#### D. Polarization Functions

In addition to the fundamental core and valence basis described above, one usually adds a set of so-called polarization functions to the basis. Polarization functions are functions of one higher angular momentum than appears in the atom's valence orbital space (e.g, d-functions for C, N , and O and p-functions for H). These polarization functions have exponents ( or ) which cause their radial sizes to be similar to the sizes of the primary valence orbitals ( i.e., the polarization p orbitals of the H atom are similar in size to the 1s orbital). Thus, they are not orbitals which provide a description of the atom's valence orbital with one higher l-value; such higher-l valence orbitals would be radially more diffuse and would therefore require the use of STOs or GTOs with smaller exponents.

The primary purpose of polarization functions is to give additional angular flexibility to the LCAO-MO process in forming the valence molecular orbitals. This is illustrated below where polarization d orbitals are seen to contribute to formation of the bonding orbital of a carbonyl group by allowing polarization of the Carbon atom's p orbital toward the right and of the Oxygen atom's p orbital toward the left.



Polarization functions are essential in strained ring compounds because they provide the angular flexibility needed to direct the electron density into regions between bonded atoms.

Functions with higher  $l$ -values and with 'sizes' more in line with those of the lower- $l$  orbitals are also used to introduce additional angular correlation into the calculation by permitting polarized orbital pairs (see Chapter 10) involving higher angular correlations to be formed. Optimal polarization functions for first and second row atoms have been tabulated (B. Roos and P. Siegbahn, *Theoret. Chim. Acta (Berl.)* 17, 199 (1970); M. J. Frisch, J. A. Pople, and J. S. Binkley, *J. Chem. Phys.* 80, 3265 (1984)).

#### E. Diffuse Functions

When dealing with anions or Rydberg states, one must augment the above basis sets by adding so-called diffuse basis orbitals. The conventional valence and polarization functions described above do not provide enough radial flexibility to adequately describe either of these cases. Energy-optimized diffuse functions appropriate to anions of most lighter main group elements have been tabulated in the literature (an excellent source of Gaussian basis set information is provided in Handbook of Gaussian Basis Sets, R. Poirier, R. Kari, and I. G. Csizmadia, Elsevier, Amsterdam (1985)) and in data bases. Rydberg diffuse basis sets are usually created by adding to conventional valence-plus-polarization bases sequences of primitive GTOs whose exponents are smaller than that (call it  $\alpha_{diff}$ ) of the most diffuse GTO which contributes strongly to the valence CGTOs. As a 'rule of thumb', one can generate a series of such diffuse orbitals which are linearly independent yet span considerably different regions of radial space by introducing primitive GTOs whose exponents are  $\alpha_{diff}/3$ ,  $\alpha_{diff}/9$ ,  $\alpha_{diff}/27$ , etc.

Once one has specified an atomic orbital basis for each atom in the molecule, the LCAO-MO procedure can be used to determine the  $C_{ij}$  coefficients that describe the

occupied and virtual orbitals in terms of the chosen basis set. It is important to keep in mind that the basis orbitals are not themselves the true orbitals of the isolated atoms; even the proper atomic orbitals are combinations (with atomic values for the  $C_{\mu,i}$  coefficients) of the basis functions. For example, in a minimal-basis-level treatment of the Carbon atom, the 2s atomic orbital is formed by combining, with opposite sign to achieve the radial node, the two CGTOs (or STOs); the more diffuse s-type basis function will have a larger  $C_{\mu,i}$  coefficient in the 2s atomic orbital. The 1s atomic orbital is formed by combining the same two CGTOs but with the same sign and with the less diffuse basis function having a larger  $C_{\mu,i}$  coefficient. The LCAO-MO-SCF process itself determines the magnitudes and signs of the  $C_{\mu,i}$ .

## VI. The Roothaan Matrix SCF Process

The matrix SCF equations introduced earlier

$$F_{\mu,i} C_{\mu,i} = \epsilon_i S_{\mu,i} C_{\mu,i}$$

must be solved both for the occupied and virtual orbitals' energies  $\epsilon_i$  and  $C_{\mu,i}$  values. Only the occupied orbitals'  $C_{\mu,i}$  coefficients enter into the Fock operator

$$F_{\mu,i} = \langle \mu | h | \mu \rangle + \sum_j^{\text{occ}} [ 2 \langle \mu | g | \mu \rangle - \langle \mu | g | j \rangle - \langle j | g | \mu \rangle ],$$

but both the occupied and virtual orbitals are solutions of the SCF equations. Once atomic basis sets have been chosen for each atom, the one- and two-electron integrals appearing in  $F_{\mu,i}$  must be evaluated. Doing so is a time consuming process, but there are presently several highly efficient computer codes which allow such integrals to be computed for s, p, d, f, and even g, h, and i basis functions. After executing one of these 'integral packages' for a basis with a total of N functions, one has available (usually on the computer's hard disk) of the order of  $N^2/2$  one-electron and  $N^4/8$  two-electron integrals over these atomic basis orbitals (the factors of 1/2 and 1/8 arise from permutational symmetries of the integrals). When treating extremely large atomic orbital basis sets (e.g., 200 or more basis functions), modern computer programs calculate the requisite integrals but never store them

on the disk. Instead, their contributions to  $F_{\mu, i}$  are accumulated 'on the fly' after which the integrals are discarded.

To begin the SCF process, one must input to the computer routine which computes  $F_{\mu, i}$  initial 'guesses' for the  $C_{,i}$  values corresponding to the occupied orbitals. These initial guesses are typically made in one of the following ways:

1. If one has available  $C_{,i}$  values for the system from an SCF calculation performed earlier at a nearby molecular geometry, one can use these  $C_{,i}$  values to begin the SCF process.
2. If one has  $C_{,i}$  values appropriate to fragments of the system (e.g., for C and O atoms if the CO molecule is under study or for CH<sub>2</sub> and O if H<sub>2</sub>CO is being studied), one can use these.
3. If one has no other information available, one can carry out one iteration of the SCF process in which the two-electron contributions to  $F_{\mu, i}$  are ignored (i.e., take  $F_{\mu, i} = \langle \mu | h | i \rangle$ ) and use the resultant solutions to  $F_{\mu, i} = \sum_j C_{,j} S_{\mu, j}$  as initial guesses for the  $C_{,i}$ . Using only the one-electron part of the Hamiltonian to determine initial values for the LCAO-MO coefficients may seem like a rather severe step; it is, and the resultant  $C_{,i}$  values are usually far from the converged values which the SCF process eventually produces. However, the initial  $C_{,i}$  obtained in this manner have proper symmetries and nodal patterns because the one-electron part of the Hamiltonian has the same symmetry as the full Hamiltonian.

Once initial guesses are made for the  $C_{,i}$  of the occupied orbitals, the full  $F_{\mu, i}$  matrix is formed and new  $\epsilon_i$  and  $C_{,i}$  values are obtained by solving  $F_{\mu, i} = \epsilon_i S_{\mu, i}$ . These new orbitals are then used to form a new  $F_{\mu, i}$  matrix from which new  $\epsilon_i$  and  $C_{,i}$  are obtained. This iterative process is carried on until the  $\epsilon_i$  and  $C_{,i}$  do not vary (within specified tolerances) from iteration to iteration, at which time one says that the SCF process has converged and reached self-consistency.

As presented, the Roothaan SCF process is carried out in a fully ab initio manner in that all one- and two-electron integrals are computed in terms of the specified basis set; no experimental data or other input is employed. As described in Appendix F, it is possible to introduce approximations to the coulomb and exchange integrals entering into the Fock matrix elements that permit many of the requisite  $F_{\mu, i}$  elements to be evaluated in terms of experimental data or in terms of a small set of 'fundamental' orbital-level coulomb interaction integrals that can be computed in an ab initio manner. This approach forms the basis of so-called 'semi-empirical' methods. Appendix F provides the reader with a brief introduction to such approaches to the electronic structure problem and deals in some detail with the well known Hückel and CNDO- level approximations.

## VII. Observations on Orbitals and Orbital Energies

### A. The Meaning of Orbital Energies

The physical content of the Hartree-Fock orbital energies can be seen by observing that  $F_i = \epsilon_i$  implies that  $\psi_i$  can be written as:

$$\begin{aligned} \epsilon_i &= \langle \psi_i | F | \psi_i \rangle = \langle \psi_i | h | \psi_i \rangle + \sum_{j(\text{occupied})} \langle \psi_i | J_j - K_j | \psi_i \rangle \\ &= \langle \psi_i | h | \psi_i \rangle + \sum_{j(\text{occupied})} [J_{i,j} - K_{i,j}]. \end{aligned}$$

In this form, it is clear that  $\epsilon_i$  is equal to the average value of the kinetic energy plus coulombic attraction to the nuclei for an electron in  $\psi_i$  plus the sum over all of the spin-orbitals occupied in  $\psi$  of coulomb minus exchange interactions between  $\psi_i$  and these occupied spin-orbitals. If  $\psi_i$  itself is an occupied spin-orbital, the term  $[J_{i,i} - K_{i,i}]$  disappears and the latter sum represents the coulomb minus exchange interaction of  $\psi_i$  with all of the  $N-1$  other occupied spin-orbitals. If  $\psi_i$  is a virtual spin-orbital, this cancellation does not occur, and one obtains the coulomb minus exchange interaction of  $\psi_i$  with all  $N$  of the occupied spin-orbitals.

In this sense, the orbital energies for occupied orbitals pertain to interactions which are appropriate to a total of  $N$  electrons, while the orbital energies of virtual orbitals pertain to a system with  $N+1$  electrons. It is this fact that makes SCF virtual orbitals not optimal (in fact, not usually very good) for use in subsequent correlation calculations where, for instance, they are used, in combination with the occupied orbitals, to form polarized orbital pairs as discussed in Chapter 12. To correlate a pair of electrons that occupy a valence orbital requires double excitations into a virtual orbital that is not too dislike in size. Although the virtual SCF orbitals themselves suffer these drawbacks, the space they span can indeed be used for treating electron correlation. To do so, it is useful to recombine (in a unitary manner to preserve orthonormality) the virtual orbitals to 'focus' the correlating power into as few orbitals as possible so that the multiconfigurational wavefunction can be formed with as few CSFs as possible. Techniques for effecting such reoptimization or improvement of the virtual orbitals are treated later in this text.

### B.. Koopmans' Theorem

Further insight into the meaning of the energies of occupied and virtual orbitals can be gained by considering the following model of the vertical (i.e., at fixed molecular geometry) detachment or attachment of an electron to the original N-electron molecule:

1. In this model, both the parent molecule and the species generated by adding or removing an electron are treated at the single-determinant level.
2. In this model, the Hartree-Fock orbitals of the parent molecule are used to describe both the parent and the species generated by electron addition or removal. It is said that such a model neglects 'orbital relaxation' which would accompany the electron addition or removal (i.e., the reoptimization of the spin-orbitals to allow them to become appropriate to the daughter species).

Within this simplified model, the energy difference between the daughter and the parent species can be written as follows (  $\epsilon_k$  represents the particular spin-orbital that is added or removed):

1. For electron detachment:

$$E^{N-1} - E^N = \langle | 1 2 \dots k-1 \dots N | H | 1 2 \dots k-1 \dots N | \rangle - \langle | 1 2 \dots k-1 k \dots N | H | | 1 2 \dots k-1 k \dots N | \rangle = - \epsilon_k ;$$

2. For electron attachment:

$$E^N - E^{N+1} = \langle | 1 2 \dots N | H | 1 2 \dots N | \rangle - \langle | 1 2 \dots N k | H | | 1 2 \dots N k | \rangle = - \epsilon_k .$$

So, within the limitations of the single-determinant, frozen-orbital model set forth, the ionization potentials (IPs) and electron affinities (EAs) are given as the negative of the occupied and virtual spin-orbital energies, respectively. This statement is referred to as Koopmans' theorem (T. Koopmans, *Physica* 1, 104 (1933)); it is used extensively in quantum chemical calculations as a means for estimating IPs and EAs and often yields results that are at least qualitatively correct (i.e.,  $\pm 0.5$  eV).

### C. Orbital Energies and the Total Energy

For the N-electron species whose Hartree-Fock orbitals and orbital energies have been determined, the total SCF electronic energy can be written, by using the Slater-Condon rules, as:

$$E = \sum_{i(\text{occupied})} \langle i | h | i \rangle + \sum_{i>j(\text{occupied})} [J_{i,j} - K_{i,j}].$$

For this same system, the sum of the orbital energies of the occupied spin-orbitals is given by:

$$\sum_{i(\text{occupied})} \epsilon_i = \sum_{i(\text{occupied})} \langle i | h | i \rangle + \sum_{i,j(\text{occupied})} [J_{i,j} - K_{i,j}].$$

These two seemingly very similar expressions differ in a very important way; the sum of occupied orbital energies, when compared to the total energy, double counts the coulomb minus exchange interaction energies. Thus, within the Hartree-Fock approximation, the sum of the occupied orbital energies is not equal to the total energy. The total SCF energy can be computed in terms of the sum of occupied orbital energies by taking one-half of  $\sum_{i(\text{occupied})} \epsilon_i$  and then adding to this one-half of  $\sum_{i(\text{occupied})} \langle i | h | i \rangle$ :

$$E = 1/2 \left[ \sum_{i(\text{occupied})} \langle i | h | i \rangle + \sum_{i(\text{occupied})} \epsilon_i \right].$$

The fact that the sum of orbital energies is not the total SCF energy also means that as one attempts to develop a qualitative picture of the energies of CSFs along a reaction path, as when orbital and configuration correlation diagrams are constructed, one must be careful not to equate the sum of orbital energies with the total configurational energy; the former is higher than the latter by an amount equal to the sum of the coulomb minus exchange interactions.

### D. The Brillouin Theorem

The condition that the SCF energy  $\langle \psi_{1\dots N} | H | \psi_{1\dots N} \rangle$  be stationary with respect to variations  $\psi_i$  in the occupied spin-orbitals (that preserve orthonormality) can be written

$$\langle \phi_{1\dots i\dots N} | \mathbf{H} | \phi_{1\dots i\dots N} \rangle = 0.$$

The infinitesimal variation of  $\phi_i$  can be expressed in terms of its (small) components along the other occupied  $\phi_j$  and along the virtual  $\phi_m$  as follows:

$$\delta \phi_i = \sum_{j=\text{occ}} U_{ij} \phi_j + \sum_m U_{im} \phi_m.$$

When substituted into  $\langle \phi_{1\dots i\dots N} |$ , the terms  $\sum_{j=\text{occ}} \langle \phi_{1\dots j\dots N} | U_{ij}$  vanish because  $\phi_j$  already appears in the original Slater determinant  $|\phi_{1\dots N}\rangle$ , so  $|\phi_{1\dots j\dots N}\rangle$  contains  $\phi_j$  twice. Only the sum over virtual orbitals remains, and the stationary property written above becomes

$$\sum_m U_{im} \langle \phi_{1\dots m\dots N} | \mathbf{H} | \phi_{1\dots i\dots N} \rangle = 0.$$

The Slater-Condon rules allow one to express the Hamiltonian matrix elements appearing here as

$$\langle \phi_{1\dots m\dots N} | \mathbf{H} | \phi_{1\dots i\dots N} \rangle = \langle \phi_m | \mathbf{h} | \phi_i \rangle + \sum_{j=\text{occ}, j \neq i} \langle \phi_m | [J_j - K_j] | \phi_i \rangle,$$

which (because the term with  $j=i$  can be included since it vanishes) is equal to the following element of the Fock operator:  $\langle \phi_m | \mathbf{F} | \phi_i \rangle = \delta_{im} = 0$ . This result proves that Hamiltonian matrix elements between the SCF determinant and those that are singly excited relative to the SCF determinant vanish because they reduce to Fock-operator integrals connecting the pair of orbitals involved in the 'excitation'. This stability property of the SCF energy is known as the Brillouin theorem (i.e., that  $|\phi_{1\dots i\dots N}\rangle$  and  $|\phi_{1\dots m\dots N}\rangle$  have zero Hamiltonian matrix elements if the  $\phi$ s are SCF orbitals). It is exploited in quantum chemical calculations in two manners:

(i) When multiconfiguration wavefunctions are formed from SCF spin-orbitals, it allows one to neglect Hamiltonian matrix elements between the SCF configuration and those that are 'singly excited' in constructing the secular matrix.

(ii) A so-called generalized Brillouin theorem (GBT) arises when one deals with energy optimization for a multiconfigurational variational trial wavefunction for which the orbitals and  $C_I$  mixing coefficients are simultaneously optimized. This GBT causes certain Hamiltonian matrix elements to vanish, which, in turn, simplifies the treatment of electron correlation for such wavefunctions. This matter is treated in more detail later in this text.



## Chapter 19

*Corrections to the mean-field model are needed to describe the instantaneous Coulombic interactions among the electrons. This is achieved by including more than one Slater determinant in the wavefunction.*

Much of the development of the previous chapter pertains to the use of a single Slater determinant trial wavefunction. As presented, it relates to what has been called the unrestricted Hartree-Fock (UHF) theory in which each spin-orbital  $\phi_i$  has its own orbital energy  $\epsilon_i$  and LCAO-MO coefficients  $C_{\mu i}$ ; there may be different  $C_{\mu i}$  for  $\alpha$  spin-orbitals than for  $\beta$  spin-orbitals. Such a wavefunction suffers from the spin contamination difficulty detailed earlier.

To allow for a properly spin- and space- symmetry adapted trial wavefunction and to permit  $\Psi$  to contain more than a single CSF, methods which are more flexible than the single-determinant HF procedure are needed. In particular, it may be necessary to use a combination of determinants to describe such a proper symmetry function. Moreover, as emphasized earlier, whenever two or more CSFs have similar energies (i.e., Hamiltonian expectation values) and can couple strongly through the Hamiltonian (e.g., at avoided crossings in configuration correlation diagrams), the wavefunction must be described in a multiconfigurational manner to permit the wavefunction to evolve smoothly from reactants to products. Also, whenever dynamical electron correlation effects are to be treated, a multiconfigurational  $\Psi$  must be used; in this case, CSFs that are doubly excited relative to one or more of the essential CSFs (i.e., the dominant CSFs that are included in the so-called reference wavefunction) are included to permit polarized-orbital-pair formation.

Multiconfigurational functions are needed not only to account for electron correlation but also to permit orbital readjustments to occur. For example, if a set of SCF orbitals is employed in forming a multi-CSF wavefunction, the variational condition that the energy is stationary with respect to variations in the LCAO-MO coefficients is no longer obeyed (i.e., the SCF energy functional is stationary when SCF orbitals are employed, but the MC-energy functional is generally not stationary if SCF orbitals are employed). For such reasons, it is important to include CSFs that are singly excited relative to the dominant CSFs in the reference wavefunction.

That singly excited CSFs allow for orbital relaxation can be seen as follows. Consider a wavefunction consisting of one CSF  $|\phi_1 \dots \phi_i \dots \phi_N\rangle$  to which singly excited CSFs of the form  $|\phi_1 \dots \phi_m \dots \phi_N\rangle$  have been added with coefficients  $C_{i,m}$ :

$$\Psi = \sum_m C_{i,m} |\phi_1 \dots \phi_m \dots \phi_N\rangle + |\phi_1 \dots \phi_i \dots \phi_N\rangle$$

All of these determinants have all of their columns equal except the  $i^{\text{th}}$  column; therefore, they can be combined into a single new determinant:

$$= | 1 \dots i' \dots N |,$$

where the relaxed orbital  $i'$  is given by

$$i' = i + \sum_m C_{i,m} \phi_m.$$

The sum of CSFs that are singly excited in the  $i^{\text{th}}$  spin-orbital with respect to  $| 1 \dots i \dots N |$  is therefore seen to allow the spin-orbital  $i$  to relax into the new spin-orbital  $i'$ . It is in this sense that singly excited CSFs allow for orbital reoptimization.

In summary, doubly excited CSFs are often employed to permit polarized orbital pair formation and hence to allow for electron correlations. Singly excited CSFs are included to permit orbital relaxation (i.e., orbital reoptimization) to occur.

## I. Different Methods

There are numerous procedures currently in use for determining the 'best' wavefunction of the form:

$$= \sum_I C_I \Phi_I,$$

where  $\Phi_I$  is a spin-and space- symmetry adapted CSF consisting of determinants of the form  $| \Phi_1 \Phi_2 \Phi_3 \dots \Phi_N |$ . Excellent overviews of many of these methods are included in Modern Theoretical Chemistry Vols. 3 and 4, H. F. Schaefer, III Ed., Plenum Press, New York (1977) and in Advances in Chemical Physics, Vols. LXVII and LXIX, K. P. Lawley, Ed., Wiley-Interscience, New York (1987). Within the present Chapter, these two key references will be denoted MTC, Vols. 3 and 4, and ACP, Vols. 67 and 69, respectively.

In all such trial wavefunctions, there are two fundamentally different kinds of parameters that need to be determined- the CI coefficients  $C_I$  and the LCAO-MO coefficients describing the  $\Phi_k$ . The most commonly employed methods used to determine these parameters include:

1. The **multiconfigurational self-consistent field** (MCSCF) method in which the expectation value  $\langle \Psi | H | \Psi \rangle / \langle \Psi | \Psi \rangle$  is treated variationally and simultaneously made stationary with respect to variations in the  $C_I$  and  $C_{\mu,i}$  coefficients subject to the constraints that the spin-orbitals and the full N-electron wavefunction remain normalized:

$$\langle \Psi | \Psi \rangle = \sum_{i,j} C_i C_j S_{ij} = 1, \text{ and}$$

$$\sum_I C_I^2 = 1.$$

The articles by H.-J. Werner and by R. Shepard in ACP Vol. 69 provide up to date reviews of the status of this approach. The article by A. C. Wahl and G. Das in MTC Vol. 3 covers the 'earlier' history on this topic. F. W. Bobrowicz and W. A. Goddard, III provide, in MTC Vol. 3, an overview of the GVB approach, which, as discussed in Chapter 12, can be viewed as a specific kind of MCSCF calculation.

2. The **configuration interaction** (CI) method in which the LCAO-MO coefficients are determined first (and independently) via either a single-configuration SCF calculation or an MCSCF calculation using a small number of CSFs. The CI coefficients are subsequently determined by making the expectation value  $\langle \Psi | H | \Psi \rangle / \langle \Psi | \Psi \rangle$  stationary with respect to variations in the  $C_I$  only. In this process, the optimizations of the orbitals and of the CSF amplitudes are done in separate steps. The articles by I. Shavitt and by B. O. Ross and P. E. M. Siegbahn in MTC, Vol. 3 give excellent early overviews of the CI method.

3. The **Møller-Plesset perturbation method** (MPPT) uses the single-configuration SCF process (usually the UHF implementation) to first determine a set of LCAO-MO coefficients and, hence, a set of orbitals that obey  $F_i = \epsilon_i \phi_i$ . Then, using an unperturbed Hamiltonian equal to the sum of these Fock operators for each of the N electrons  $H^0 = \sum_{i=1,N} F(i)$ , perturbation theory (see Appendix D for an introduction to time-independent perturbation theory) is used to determine the  $C_I$  amplitudes for the CSFs. The MPPT procedure is also referred to as the many-body perturbation theory (MBPT) method. The two names arose because two different schools of physics and chemistry developed them for somewhat different applications. Later, workers realized that they were identical in their working equations when the UHF  $H^0$  is employed as the unperturbed Hamiltonian. In this text, we will therefore refer to this approach as MPPT/MBPT.

The amplitude for the so-called reference CSF used in the SCF process is taken as unity and the other CSFs' amplitudes are determined, relative to this one, by Rayleigh-Schrödinger perturbation theory using the full N-electron Hamiltonian minus the sum of Fock operators  $H-H^0$  as the perturbation. The Slater-Condon rules are used for evaluating matrix elements of  $(H-H^0)$  among these CSFs. The essential features of the MPPT/MBPT approach are described in the following articles: J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, *Int. J. Quantum Chem.* 14, 545 (1978); R. J. Bartlett and D. M. Silver, *J. Chem. Phys.* 62, 3258 (1975); R. Krishnan and J. A. Pople, *Int. J. Quantum Chem.* 14, 91 (1978).

4. The **Coupled-Cluster method** expresses the CI part of the wavefunction in a somewhat different manner (the early work in chemistry on this method is described in J. Cizek, *J. Chem. Phys.* 45, 4256 (1966); J. Paldus, J. Cizek, and I. Shavitt, *Phys. Rev.* A5, 50 (1972); R. J. Bartlett and G. D. Purvis, *Int. J. Quantum Chem.* 14, 561 (1978); G. D. Purvis and R. J. Bartlett, *J. Chem. Phys.* 76, 1910 (1982)):

$$= \exp(T) \quad ,$$

where  $\Phi_0$  is a single CSF (usually the UHF single determinant) which has been used to independently determine a set of spin-orbitals and LCAO-MO coefficients via the SCF process. The operator T generates, when acting on  $\Phi_0$ , single, double, etc. 'excitations' (i.e., CSFs in which one, two, etc. of the occupied spin-orbitals in  $\Phi_0$  have been replaced by virtual spin-orbitals). T is commonly expressed in terms of operators that effect such spin-orbital removals and additions as follows:

$$T = \sum_{i,m} t_i^m m^+ i + \sum_{i,j,m,n} t_{i,j}^{m,n} m^+ n^+ j i + \dots,$$

where the operator  $m^+$  is used to denote creation of an electron in virtual spin-orbital  $m$  and the operator  $j$  is used to denote removal of an electron from occupied spin-orbital  $j$ .

The  $t_i^m$ ,  $t_{i,j}^{m,n}$ , etc. amplitudes, which play the role of the CI coefficients in CC theory, are determined through the set of equations generated by projecting the Schrödinger equation in the form

$$\exp(-T) H \exp(T) \Phi_0 = E \Phi_0$$

against CSFs which are single, double, etc. excitations relative to  $|\Phi\rangle$ . For example, for double excitations  $|\Phi_{ij}^{m,n}\rangle$  the equations read:

$$\langle \Phi_{ij}^{m,n} | \exp(-T) H \exp(T) | \Phi \rangle = E \langle \Phi_{ij}^{m,n} | \Phi \rangle = 0;$$

zero is obtained on the right hand side because the excited CSFs  $|\Phi_{ij}^{m,n}\rangle$  are orthogonal to the reference function  $|\Phi\rangle$ . The elements on the left hand side of the CC equations can be expressed, as described below, in terms of one- and two-electron integrals over the spin-orbitals used in forming the reference and excited CSFs.

### A. Integral Transformations

All of the above methods require the evaluation of one- and two-electron integrals over the  $N$  atomic orbital basis:  $\langle a | f | b \rangle$  and  $\langle a b | g | c d \rangle$ . Eventually, all of these methods provide their working equations and energy expressions in terms of one- and two-electron integrals over the  $N$  final molecular orbitals:  $\langle i | f | j \rangle$  and  $\langle i j | g | k l \rangle$ .

The mo-based integrals can only be evaluated by transforming the AO-based integrals as follows:

$$\langle i j | g | k l \rangle = \sum_{a,b,c,d} C_{a,i} C_{b,j} C_{c,k} C_{d,l} \langle a b | g | c d \rangle,$$

and

$$\langle i | f | j \rangle = \sum_{a,b} C_{a,i} C_{b,j} \langle a | f | b \rangle.$$

It would seem that the process of evaluating all  $N^4$  of the  $\langle i j | g | k l \rangle$ , each of which requires  $N^4$  additions and multiplications, would require computer time proportional to  $N^8$ . However, it is possible to perform the full transformation of the two-electron integral list in a time that scales as  $N^5$ . This is done by first performing a transformation of the  $\langle a b | g | c d \rangle$  to an intermediate array labeled  $\langle a b | g | c l \rangle$  as follows:

$$\langle a b | g | c l \rangle = \sum_d C_{d,l} \langle a b | g | c d \rangle.$$

This partial transformation requires  $N^5$  multiplications and additions.

The list  $\langle a b | g | c l \rangle$  is then transformed to a second-level transformed array

$$\langle a b | g | k l \rangle:$$

$$\langle a | b | g | k | l \rangle = \sum_c C_{c,k} \langle a | b | g | c | l \rangle,$$

which requires another  $N^5$  operations. This sequential, one-index-at-a-time transformation is repeated four times until the final  $\langle i | j | g | k | l \rangle$  array is in hand. The entire transformation done this way requires  $4N^5$  multiplications and additions.

Once the requisite one- and two-electron integrals are available in the molecular orbital basis, the multiconfigurational wavefunction and energy calculation can begin. These transformations consume a large fraction of the computer time used in most such calculations, and represent a severe bottleneck to progress in applying ab initio electronic structure methods to larger systems.

## B. Configuration List Choices

Once the requisite one- and two-electron integrals are available in the molecular orbital basis, the multiconfigurational wavefunction and energy calculation can begin. Each of these methods has its own approach to describing the configurations  $\{ J \}$  included in the calculation and how the  $\{ C_J \}$  amplitudes and the total energy  $E$  is to be determined.

The number of configurations ( $N_C$ ) varies greatly among the methods and is an important factor to keep in mind when planning to carry out an *ab initio* calculation. Under certain circumstances (e.g., when studying Woodward-Hoffmann forbidden reactions where an avoided crossing of two configurations produces an activation barrier), it may be essential to use more than one electronic configuration. Sometimes, one configuration (e.g., the SCF model) is adequate to capture the qualitative essence of the electronic structure. In all cases, many configurations will be needed if highly accurate treatment of electron-electron correlations are desired.

The value of  $N_C$  determines how much computer time and memory is needed to solve the  $N_C$ -dimensional  $\sum_J H_{I,J} C_J = E C_I$  secular problem in the CI and MCSCF methods. Solution of these matrix eigenvalue equations requires computer time that scales as  $N_C^2$  (if few eigenvalues are computed) to  $N_C^3$  (if most eigenvalues are obtained).

So-called complete-active-space (CAS) methods form all CSFs that can be created by distributing  $N$  valence electrons among  $P$  valence orbitals. For example, the eight non-core electrons of  $H_2O$  might be distributed, in a manner that gives  $M_S = 0$ , among six valence orbitals (e.g., two lone-pair orbitals, two OH bonding orbitals, and two OH\* antibonding orbitals). The number of configurations thereby created is 225. If the same eight electrons were distributed among ten valence orbitals 44,100 configurations results;

for twenty and thirty valence orbitals, 23,474,025 and 751,034,025 configurations arise, respectively. Clearly, practical considerations dictate that CAS-based approaches be limited to situations in which a few electrons are to be correlated using a few valence orbitals. The primary advantage of CAS configurations is discussed below in Sec. II. C.

## II. Strengths and Weaknesses of Various Methods

### A. Variational Methods Such as MCSCF, SCF, and CI Produce Energies that are Upper Bounds, but These Energies are not Size-Extensive

Methods that are based on making the energy functional  $\langle \Psi | H | \Psi \rangle / \langle \Psi | \Psi \rangle$  stationary (i.e., variational methods) yield upper bounds to the lowest energy of the symmetry which characterizes the CSFs which comprise  $\Psi$ . These methods also can provide approximate excited-state energies and wavefunctions (e. g., in the form of other solutions of the secular equation  $\sum_J H_{IJ} C_J = E C_I$  that arises in the CI and MCSCF methods). Excited-state energies obtained in this manner can be shown to 'bracket' the true energies of the given symmetry in that between any two approximate energies obtained in the variational calculation, there exists at least one true eigenvalue. This characteristic is commonly referred to as the 'bracketing theorem' (E. A. Hylleraas and B. Undheim, *Z. Phys.* **65**, 759 (1930); J. K. L. MacDonald, *Phys. Rev.* **43**, 830 (1933)). These are strong attributes of the variational methods, as is the long and rich history of developments of analytical and computational tools for efficiently implementing such methods (see the discussions of the CI and MCSCF methods in MTC and ACP).

However, all variational techniques suffer from at least one serious drawback; they are not **size-extensive** (J. A. Pople, pg. 51 in *Energy, Structure, and Reactivity*, D. W. Smith and W. B. McRae, Eds., Wiley, New York (1973)). This means that the energy computed using these tools can not be trusted to scale with the size of the system. For example, a calculation performed on two CH<sub>3</sub> species at large separation may not yield an energy equal to twice the energy obtained by performing the same kind of calculation on a single CH<sub>3</sub> species. Lack of size-extensivity precludes these methods from use in extended systems (e.g., solids) where errors due to improper scaling of the energy with the number of molecules produce nonsensical results.

By carefully adjusting the kind of variational wavefunction used, it is possible to circumvent size-extensivity problems for selected species. For example, a CI calculation on Be<sub>2</sub> using all <sup>1</sup>g CSFs that can be formed by placing the four valence electrons into the orbitals 2<sub>g</sub>, 2<sub>u</sub>, 3<sub>g</sub>, 3<sub>u</sub>, 1<sub>u</sub>, and 1<sub>g</sub> can yield an energy equal to twice that of the Be

atom described by CSFs in which the two valence electrons of the Be atom are placed into the 2s and 2p orbitals in all ways consistent with a  $^1S$  symmetry. Such special choices of configurations give rise to what are called complete-active-space (CAS) MCSCF or CI calculations (see the article by B. O. Roos in ACP for an overview of this approach).

Let us consider an example to understand why the CAS choice of configurations works. The  $^1S$  ground state of the Be atom is known to form a wavefunction that is a strong mixture of CSFs that arise from the  $2s^2$  and  $2p^2$  configurations:

$$\Psi_{\text{Be}} = C_1 |1s^2 2s^2| + C_2 |1s^2 2p^2|,$$

where the latter CSF is a short-hand representation for the proper spin- and space-symmetry adapted CSF

$$|1s^2 2p^2| = \frac{1}{\sqrt{3}} [ |1s \uparrow 1s \uparrow 2p_0 \uparrow 2p_0 \uparrow| - |1s \uparrow 1s \uparrow 2p_1 \uparrow 2p_{-1} \uparrow| - |1s \uparrow 1s \uparrow 2p_{-1} \uparrow 2p_1 \uparrow| ].$$

The reason the CAS process works is that the  $\text{Be}_2$  CAS wavefunction has the flexibility to dissociate into the product of two CAS Be wavefunctions:

$$\Psi_{\text{Be}_2} = \Psi_{\text{Be}_a} \Psi_{\text{Be}_b} = \{C_1 |1s^2 2s^2| + C_2 |1s^2 2p^2|\}_a \{C_1 |1s^2 2s^2| + C_2 |1s^2 2p^2|\}_b,$$

where the subscripts a and b label the two Be atoms, because the four electron CAS function distributes the four electrons in all ways among the  $2s_a$ ,  $2s_b$ ,  $2p_a$ , and  $2p_b$  orbitals. In contrast, if the  $\text{Be}_2$  calculation had been carried out using only the following CSFs :  $|1 \uparrow 2_g 1 \uparrow 2_u 2 \uparrow 2_g 2 \uparrow 2_u|$  and all single and double excitations relative to this (dominant) CSF, which is a very common type of CI procedure to follow, the  $\text{Be}_2$  wavefunction would not have contained the particular CSFs  $|1s^2 2p^2|_a |1s^2 2p^2|_b$  because these CSFs are four-fold excited relative to the  $|1 \uparrow 2_g 1 \uparrow 2_u 2 \uparrow 2_g 2 \uparrow 2_u|$  'reference' CSF.

In general, one finds that if the 'monomer' uses CSFs that are K-fold excited relative to its dominant CSF to achieve an accurate description of its electron correlation, a size-extensive variational calculation on the 'dimer' will require the inclusion of CSFs that are 2K-fold excited relative to the dimer's dominant CSF. To perform a size-extensive

variational calculation on a species containing  $M$  monomers therefore requires the inclusion of CSFs that are  $M \times K$ -fold excited relative to the  $M$ -mer's dominant CSF.

## B. Non-Variational Methods Such as MPPT/MBPT and CC do not Produce Upper Bounds, but Yield Size-Extensive Energies

In contrast to variational methods, perturbation theory and coupled-cluster methods achieve their energies from a 'transition formula'  $\langle \Psi | H | \Phi \rangle$  rather than from an expectation value

$\langle \Psi | H | \Psi \rangle$ . It can be shown (H. P. Kelly, Phys. Rev. 131, 684 (1963)) that this difference allows non-variational techniques to yield size-extensive energies. This can be seen in the MPPT/MBPT case by considering the energy of two non-interacting Be atoms. The reference CSF is  $|\Phi\rangle = |1s_a^2 2s_a^2 1s_b^2 2s_b^2|$ ; the Slater-Condon rules limit the CSFs in which can contribute to

$$E = \langle \Psi | H | \Psi \rangle = \langle \Psi | H | \sum_J C_J \Psi_J \rangle,$$

to be  $\Psi$  itself and those CSFs that are singly or doubly excited relative to  $\Psi$ . These 'excitations' can involve atom a, atom b, or both atoms. However, any CSFs that involve excitations on both atoms

(e.g.,  $|1s_a^2 2s_a 2p_a 1s_b^2 2s_b 2p_b|$ ) give rise, via the SC rules, to one- and two- electron integrals over orbitals on both atoms; these integrals (e.g.,  $\langle 2s_a 2p_a | g | 2s_b 2p_b \rangle$ ) vanish if the atoms are far apart, as a result of which the contributions due to such CSFs vanish in our consideration of size-extensivity. Thus, only CSFs that are excited on one or the other atom contribute to the energy:

$$E = \langle \Psi_a | H | \sum_{J_a} C_{J_a} \Psi_{J_a} + \sum_{J_b} C_{J_b} \Psi_{J_b} \rangle,$$

where  $\Psi_a$  and  $\Psi_b$  as well as  $\Psi_{J_a}$  and  $\Psi_{J_b}$  are used to denote the a and b parts of the reference and excited CSFs, respectively.

This expression, once the SC rules are used to reduce it to one- and two- electron integrals, is of the additive form required of any size-extensive method:

$$E = \langle \Psi_a | H | \sum_{J_a} C_{J_a} \Psi_{J_a} \rangle + \langle \Psi_b | H | \sum_{J_b} C_{J_b} \Psi_{J_b} \rangle,$$

and will yield a size-extensive energy if the equations used to determine the  $C_{Ja}$  and  $C_{Jb}$  amplitudes are themselves separable. In MPPT/MBPT, these amplitudes are expressed, in first order, as:

$$C_{Ja} = \langle a | b | H | \psi_{Jab}^* \rangle / [E_a^0 + E_b^0 - E_{Ja}^* - E_b^0]$$

(and analogously for  $C_{Jb}$ ). Again using the SC rules, this expression reduces to one that involves only atom a:

$$C_{Ja} = \langle a | H | \psi_{Ja}^* \rangle / [E_a^0 - E_{Ja}^*].$$

The additivity of  $E$  and the separability of the equations determining the  $C_J$  coefficients make the MPPT/MBPT energy size-extensive. This property can also be demonstrated for the Coupled-Cluster energy (see the references given above in Chapter 19. I.4). However, size-extensive methods have at least one serious weakness; their energies do not provide upper bounds to the true energies of the system (because their energy functional is not of the expectation-value form for which the upper bound property has been proven).

### C. Which Method is Best?

At this time, it may not possible to say which method is preferred for applications where all are practical. Nor is it possible to assess, in a way that is applicable to most chemical species, the accuracies with which various methods predict bond lengths and energies or other properties. However, there are reasons to recommend some methods over others in specific cases. For example, certain applications require a size-extensive energy (e.g., extended systems that consist of a large or macroscopic number of units or studies of weak intermolecular interactions), so MBPT/MPPT or CC or CAS-based MCSCF are preferred. Moreover, certain chemical reactions (e.g., Woodward-Hoffmann forbidden reactions) and certain bond-breaking events require two or more 'essential' electronic configurations. For them, single-configuration-based methods such as conventional CC and MBTP/MPPT should not be used; MCSCF or CI calculations would be better. Very large molecules, in which thousands of atomic orbital basis functions are required, may be impossible to treat by methods whose effort scales as  $N^4$  or higher; density functional methods would be better to use then.

For all calculations, the choice of atomic orbital basis set must be made carefully, keeping in mind the  $N^4$  scaling of the one- and two-electron integral evaluation step and the

$N^5$  scaling of the two-electron integral transformation step. Of course, basis functions that describe the essence of the states to be studied are essential (e.g., Rydberg or anion states require diffuse functions, and strained rings require polarization functions).

As larger atomic basis sets are employed, the size of the CSF list used to treat dynamic correlation increases rapidly. For example, most of the above methods use singly and doubly excited CSFs for this purpose. For large basis sets, the number of such CSFs,  $N_C$ , scales as the number of electrons squared,  $n_e^2$ , times the number of basis functions squared,  $N^2$ . Since the effort needed to solve the CI secular problem varies as  $N_C^2$  or  $N_C^3$ , a dependence as strong as  $N^4$  to  $N^6$  can result. To handle such large CSF spaces, all of the multiconfigurational techniques mentioned in this paper have been developed to the extent that calculations involving of the order of 100 to 5,000 CSFs are routinely performed and calculations using 10,000, 100,000, and even several million CSFs are practical.

Other methods, most of which can be viewed as derivatives of the techniques introduced above, have been and are still being developed. This ongoing process has been, in large part, stimulated by the explosive growth in computer power and change in computer architecture that has been realized in recent years. All indications are that this growth pattern will continue, so *ab initio* quantum chemistry will likely have an even larger impact on future chemistry research and education (through new insights and concepts).

### III. Further Details on Implementing Multiconfigurational Methods

#### A. The MCSCF Method

The simultaneous optimization of the LCAO-MO and CI coefficients performed within an MCSCF calculation is a quite formidable task. The variational energy functional is a quadratic function of the CI coefficients, and so one can express the stationary conditions for these variables in the secular form:

$$\sum_J H_{IJ} C_J = E C_I.$$

However,  $E$  is a quartic function of the  $C_{,i}$  coefficients because each matrix element  $\langle I | H | J \rangle$  involves one- and two-electron integrals over the mos  $\psi_i$ , and the two-electron integrals depend quartically on the  $C_{,i}$  coefficients. The stationary conditions with respect to these  $C_{,i}$  parameters must be solved iteratively because of this quartic dependence.

It is well known that minimization of a function ( $E$ ) of several non-linear parameters (the  $C_{i,j}$ ) is a difficult task that can suffer from poor convergence and may locate local rather than global minima. In an MCSCF wavefunction containing many CSFs, the energy is only weakly dependent on the orbitals that are weakly occupied (i.e., those that appear in CSFs with small  $C_I$  values); in contrast,  $E$  is strongly dependent on the  $C_{i,j}$  coefficients of those orbitals that appear in the CSFs with larger  $C_I$  values. One is therefore faced with minimizing a function of many variables (there may be as many  $C_{i,j}$  as the square of the number of orbital basis functions) that depends strongly on several of the variables and weakly on many others. This is a very difficult job.

For these reasons, in the MCSCF method, the number of CSFs is usually kept to a small to moderate number (e.g., a few to several hundred) chosen to describe essential correlations (i.e., configuration crossings, proper dissociation) and important dynamical correlations (those electron-pair correlations of angular, radial, left-right, etc. nature that arise when low-lying 'virtual' orbitals are present). In such a compact wavefunction, only spin-orbitals with reasonably large occupations (e.g., as characterized by the diagonal elements of the one-particle density matrix  $\rho_{i,j}$ ) appear. As a result, the energy functional is expressed in terms of variables on which it is strongly dependent, in which case the non-linear optimization process is less likely to be pathological.

Such a compact MCSCF wavefunction is designed to provide a good description of the set of strongly occupied spin-orbitals and of the CI amplitudes for CSFs in which only these spin-orbitals appear. It, of course, provides no information about the spin-orbitals that are not used to form the CSFs on which the MCSCF calculation is based. As a result, the MCSCF energy is invariant to a unitary transformation among these 'virtual' orbitals.

In addition to the references mentioned earlier in ACP and MTC, the following papers describe several of the advances that have been made in the MCSCF method, especially with respect to enhancing its rate and range of convergence: E. Dalggaard and P. Jørgensen, *J. Chem. Phys.* 69, 3833 (1978); H. J. Aa. Jensen, P. Jørgensen, and H. Ågren, *J. Chem. Phys.* 87, 457 (1987); B. H. Lengsfeld, III and B. Liu, *J. Chem. Phys.* 75, 478 (1981).

## B. The Configuration Interaction Method

In the CI method, one usually attempts to realize a high-level treatment of electron correlation. A set of orthonormal molecular orbitals are first obtained from an SCF or MCSCF calculation (usually involving a small to moderate list of CSFs). The LCAO-MO

coefficients of these orbitals are no longer considered as variational parameters in the subsequent CI calculation; only the  $C_I$  coefficients are to be further optimized.

The CI wavefunction

$$= \sum_J C_J \Psi_J$$

is most commonly constructed from CSFs  $\Psi_J$  that include:

1. All of the CSFs in the SCF (in which case only a single CSF is included) or MCSCF wavefunction that was used to generate the molecular orbitals  $\psi_i$ . This set of CSFs are referred to as spanning the 'reference space' of the subsequent CI calculation, and the particular combination of these CSFs used in this orbital optimization (i.e., the SCF or MCSCF wavefunction) is called the reference function.
2. CSFs that are generated by carrying out single, double, triple, etc. level 'excitations' (i.e., orbital replacements) relative to reference CSFs. CI wavefunctions limited to include contributions through various levels of excitation (e.g., single, double, etc.) are denoted S (singly excited), D (doubly), SD (singly and doubly), SDT (singly, doubly, and triply), and so on.

The orbitals from which electrons are removed and those into which electrons are excited can be restricted to focus attention on correlations among certain orbitals. For example, if excitations out of core electrons are excluded, one computes a total energy that contains no correlation corrections for these core orbitals. Often it is possible to so limit the nature of the orbital excitations to focus on the energetic quantities of interest (e.g., the CC bond breaking in ethane requires correlation of the  $\psi_{CC}$  orbital but the 1s Carbon core orbitals and the CH bond orbitals may be treated in a non-correlated manner).

Clearly, the number of CSFs included in the CI calculation can be far in excess of the number considered in typical MCSCF calculations; CI wavefunctions including 5,000 to 50,000 CSFs are routinely used, and functions with one to several million CSFs are within the realm of practicality (see, for example, J. Olsen, B. Roos, Poul Jørgensen, and H. J. Aa. Jensen, J. Chem. Phys. 89, 2185 (1988) and J. Olsen, P. Jørgensen, and J. Simons, Chem. Phys. Letters 169, 463 (1990)).

The need for such large CSF expansions should not come as a surprise once one considers that (i) each electron pair requires at least two CSFs (let us say it requires P of them, on average, a dominant one and P-1 others which are doubly excited) to form

polarized orbital pairs, (ii) there are of the order of  $N(N-1)/2 = X$  electron pairs in an atom or molecule containing  $N$  electrons, and (iii) that the number of terms in the CI wavefunction scales as  $P^X$ . So, for an  $H_2O$  molecule containing ten electrons, there would be  $P^{55}$  terms in the CI expansion. This is  $3.6 \times 10^{16}$  terms if  $P=2$  and  $1.7 \times 10^{26}$  terms if  $P=3$ . Undoubtedly, this is an over estimate of the number of CSFs needed to describe electron correlation in  $H_2O$ , but it demonstrates how rapidly the number of CSFs can grow with the number of electrons in the system.

The  $H_{I,J}$  matrices that arise in CI calculations are evaluated in terms of one- and two- electron integrals over the molecular orbitals using the equivalent of the Slater-Condon rules. For large CI calculations, the full  $H_{I,J}$  matrix is not actually evaluated and stored in the computer's memory (or on its disk); rather, so-called 'direct CI' methods (see the article by Roos and Siegbahn in MTC) are used to compute and immediately sum contributions to the sum  $\sum_I H_{I,J} C_J$  in terms of integrals, density matrix elements, and approximate values of the  $C_J$  amplitudes. Iterative methods (see, for example, E. R. Davidson, *J. Comput. Phys.* **17**, 87 (1975)), in which approximate values for the  $C_J$  coefficients and energy  $E$  are refined through sequential application of  $\sum_I H_{I,J}$  to the preceding estimate of the  $C_J$  vector, are employed to solve these large CI matrix eigenvalue problems.

### C. The MPPT/MBPT Method

In the MPPT/MBPT method, once the reference CSF is chosen and the SCF orbitals belonging to this CSF are determined, the wavefunction and energy  $E$  are determined in an order-by-order manner. This is one of the primary strengths of the MPPT/MBPT technique; it does not require one to make further (potentially arbitrary) choices once the basis set and dominant (SCF) configuration are specified. In contrast to the MCSCF and CI treatments, one need not make choices of CSFs to include in or exclude from . The MPPT/MBPT perturbation equations determine what CSFs must be included through any particular order.

For example, the first-order wavefunction correction  $\psi^1$  (i.e.,  $\psi = \psi^0 + \psi^1$  through first order) is given by:

$$\psi^1 = - \sum_{i < j, m < n} \langle i, j | g | m, n \rangle \langle i, j | g | m, n \rangle^{-1} [ \langle m | h | i \rangle + \langle n | h | j \rangle ]^{-1} | i, j |^{m, n} \rangle$$

$$= - \sum_{i < j, m < n} [ \langle i, j | g | m, n \rangle - \langle i, j | g | n, m \rangle ] [ \langle m | h | i \rangle + \langle n | h | j \rangle ]^{-1} | i, j |^{m, n} \rangle$$

where the SCF orbital energies are denoted  $\epsilon_k$  and  $\phi_{i,j}^{m,n}$  represents a CSF that is doubly excited relative to  $\phi_0$ . Thus, only doubly excited CSFs contribute to the first-order wavefunction; as a result, the energy  $E$  is given through second order as:

$$\begin{aligned}
 E &= \langle \phi_0 | H^0 | \phi_0 \rangle + \langle \phi_0 | H - H^0 | \phi_0 \rangle + \langle \phi_0 | H - H^0 | \phi_1 \rangle \\
 &= \langle \phi_0 | H | \phi_0 \rangle - \sum_{i<j, m<n} \langle \phi_0 | \phi_{i,j}^{m,n} | H - H^0 | \phi_0 \rangle^2 / [\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j] \\
 &= E_{\text{SCF}} - \sum_{i<j, m<n} |\langle \phi_0 | \phi_{i,j}^{m,n} | H - H^0 | \phi_0 \rangle|^2 / [\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j] \\
 &= E^0 + E^1 + E^2.
 \end{aligned}$$

These contributions have been expressed, using the SC rules, in terms of the two-electron integrals  $\langle \phi_0 | \phi_{i,j}^{m,n} | H - H^0 | \phi_0 \rangle$  coupling the excited spin-orbitals to the spin-orbitals from which electrons were excited as well as the orbital energy differences  $[\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j]$  accompanying such excitations. In this form, it becomes clear that major contributions to the correlation energy of the pair of occupied orbitals  $i, j$  are made by double excitations into virtual orbitals  $m, n$  that have large coupling (i.e., large  $\langle \phi_0 | \phi_{i,j}^{m,n} | H - H^0 | \phi_0 \rangle$  integrals) and small orbital energy gaps,  $[\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j]$ .

In higher order corrections to the wavefunction and to the energy, contributions from CSFs that are singly, triply, etc. excited relative to  $\phi_0$  appear, and additional contributions from the doubly excited CSFs also enter. It is relatively common to carry MPPT/MBPT calculations (see the references given above in Chapter 19.I.3 where the contributions of the Pople and Bartlett groups to the development of MPPT/MBPT are documented) through to third order in the energy (whose evaluation can be shown to require only  $\epsilon_0$  and  $\epsilon_1$ ). The entire GAUSSIAN-8X series of programs, which have been used in thousands of important chemical studies, calculate  $E$  through third order in this manner.

In addition to being size-extensive and not requiring one to specify input beyond the basis set and the dominant CSF, the MPPT/MBPT approach is able to include the effect of all CSFs (that contribute to any given order) without having to find any eigenvalues of a matrix. This is an important advantage because matrix eigenvalue determination, which is necessary in MCSCF and CI calculations, requires computer time in proportion to the third power of the dimension of the  $H_{I,J}$  matrix. Despite all of these advantages, it is important to remember the primary disadvantages of the MPPT/MBPT approach; its energy is not an upper bound to the true energy and it may not be able to treat cases for which two or more

CSFs have equal or nearly equal amplitudes because it obtains the amplitudes of all but the dominant CSF from perturbation theory formulas that assume the perturbation is 'small'.

#### D. The Coupled-Cluster Method

The implementation of the CC method begins much as in the MPPT/MBPT case; one selects a reference CSF that is used in the SCF process to generate a set of spin-orbitals to be used in the subsequent correlated calculation. The set of working equations of the CC technique given above in Chapter 19.I.4 can be written explicitly by introducing the form of the so-called cluster operator  $T$ ,

$$T = \sum_{i,m} t_i^m m^+ i + \sum_{i,j,m,n} t_{i,j}^{m,n} m^+ n^+ j i + \dots,$$

where the combination of operators  $m^+ i$  denotes creation of an electron in virtual spin-orbital  $m$  and removal of an electron from occupied spin-orbital  $i$  to generate a single excitation. The operation  $m^+ n^+ j i$  therefore represents a double excitation from  $i j$  to  $m n$ . Expressing the cluster operator  $T$  in terms of the amplitudes  $t_i^m$ ,  $t_{i,j}^{m,n}$ , etc. for singly, doubly, etc. excited CSFs, and expanding the exponential operators in  $\exp(-T) H \exp(T)$  one obtains:

$$\langle i^m | H + [H,T] + 1/2 [[H,T],T] + 1/6 [[[H,T],T],T] + 1/24 [[[[H,T],T],T],T] | \rangle = 0;$$

$$\langle i,j^{m,n} | H + [H,T] + 1/2 [[H,T],T] + 1/6 [[[H,T],T],T] + 1/24 [[[[H,T],T],T],T] | \rangle = 0;$$

$$\langle i,j,k^{m,n,p} | H + [H,T] + 1/2 [[H,T],T] + 1/6 [[[H,T],T],T] + 1/24 [[[[H,T],T],T],T] | \rangle = 0,$$

and so on for higher order excited CSFs. It can be shown, because of the one- and two-electron operator nature of  $H$ , that the expansion of the exponential operators truncates exactly at the fourth power; that is terms such as  $[[[[[H,T],T],T],T],T]$  and higher commutators vanish identically (this is demonstrated in Chapter 4 of Second Quantization Based Methods in Quantum Chemistry, P. Jørgensen and J. Simons, Academic Press, New York (1981).

As a result, the exact CC equations are quartic equations for the  $t_i^m$ ,  $t_{i,j}^{m,n}$ , etc. amplitudes. Although it is a rather formidable task to evaluate all of the commutator matrix elements appearing in the above CC equations, it can be and has been done (the references given above to Purvis and Bartlett are especially relevant in this context). The result is to express each such matrix element, via the Slater-Condon rules, in terms of one- and two-electron integrals over the spin-orbitals used in determining  $\Psi$ , including those in  $\Psi$  itself and the 'virtual' orbitals not in  $\Psi$ .

In general, these quartic equations must then be solved in an iterative manner and are susceptible to convergence difficulties that are similar to those that arise in MCSCF-type calculations. In any such iterative process, it is important to start with an approximation (to the  $t$  amplitudes, in this case) which is reasonably close to the final converged result. Such an approximation is often achieved, for example, by neglecting all of the terms that are non-linear in the  $t$  amplitudes (because these amplitudes are assumed to be less than unity in magnitude). This leads, for the CC working equations obtained by projecting onto the doubly excited CSFs, to:

$$\langle i,j | g | m,n \rangle' + [ \epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j ] t_{i,j}^{m,n} + \sum_{i',j',m',n'} \langle i',j',m',n' | H - H^0 | i,j,m,n \rangle t_{i',j'}^{m',n'} = 0 ,$$

where the notation  $\langle i,j | g | m,n \rangle'$  is used to denote the two-electron integral difference  $\langle i,j | g | m,n \rangle - \langle i,j | g | n,m \rangle$ . If, in addition, the factors that couple different doubly excited CSFs are ignored (i.e., the sum over  $i',j',m',n'$ ), the equations for the  $t$  amplitudes reduce to the equations for the CSF amplitudes of the first-order MPPT/MBPT wavefunction:

$$t_{i,j}^{m,n} = - \langle i,j | g | m,n \rangle' / [ \epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j ] .$$

As Bartlett and Pople have both demonstrated, there is, in fact, close relationship between the MPPT/MBPT and CC methods when the CC equations are solved iteratively starting with such an MPPT/MBPT-like initial 'guess' for these double-excitation amplitudes.

The CC method, as presented here, suffers from the same drawbacks as the MPPT/MBPT approach; its energy is not an upper bound and it may not be able to accurately describe wavefunctions which have two or more CSFs with approximately equal amplitude. Moreover, solution of the non-linear CC equations may be difficult and slowly (if at all) convergent. It has the same advantages as the MPPT/MBPT method; its energy is

size-extensive, it requires no large matrix eigenvalue solution, and its energy and wavefunction are determined once one specifies the basis and the dominant CSF.

### E. Density Functional Methods

These approaches provide alternatives to the conventional tools of quantum chemistry. The CI, MCSCF, MPPT/MBPT, and CC methods move beyond the single-configuration picture by adding to the wave function more configurations whose amplitudes they each determine in their own way. This can lead to a very large number of CSFs in the correlated wave function, and, as a result, a need for extraordinary computer resources.

The density functional approaches are different. Here one solves a set of orbital-level equations

$$\left[ -\frac{\hbar^2}{2m_e} \nabla^2 - \sum_A \frac{Z_A e^2}{|\mathbf{r} - \mathbf{R}_A|} + \int (\mathbf{r}') \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + U(\mathbf{r}) \right] \psi_i = \epsilon_i \psi_i$$

in which the orbitals  $\{\psi_i\}$  'feel' potentials due to the nuclear centers (having charges  $Z_A$ ), Coulombic interaction with the total electron density  $\rho(\mathbf{r}')$ , and a so-called exchange-correlation potential denoted  $U(\mathbf{r}')$ . The particular electronic state for which the calculation is being performed is specified by forming a corresponding density  $\rho(\mathbf{r}')$ . Before going further in describing how DFT calculations are carried out, let us examine the origins underlying this theory.

The so-called Hohenberg-Kohn theorem states that the ground-state electron density  $\rho(\mathbf{r})$  describing an N-electron system uniquely determines the potential  $V(\mathbf{r})$  in the Hamiltonian

$$H = \sum_j \left\{ -\frac{\hbar^2}{2m_e} \nabla_j^2 + V(\mathbf{r}_j) + \frac{e^2}{2} \sum_{k \neq j} \frac{1}{r_{j,k}} \right\},$$

and, because  $H$  determines the ground-state energy and wave function of the system, the ground-state density  $\rho(\mathbf{r})$  determines the ground-state properties of the system. The proof of this theorem proceeds as follows:

- a.  $\rho(\mathbf{r})$  determines  $N$  because  $\int \rho(\mathbf{r}) d^3r = N$ .
- b. Assume that there are two distinct potentials (aside from an additive constant that simply shifts the zero of total energy)  $V(\mathbf{r})$  and  $V'(\mathbf{r})$  which, when used in  $H$  and  $H'$ , respectively, to solve for a ground state produce  $E_0, \psi_0(\mathbf{r})$  and  $E_0', \psi_0'(\mathbf{r})$  that have the same one-electron density:  $|\psi_0|^2 dr_2 dr_3 \dots dr_N = \rho(\mathbf{r}) = |\psi_0'|^2 dr_2 dr_3 \dots dr_N$ .
- c. If we think of  $\psi_0'$  as trial variational wave function for the Hamiltonian  $H$ , we know that  $E_0 < \langle \psi_0' | H | \psi_0' \rangle = \langle \psi_0' | H' | \psi_0' \rangle + \int \rho(\mathbf{r}) [V(\mathbf{r}) - V'(\mathbf{r})] d^3r = E_0' + \int \rho(\mathbf{r}) [V(\mathbf{r}) - V'(\mathbf{r})] d^3r$ .
- d. Similarly, taking  $\psi_0$  as a trial function for the  $H'$  Hamiltonian, one finds that  $E_0' < E_0 + \int \rho(\mathbf{r}) [V'(\mathbf{r}) - V(\mathbf{r})] d^3r$ .
- e. Adding the equations in c and d gives  $E_0 + E_0' < E_0 + E_0'$ , a clear contradiction.

Hence, there cannot be two distinct potentials  $V$  and  $V'$  that give the same ground-state  $\rho(\mathbf{r})$ . So, the ground-state density  $\rho(\mathbf{r})$  uniquely determines  $N$  and  $V$ , and thus  $H$ , and therefore  $E_0$  and  $\psi_0$ . Furthermore, because  $\rho(\mathbf{r})$  determines all properties of the ground state, then  $\rho(\mathbf{r})$ , in principle, determines all such properties. This means that even the kinetic energy and the electron-electron interaction energy of the ground-state are determined by  $\rho(\mathbf{r})$ . It is easy to see that  $\int \rho(\mathbf{r}) V(\mathbf{r}) d^3r = V[\rho]$  gives the average value of the electron-nuclear (plus any additional one-electron additive potential) interaction in terms of the ground-state density  $\rho(\mathbf{r})$ , but how are the kinetic energy  $T[\rho]$  and the electron-electron interaction  $V_{ee}[\rho]$  energy expressed in terms of  $\rho$ ?

The main difficulty with DFT is that the Hohenberg-Kohn theorem shows that the ground-state values of  $T, V_{ee}, V$ , etc. are all unique functionals of the ground-state (i.e.,

that they can, in principle, be determined once  $\rho$  is given), but it does not tell us what these functional relations are.

To see how it might make sense that a property such as the kinetic energy, whose operator  $-\hbar^2/2m_e \nabla^2$  involves derivatives, can be related to the electron density, consider a simple system of  $N$  non-interacting electrons moving in a three-dimensional cubic “box” potential. The energy states of such electrons are known to be

$$E = (\hbar^2/2m_e L^2) (n_x^2 + n_y^2 + n_z^2),$$

where  $L$  is the length of the box along the three axes, and  $n_x$ ,  $n_y$ , and  $n_z$  are the quantum numbers describing the state. We can view  $n_x^2 + n_y^2 + n_z^2 = R^2$  as defining the squared radius of a sphere in three dimensions, and we realize that the density of quantum states in this space is one state per unit volume in the  $n_x$ ,  $n_y$ ,  $n_z$  space. Because  $n_x$ ,  $n_y$ , and  $n_z$  must be positive integers, the volume covering all states with energy less than or equal to a specified energy  $E = (\hbar^2/2m_e L^2) R^2$  is 1/8 the volume of the sphere of radius  $R$ :

$$V(E) = 1/8 (4\pi/3) R^3 = (\pi/6) (8m_e L^2 E/\hbar^2)^{3/2}.$$

Since there is one state per unit of such volume,  $V(E)$  is also the number of states with energy less than or equal to  $E$ , and is called the integrated density of states. The number of states  $g(E) dE$  with energy between  $E$  and  $E+dE$ , the density of states, is the derivative of  $V(E)$ :

$$g(E) = dV/dE = (\pi/4) (8m_e L^2/\hbar^2)^{3/2} E^{1/2}.$$

If we calculate the total energy for  $N$  electrons, with the states having energies up to the so-called Fermi energy (i.e., the energy of the highest occupied molecular orbital HOMO) doubly occupied, we obtain the ground-state energy:

$$E_0 = 2 \int_0^{E_F} g(E) E dE = (8\pi/5) (2m_e/\hbar^2)^{3/2} L^3 E_F^{5/2}.$$

The total number of electrons  $N$  can be expressed as

$$N = 2 \int_0^{E_F} g(E) dE = (8/3) (2m_e/h^2)^{3/2} L^3 E_F^{3/2},$$

which can be solved for  $E_F$  in terms of  $N$  to then express  $E_0$  in terms of  $N$  instead of  $E_F$ :

$$E_0 = (3h^2/10m_e) (3/8)^{2/3} L^3 (N/L^3)^{5/3}.$$

This gives the total energy, which is also the kinetic energy in this case because the potential energy is zero within the “box”, in terms of the electron density  $n(x,y,z) = (N/L^3)$ . It therefore may be plausible to express kinetic energies in terms of electron densities  $n(\mathbf{r})$ , but it is by no means clear how to do so for “real” atoms and molecules with electron-nuclear and electron-electron interactions operative.

In one of the earliest DFT models, the Thomas-Fermi theory, the kinetic energy of an atom or molecule is approximated using the above kind of treatment on a “local” level. That is, for each volume element in  $\mathbf{r}$  space, one assumes the expression given above to be valid, and then one integrates over all  $\mathbf{r}$  to compute the total kinetic energy:

$$T_{TF}[n] = (3h^2/10m_e) (3/8)^{2/3} \int [n(\mathbf{r})]^{5/3} d^3r = C_F \int [n(\mathbf{r})]^{5/3} d^3r,$$

where the last equality simply defines the  $C_F$  constant (which is 2.8712 in atomic units). Ignoring the correlation and exchange contributions to the total energy, this  $T$  is combined with the electron-nuclear  $V$  and Coulombic electron-electron potential energies to give the Thomas-Fermi total energy:

$$E_{0,TF}[n] = C_F \int [n(\mathbf{r})]^{5/3} d^3r + \int V(\mathbf{r}) n(\mathbf{r}) d^3r + e^2/2 \int \int n(\mathbf{r}) n(\mathbf{r}')/|\mathbf{r}-\mathbf{r}'| d^3r d^3r',$$

This expression is an example of how  $E_0$  is given as a local density functional approximation (LDA). The term local means that the energy is given as a functional (i.e., a function of  $n$ ) which depends only on  $n(\mathbf{r})$  at points in space but not on  $n(\mathbf{r})$  at more than one point in space.

Unfortunately, the Thomas-Fermi energy functional does not produce results that are of sufficiently high accuracy to be of great use in chemistry. What is missing in this

theory are a. the exchange energy and b. the correlation energy; moreover, the kinetic energy is treated only in the approximate manner described.

In the book by Parr and Yang, it is shown how Dirac was able to address the exchange energy for the 'uniform electron gas' (N Coulomb interacting electrons moving in a uniform positive background charge whose magnitude balances the charge of the N electrons). If the exact expression for the exchange energy of the uniform electron gas is applied on a local level, one obtains the commonly used Dirac local density approximation to the exchange energy:

$$E_{\text{ex,Dirac}}[\rho] = -C_x \int \rho(\mathbf{r})^{4/3} d^3r,$$

with  $C_x = (3/4) (3/\pi)^{1/3} = 0.7386$  in atomic units. Adding this exchange energy to the Thomas-Fermi total energy  $E_{0,\text{TF}}[\rho]$  gives the so-called Thomas-Fermi-Dirac (TFD) energy functional.

Because electron densities vary rather strongly spatially near the nuclei, corrections to the above approximations to  $T[\rho]$  and  $E_{\text{ex,Dirac}}$  are needed. One of the more commonly used so-called gradient-corrected approximations is that invented by Becke, and referred to as the Becke88 exchange functional:

$$E_{\text{ex}}(\text{Becke88}) = E_{\text{ex,Dirac}}[\rho] - \int x^2 \rho^{4/3} (1 + 6 \rho^{-1/3} |\nabla \rho|^2)^{-1} d\mathbf{r},$$

where  $x = \alpha |\nabla \rho| / \rho^{4/3}$ , and  $\alpha$  is a parameter chosen so that the above exchange energy can best reproduce the known exchange energies of specific electronic states of the inert gas atoms (Becke finds  $\alpha$  to equal 0.0042). A common gradient correction to the earlier  $T[\rho]$  is called the Weizsacker correction and is given by

$$T_{\text{Weizsacker}} = (1/72)(\hbar/m_e) \int |\nabla \rho|^2 / \rho d\mathbf{r}.$$

Although the above discussion suggests how one might compute the ground-state energy once the ground-state density  $\rho(\mathbf{r})$  is given, one still needs to know how to obtain

. Kohn and Sham (KS) introduced a set of so-called KS orbitals obeying the following equation:

$$\left\{ -\frac{1}{2} \nabla^2 + V(\mathbf{r}) + \frac{e^2}{2} \int \frac{|\mathbf{r}'|}{|\mathbf{r}-\mathbf{r}'|} |\mathbf{d}\mathbf{r}' + U_{xc}(\mathbf{r}) \right\} \psi_j = \epsilon_j \psi_j,$$

where the so-called exchange-correlation potential  $U_{xc}(\mathbf{r}) = \delta E_{xc}[\rho] / \delta \rho(\mathbf{r})$  could be obtained by functional differentiation if the exchange-correlation energy functional  $E_{xc}[\rho]$  were known. KS also showed that the KS orbitals  $\{\psi_j\}$  could be used to compute the density by simply adding up the orbital densities multiplied by orbital occupancies  $n_j$ :

$$\rho(\mathbf{r}) = \sum_j n_j |\psi_j(\mathbf{r})|^2.$$

(here  $n_j = 0, 1, \text{ or } 2$  is the occupation number of the orbital  $\psi_j$  in the state being studied) and that the kinetic energy should be calculated as

$$T = \sum_j n_j \int |\psi_j(\mathbf{r})|^{-1/2} \nabla^2 |\psi_j(\mathbf{r})|^{1/2} d\mathbf{r}.$$

The same investigations of the idealized 'uniform electron gas' that identified the Dirac exchange functional, found that the correlation energy (per electron) could also be written exactly as a function of the electron density of the system, but only in two limiting cases- the high-density limit (large  $\rho$ ) and the low-density limit. There still exists no exact expression for the correlation energy even for the uniform electron gas that is valid at arbitrary values of  $\rho$ . Therefore, much work has been devoted to creating efficient and accurate interpolation formulas connecting the low- and high- density uniform electron gas expressions. One such expression is

$$E_c[\rho] = \int \epsilon_c(\rho(\mathbf{r})) d\mathbf{r},$$

where

$$\epsilon_c(\rho) = A/2 \left\{ \ln(x/X) + 2b/Q \tan^{-1}(Q/(2x+b)) - bx_0/X_0 \left[ \ln((x-x_0)^2/X) \right] \right\}$$

$$+2(b+2x_0)/Q \tan^{-1}(Q/(2x+b))]$$

is the correlation energy per electron. Here  $x = r_s^{1/2}$ ,  $X = x^2 + bx + c$ ,  $X_0 = x_0^2 + bx_0 + c$  and  $Q = (4c - b^2)^{1/2}$ ,  $A = 0.0621814$ ,  $x_0 = -0.409286$ ,  $b = 13.0720$ , and  $c = 42.7198$ . The parameter  $r_s$  is how the density enters since  $4/3 r_s^3$  is equal to  $1/\rho$ ; that is,  $r_s$  is the radius of a sphere whose volume is the effective volume occupied by one electron. A reasonable approximation to the full  $E_{xc}[\rho]$  would contain the Dirac (and perhaps gradient corrected) exchange functional plus the above  $E_c[\rho]$ , but there are many alternative approximations to the exchange-correlation energy functional. Currently, many workers are doing their best to “cook up” functionals for the correlation and exchange energies, but no one has yet invented functionals that are so reliable that most workers agree to use them.

To summarize, in implementing any DFT, one usually proceeds as follows:

1. An atomic orbital basis is chosen in terms of which the KS orbitals are to be expanded.
2. Some initial guess is made for the LCAO-KS expansion coefficients  $C_{j,a}$ :  $\psi_j = \sum_a C_{j,a} \phi_a$ .
3. The density is computed as  $\rho(\mathbf{r}) = \sum_j n_j |\psi_j(\mathbf{r})|^2$ . Often,  $\rho(\mathbf{r})$  is expanded in an atomic orbital basis, which need not be the same as the basis used for the  $\psi_j$ , and the expansion coefficients of  $\rho$  are computed in terms of those of the  $\psi_j$ . It is also common to use an atomic orbital basis to expand  $\rho^{1/3}(\mathbf{r})$  which, together with  $\rho$ , is needed to evaluate the exchange-correlation functional's contribution to  $E_0$ .
4. The current iteration's density is used in the KS equations to determine the Hamiltonian  $\{-1/2 \nabla^2 + V(\mathbf{r}) + e^2/2 \sum_{j'} \rho(\mathbf{r}')/|\mathbf{r}-\mathbf{r}'| - \mu\} \psi_j + U_{xc}(\mathbf{r}) \psi_j = \epsilon_j \psi_j$  whose “new” eigenfunctions  $\{\psi_j\}$  and eigenvalues  $\{\epsilon_j\}$  are found by solving the KS equations.
5. These new  $\psi_j$  are used to compute a new density, which, in turn, is used to solve a new set of KS equations. This process is continued until convergence is reached (i.e., until the  $\psi_j$  used to determine the current iteration's  $\rho$  are the same  $\psi_j$  that arise as solutions on the next iteration).
6. Once the converged  $\rho(\mathbf{r})$  is determined, the energy can be computed using the earlier expression

$$E[\psi] = \int \psi^*(\mathbf{r}) \left[ -\frac{1}{2} \nabla^2 + V(\mathbf{r}) \right] \psi(\mathbf{r}) d\mathbf{r} + \frac{e^2}{2} \iint \frac{|\psi(\mathbf{r})|^2 |\psi(\mathbf{r}')|^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E_{xc}[\psi].$$

In closing this section, it should once again be emphasized that this area is currently undergoing explosive growth and much scrutiny. As a result, it is nearly certain that many of the specific functionals discussed above will be replaced in the near future by improved and more rigorously justified versions. It is also likely that extensions of DFT to excited states (many workers are actively pursuing this) will be placed on more solid ground and made applicable to molecular systems. Because the computational effort involved in these approaches scales much less strongly with basis set size than for conventional (SCF, MCSCF, CI, etc.) methods, density functional methods offer great promise and are likely to contribute much to quantum chemistry in the next decade.

## Chapter 19

*Corrections to the mean-field model are needed to describe the instantaneous Coulombic interactions among the electrons. This is achieved by including more than one Slater determinant in the wavefunction.*

Much of the development of the previous chapter pertains to the use of a single Slater determinant trial wavefunction. As presented, it relates to what has been called the unrestricted Hartree-Fock (UHF) theory in which each spin-orbital  $\phi_i$  has its own orbital energy  $\epsilon_i$  and LCAO-MO coefficients  $C_{\mu i}$ ; there may be different  $C_{\mu i}$  for  $\alpha$  spin-orbitals than for  $\beta$  spin-orbitals. Such a wavefunction suffers from the spin contamination difficulty detailed earlier.

To allow for a properly spin- and space- symmetry adapted trial wavefunction and to permit  $\Psi$  to contain more than a single CSF, methods which are more flexible than the single-determinant HF procedure are needed. In particular, it may be necessary to use a combination of determinants to describe such a proper symmetry function. Moreover, as emphasized earlier, whenever two or more CSFs have similar energies (i.e., Hamiltonian expectation values) and can couple strongly through the Hamiltonian (e.g., at avoided crossings in configuration correlation diagrams), the wavefunction must be described in a multiconfigurational manner to permit the wavefunction to evolve smoothly from reactants to products. Also, whenever dynamical electron correlation effects are to be treated, a multiconfigurational  $\Psi$  must be used; in this case, CSFs that are doubly excited relative to one or more of the essential CSFs (i.e., the dominant CSFs that are included in the so-called reference wavefunction) are included to permit polarized-orbital-pair formation.

Multiconfigurational functions are needed not only to account for electron correlation but also to permit orbital readjustments to occur. For example, if a set of SCF orbitals is employed in forming a multi-CSF wavefunction, the variational condition that the energy is stationary with respect to variations in the LCAO-MO coefficients is no longer obeyed (i.e., the SCF energy functional is stationary when SCF orbitals are employed, but the MC-energy functional is generally not stationary if SCF orbitals are employed). For such reasons, it is important to include CSFs that are singly excited relative to the dominant CSFs in the reference wavefunction.

That singly excited CSFs allow for orbital relaxation can be seen as follows. Consider a wavefunction consisting of one CSF  $|\phi_1 \dots \phi_i \dots \phi_N\rangle$  to which singly excited CSFs of the form  $|\phi_1 \dots \phi_m \dots \phi_N\rangle$  have been added with coefficients  $C_{i,m}$ :

$$\Psi = \sum_m C_{i,m} |\phi_1 \dots \phi_m \dots \phi_N\rangle + |\phi_1 \dots \phi_i \dots \phi_N\rangle.$$

All of these determinants have all of their columns equal except the  $i^{\text{th}}$  column; therefore, they can be combined into a single new determinant:

$$= | 1 \dots i' \dots N |,$$

where the relaxed orbital  $i'$  is given by

$$i' = i + \sum_m C_{i,m} \phi_m.$$

The sum of CSFs that are singly excited in the  $i^{\text{th}}$  spin-orbital with respect to  $| 1 \dots i \dots N |$  is therefore seen to allow the spin-orbital  $i$  to relax into the new spin-orbital  $i'$ . It is in this sense that singly excited CSFs allow for orbital reoptimization.

In summary, doubly excited CSFs are often employed to permit polarized orbital pair formation and hence to allow for electron correlations. Singly excited CSFs are included to permit orbital relaxation (i.e., orbital reoptimization) to occur.

## I. Different Methods

There are numerous procedures currently in use for determining the 'best' wavefunction of the form:

$$= \sum_I C_I \Phi_I,$$

where  $\Phi_I$  is a spin-and space- symmetry adapted CSF consisting of determinants of the form  $| \Phi_1 \Phi_2 \Phi_3 \dots \Phi_N |$ . Excellent overviews of many of these methods are included in Modern Theoretical Chemistry Vols. 3 and 4, H. F. Schaefer, III Ed., Plenum Press, New York (1977) and in Advances in Chemical Physics, Vols. LXVII and LXIX, K. P. Lawley, Ed., Wiley-Interscience, New York (1987). Within the present Chapter, these two key references will be denoted MTC, Vols. 3 and 4, and ACP, Vols. 67 and 69, respectively.

In all such trial wavefunctions, there are two fundamentally different kinds of parameters that need to be determined- the CI coefficients  $C_I$  and the LCAO-MO coefficients describing the  $\Phi_k$ . The most commonly employed methods used to determine these parameters include:

1. The **multiconfigurational self-consistent field** (MCSCF) method in which the expectation value  $\langle \Psi | H | \Psi \rangle / \langle \Psi | \Psi \rangle$  is treated variationally and simultaneously made stationary with respect to variations in the  $C_I$  and  $C_{\mu,i}$  coefficients subject to the constraints that the spin-orbitals and the full N-electron wavefunction remain normalized:

$$\langle \Psi | \Psi \rangle = \sum_{i,j} C_i C_j S_{\mu,i} C_{\mu,i}, \text{ and}$$

$$\sum_I C_I^2 = 1.$$

The articles by H.-J. Werner and by R. Shepard in ACP Vol. 69 provide up to date reviews of the status of this approach. The article by A. C. Wahl and G. Das in MTC Vol. 3 covers the 'earlier' history on this topic. F. W. Bobrowicz and W. A. Goddard, III provide, in MTC Vol. 3, an overview of the GVB approach, which, as discussed in Chapter 12, can be viewed as a specific kind of MCSCF calculation.

2. The **configuration interaction** (CI) method in which the LCAO-MO coefficients are determined first (and independently) via either a single-configuration SCF calculation or an MCSCF calculation using a small number of CSFs. The CI coefficients are subsequently determined by making the expectation value  $\langle \Psi | H | \Psi \rangle / \langle \Psi | \Psi \rangle$  stationary with respect to variations in the  $C_I$  only. In this process, the optimizations of the orbitals and of the CSF amplitudes are done in separate steps. The articles by I. Shavitt and by B. O. Ross and P. E. M. Siegbahn in MTC, Vol. 3 give excellent early overviews of the CI method.

3. The **Møller-Plesset perturbation method** (MPPT) uses the single-configuration SCF process (usually the UHF implementation) to first determine a set of LCAO-MO coefficients and, hence, a set of orbitals that obey  $F_i = \epsilon_i \phi_i$ . Then, using an unperturbed Hamiltonian equal to the sum of these Fock operators for each of the N electrons  $H^0 = \sum_{i=1,N} F(i)$ , perturbation theory (see Appendix D for an introduction to time-independent perturbation theory) is used to determine the  $C_I$  amplitudes for the CSFs. The MPPT procedure is also referred to as the many-body perturbation theory (MBPT) method. The two names arose because two different schools of physics and chemistry developed them for somewhat different applications. Later, workers realized that they were identical in their working equations when the UHF  $H^0$  is employed as the unperturbed Hamiltonian. In this text, we will therefore refer to this approach as MPPT/MBPT.

The amplitude for the so-called reference CSF used in the SCF process is taken as unity and the other CSFs' amplitudes are determined, relative to this one, by Rayleigh-Schrödinger perturbation theory using the full N-electron Hamiltonian minus the sum of Fock operators  $H-H^0$  as the perturbation. The Slater-Condon rules are used for evaluating matrix elements of  $(H-H^0)$  among these CSFs. The essential features of the MPPT/MBPT approach are described in the following articles: J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, *Int. J. Quantum Chem.* 14, 545 (1978); R. J. Bartlett and D. M. Silver, *J. Chem. Phys.* 62, 3258 (1975); R. Krishnan and J. A. Pople, *Int. J. Quantum Chem.* 14, 91 (1978).

4. The **Coupled-Cluster method** expresses the CI part of the wavefunction in a somewhat different manner (the early work in chemistry on this method is described in J. Cizek, *J. Chem. Phys.* 45, 4256 (1966); J. Paldus, J. Cizek, and I. Shavitt, *Phys. Rev.* A5, 50 (1972); R. J. Bartlett and G. D. Purvis, *Int. J. Quantum Chem.* 14, 561 (1978); G. D. Purvis and R. J. Bartlett, *J. Chem. Phys.* 76, 1910 (1982)):

$$= \exp(T) \quad ,$$

where  $\Phi_0$  is a single CSF (usually the UHF single determinant) which has been used to independently determine a set of spin-orbitals and LCAO-MO coefficients via the SCF process. The operator T generates, when acting on  $\Phi_0$ , single, double, etc. 'excitations' (i.e., CSFs in which one, two, etc. of the occupied spin-orbitals in  $\Phi_0$  have been replaced by virtual spin-orbitals). T is commonly expressed in terms of operators that effect such spin-orbital removals and additions as follows:

$$T = \sum_{i,m} t_i^m m^+ i + \sum_{i,j,m,n} t_{i,j}^{m,n} m^+ n^+ j i + \dots,$$

where the operator  $m^+$  is used to denote creation of an electron in virtual spin-orbital  $m$  and the operator  $j$  is used to denote removal of an electron from occupied spin-orbital  $j$ .

The  $t_i^m$ ,  $t_{i,j}^{m,n}$ , etc. amplitudes, which play the role of the CI coefficients in CC theory, are determined through the set of equations generated by projecting the Schrödinger equation in the form

$$\exp(-T) H \exp(T) \Phi_0 = E \Phi_0$$

against CSFs which are single, double, etc. excitations relative to  $|\Phi\rangle$ . For example, for double excitations  $|\Phi_{i,j}^{m,n}\rangle$  the equations read:

$$\langle \Phi_{i,j}^{m,n} | \exp(-T) H \exp(T) | \Phi \rangle = E \langle \Phi_{i,j}^{m,n} | \Phi \rangle = 0;$$

zero is obtained on the right hand side because the excited CSFs  $|\Phi_{i,j}^{m,n}\rangle$  are orthogonal to the reference function  $|\Phi\rangle$ . The elements on the left hand side of the CC equations can be expressed, as described below, in terms of one- and two-electron integrals over the spin-orbitals used in forming the reference and excited CSFs.

### A. Integral Transformations

All of the above methods require the evaluation of one- and two-electron integrals over the  $N$  atomic orbital basis:  $\langle a | f | b \rangle$  and  $\langle a b | g | c d \rangle$ . Eventually, all of these methods provide their working equations and energy expressions in terms of one- and two-electron integrals over the  $N$  final molecular orbitals:  $\langle i | f | j \rangle$  and  $\langle i j | g | k l \rangle$ .

The mo-based integrals can only be evaluated by transforming the AO-based integrals as follows:

$$\langle i j | g | k l \rangle = \sum_{a,b,c,d} C_{a,i} C_{b,j} C_{c,k} C_{d,l} \langle a b | g | c d \rangle,$$

and

$$\langle i | f | j \rangle = \sum_{a,b} C_{a,i} C_{b,j} \langle a | f | b \rangle.$$

It would seem that the process of evaluating all  $N^4$  of the  $\langle i j | g | k l \rangle$ , each of which requires  $N^4$  additions and multiplications, would require computer time proportional to  $N^8$ . However, it is possible to perform the full transformation of the two-electron integral list in a time that scales as  $N^5$ . This is done by first performing a transformation of the  $\langle a b | g | c d \rangle$  to an intermediate array labeled  $\langle a b | g | c l \rangle$  as follows:

$$\langle a b | g | c l \rangle = \sum_d C_{d,l} \langle a b | g | c d \rangle.$$

This partial transformation requires  $N^5$  multiplications and additions.

The list  $\langle a b | g | c l \rangle$  is then transformed to a second-level transformed array

$$\langle a b | g | k l \rangle:$$

$$\langle a | b | g | k | l \rangle = \sum_c C_{c,k} \langle a | b | g | c | l \rangle,$$

which requires another  $N^5$  operations. This sequential, one-index-at-a-time transformation is repeated four times until the final  $\langle i | j | g | k | l \rangle$  array is in hand. The entire transformation done this way requires  $4N^5$  multiplications and additions.

Once the requisite one- and two-electron integrals are available in the molecular orbital basis, the multiconfigurational wavefunction and energy calculation can begin. These transformations consume a large fraction of the computer time used in most such calculations, and represent a severe bottleneck to progress in applying ab initio electronic structure methods to larger systems.

## B. Configuration List Choices

Once the requisite one- and two-electron integrals are available in the molecular orbital basis, the multiconfigurational wavefunction and energy calculation can begin. Each of these methods has its own approach to describing the configurations  $\{ J \}$  included in the calculation and how the  $\{ C_J \}$  amplitudes and the total energy  $E$  is to be determined.

The number of configurations ( $N_C$ ) varies greatly among the methods and is an important factor to keep in mind when planning to carry out an *ab initio* calculation. Under certain circumstances (e.g., when studying Woodward-Hoffmann forbidden reactions where an avoided crossing of two configurations produces an activation barrier), it may be essential to use more than one electronic configuration. Sometimes, one configuration (e.g., the SCF model) is adequate to capture the qualitative essence of the electronic structure. In all cases, many configurations will be needed if highly accurate treatment of electron-electron correlations are desired.

The value of  $N_C$  determines how much computer time and memory is needed to solve the  $N_C$ -dimensional  $\sum_J H_{I,J} C_J = E C_I$  secular problem in the CI and MCSCF methods. Solution of these matrix eigenvalue equations requires computer time that scales as  $N_C^2$  (if few eigenvalues are computed) to  $N_C^3$  (if most eigenvalues are obtained).

So-called complete-active-space (CAS) methods form all CSFs that can be created by distributing  $N$  valence electrons among  $P$  valence orbitals. For example, the eight non-core electrons of  $H_2O$  might be distributed, in a manner that gives  $M_S = 0$ , among six valence orbitals (e.g., two lone-pair orbitals, two OH bonding orbitals, and two OH\* antibonding orbitals). The number of configurations thereby created is 225. If the same eight electrons were distributed among ten valence orbitals 44,100 configurations results;

for twenty and thirty valence orbitals, 23,474,025 and 751,034,025 configurations arise, respectively. Clearly, practical considerations dictate that CAS-based approaches be limited to situations in which a few electrons are to be correlated using a few valence orbitals. The primary advantage of CAS configurations is discussed below in Sec. II. C.

## II. Strengths and Weaknesses of Various Methods

### A. Variational Methods Such as MCSCF, SCF, and CI Produce Energies that are Upper Bounds, but These Energies are not Size-Extensive

Methods that are based on making the energy functional  $\langle \Psi | H | \Psi \rangle / \langle \Psi | \Psi \rangle$  stationary (i.e., variational methods) yield upper bounds to the lowest energy of the symmetry which characterizes the CSFs which comprise  $\Psi$ . These methods also can provide approximate excited-state energies and wavefunctions (e. g., in the form of other solutions of the secular equation  $\sum_J H_{IJ} C_J = E C_I$  that arises in the CI and MCSCF methods). Excited-state energies obtained in this manner can be shown to 'bracket' the true energies of the given symmetry in that between any two approximate energies obtained in the variational calculation, there exists at least one true eigenvalue. This characteristic is commonly referred to as the 'bracketing theorem' (E. A. Hylleraas and B. Undheim, *Z. Phys.* **65**, 759 (1930); J. K. L. MacDonald, *Phys. Rev.* **43**, 830 (1933)). These are strong attributes of the variational methods, as is the long and rich history of developments of analytical and computational tools for efficiently implementing such methods (see the discussions of the CI and MCSCF methods in MTC and ACP).

However, all variational techniques suffer from at least one serious drawback; they are not **size-extensive** (J. A. Pople, pg. 51 in *Energy, Structure, and Reactivity*, D. W. Smith and W. B. McRae, Eds., Wiley, New York (1973)). This means that the energy computed using these tools can not be trusted to scale with the size of the system. For example, a calculation performed on two CH<sub>3</sub> species at large separation may not yield an energy equal to twice the energy obtained by performing the same kind of calculation on a single CH<sub>3</sub> species. Lack of size-extensivity precludes these methods from use in extended systems (e.g., solids) where errors due to improper scaling of the energy with the number of molecules produce nonsensical results.

By carefully adjusting the kind of variational wavefunction used, it is possible to circumvent size-extensivity problems for selected species. For example, a CI calculation on Be<sub>2</sub> using all <sup>1</sup>g CSFs that can be formed by placing the four valence electrons into the orbitals 2<sub>g</sub>, 2<sub>u</sub>, 3<sub>g</sub>, 3<sub>u</sub>, 1<sub>u</sub>, and 1<sub>g</sub> can yield an energy equal to twice that of the Be

atom described by CSFs in which the two valence electrons of the Be atom are placed into the 2s and 2p orbitals in all ways consistent with a  $^1S$  symmetry. Such special choices of configurations give rise to what are called complete-active-space (CAS) MCSCF or CI calculations (see the article by B. O. Roos in ACP for an overview of this approach).

Let us consider an example to understand why the CAS choice of configurations works. The  $^1S$  ground state of the Be atom is known to form a wavefunction that is a strong mixture of CSFs that arise from the  $2s^2$  and  $2p^2$  configurations:

$$|Be\rangle = C_1 |1s^2 2s^2\rangle + C_2 |1s^2 2p^2\rangle,$$

where the latter CSF is a short-hand representation for the proper spin- and space-symmetry adapted CSF

$$|1s^2 2p^2\rangle = \frac{1}{\sqrt{3}} [ |1s \uparrow 1s \uparrow 2p_0 \uparrow 2p_0 \uparrow\rangle - |1s \uparrow 1s \uparrow 2p_1 \uparrow 2p_{-1} \uparrow\rangle - |1s \uparrow 1s \uparrow 2p_{-1} \uparrow 2p_1 \uparrow\rangle ].$$

The reason the CAS process works is that the  $Be_2$  CAS wavefunction has the flexibility to dissociate into the product of two CAS Be wavefunctions:

$$|Be_2\rangle = |Be_a\rangle |Be_b\rangle = \{C_1 |1s^2 2s^2\rangle + C_2 |1s^2 2p^2\rangle\}_a \{C_1 |1s^2 2s^2\rangle + C_2 |1s^2 2p^2\rangle\}_b,$$

where the subscripts a and b label the two Be atoms, because the four electron CAS function distributes the four electrons in all ways among the  $2s_a$ ,  $2s_b$ ,  $2p_a$ , and  $2p_b$  orbitals. In contrast, if the  $Be_2$  calculation had been carried out using only the following CSFs :  $|1 \uparrow 2_g 1 \uparrow 2_u 2 \uparrow 2_g 2 \uparrow 2_u\rangle$  and all single and double excitations relative to this (dominant) CSF, which is a very common type of CI procedure to follow, the  $Be_2$  wavefunction would not have contained the particular CSFs  $|1s^2 2p^2\rangle_a |1s^2 2p^2\rangle_b$  because these CSFs are four-fold excited relative to the  $|1 \uparrow 2_g 1 \uparrow 2_u 2 \uparrow 2_g 2 \uparrow 2_u\rangle$  'reference' CSF.

In general, one finds that if the 'monomer' uses CSFs that are K-fold excited relative to its dominant CSF to achieve an accurate description of its electron correlation, a size-extensive variational calculation on the 'dimer' will require the inclusion of CSFs that are 2K-fold excited relative to the dimer's dominant CSF. To perform a size-extensive

variational calculation on a species containing  $M$  monomers therefore requires the inclusion of CSFs that are  $M \times K$ -fold excited relative to the  $M$ -mer's dominant CSF.

## B. Non-Variational Methods Such as MPPT/MBPT and CC do not Produce Upper Bounds, but Yield Size-Extensive Energies

In contrast to variational methods, perturbation theory and coupled-cluster methods achieve their energies from a 'transition formula'  $\langle \Psi | H | \Phi \rangle$  rather than from an expectation value

$\langle \Psi | H | \Psi \rangle$ . It can be shown (H. P. Kelly, Phys. Rev. 131, 684 (1963)) that this difference allows non-variational techniques to yield size-extensive energies. This can be seen in the MPPT/MBPT case by considering the energy of two non-interacting Be atoms. The reference CSF is  $\Phi = | 1s_a^2 2s_a^2 1s_b^2 2s_b^2 |$ ; the Slater-Condon rules limit the CSFs in which can contribute to

$$E = \langle \Psi | H | \Psi \rangle = \langle \Psi | H | \sum_J C_J \Phi_J \rangle,$$

to be  $\Phi$  itself and those CSFs that are singly or doubly excited relative to  $\Phi$ . These 'excitations' can involve atom a, atom b, or both atoms. However, any CSFs that involve excitations on both atoms

( e.g.,  $| 1s_a^2 2s_a 2p_a 1s_b^2 2s_b 2p_b |$  ) give rise, via the SC rules, to one- and two- electron integrals over orbitals on both atoms; these integrals ( e.g.,  $\langle 2s_a 2p_a | g | 2s_b 2p_b \rangle$  ) vanish if the atoms are far apart, as a result of which the contributions due to such CSFs vanish in our consideration of size-extensivity. Thus, only CSFs that are excited on one or the other atom contribute to the energy:

$$E = \langle \Phi_a | H | \sum_{J_a} C_{J_a} \Phi_{J_a} + \sum_{J_b} C_{J_b} \Phi_{J_b} \rangle,$$

where  $\Phi_a$  and  $\Phi_b$  as well as  $\Phi_{J_a}$  and  $\Phi_{J_b}$  are used to denote the a and b parts of the reference and excited CSFs, respectively.

This expression, once the SC rules are used to reduce it to one- and two- electron integrals, is of the additive form required of any size-extensive method:

$$E = \langle \Phi_a | H | \sum_{J_a} C_{J_a} \Phi_{J_a} \rangle + \langle \Phi_b | H | \sum_{J_b} C_{J_b} \Phi_{J_b} \rangle,$$

and will yield a size-extensive energy if the equations used to determine the  $C_{Ja}$  and  $C_{Jb}$  amplitudes are themselves separable. In MPPT/MBPT, these amplitudes are expressed, in first order, as:

$$C_{Ja} = \langle a | b | H | {}^*J_{ab} \rangle / [E_a^0 + E_b^0 - E_{Ja}^* - E_b^0]$$

(and analogously for  $C_{Jb}$ ). Again using the SC rules, this expression reduces to one that involves only atom a:

$$C_{Ja} = \langle a | H | {}^*J_a \rangle / [E_a^0 - E_{Ja}^*].$$

The additivity of  $E$  and the separability of the equations determining the  $C_J$  coefficients make the MPPT/MBPT energy size-extensive. This property can also be demonstrated for the Coupled-Cluster energy (see the references given above in Chapter 19. I.4). However, size-extensive methods have at least one serious weakness; their energies do not provide upper bounds to the true energies of the system (because their energy functional is not of the expectation-value form for which the upper bound property has been proven).

### C. Which Method is Best?

At this time, it may not possible to say which method is preferred for applications where all are practical. Nor is it possible to assess, in a way that is applicable to most chemical species, the accuracies with which various methods predict bond lengths and energies or other properties. However, there are reasons to recommend some methods over others in specific cases. For example, certain applications require a size-extensive energy (e.g., extended systems that consist of a large or macroscopic number of units or studies of weak intermolecular interactions), so MBPT/MPPT or CC or CAS-based MCSCF are preferred. Moreover, certain chemical reactions (e.g., Woodward-Hoffmann forbidden reactions) and certain bond-breaking events require two or more 'essential' electronic configurations. For them, single-configuration-based methods such as conventional CC and MBTP/MPPT should not be used; MCSCF or CI calculations would be better. Very large molecules, in which thousands of atomic orbital basis functions are required, may be impossible to treat by methods whose effort scales as  $N^4$  or higher; density functional methods would be better to use then.

For all calculations, the choice of atomic orbital basis set must be made carefully, keeping in mind the  $N^4$  scaling of the one- and two-electron integral evaluation step and the

$N^5$  scaling of the two-electron integral transformation step. Of course, basis functions that describe the essence of the states to be studied are essential (e.g., Rydberg or anion states require diffuse functions, and strained rings require polarization functions).

As larger atomic basis sets are employed, the size of the CSF list used to treat dynamic correlation increases rapidly. For example, most of the above methods use singly and doubly excited CSFs for this purpose. For large basis sets, the number of such CSFs,  $N_C$ , scales as the number of electrons squared,  $n_e^2$ , times the number of basis functions squared,  $N^2$ . Since the effort needed to solve the CI secular problem varies as  $N_C^2$  or  $N_C^3$ , a dependence as strong as  $N^4$  to  $N^6$  can result. To handle such large CSF spaces, all of the multiconfigurational techniques mentioned in this paper have been developed to the extent that calculations involving of the order of 100 to 5,000 CSFs are routinely performed and calculations using 10,000, 100,000, and even several million CSFs are practical.

Other methods, most of which can be viewed as derivatives of the techniques introduced above, have been and are still being developed. This ongoing process has been, in large part, stimulated by the explosive growth in computer power and change in computer architecture that has been realized in recent years. All indications are that this growth pattern will continue, so *ab initio* quantum chemistry will likely have an even larger impact on future chemistry research and education (through new insights and concepts).

### III. Further Details on Implementing Multiconfigurational Methods

#### A. The MCSCF Method

The simultaneous optimization of the LCAO-MO and CI coefficients performed within an MCSCF calculation is a quite formidable task. The variational energy functional is a quadratic function of the CI coefficients, and so one can express the stationary conditions for these variables in the secular form:

$$\sum_J H_{IJ} C_J = E C_I.$$

However,  $E$  is a quartic function of the  $C_{,i}$  coefficients because each matrix element  $\langle I | H | J \rangle$  involves one- and two-electron integrals over the mos  $\psi_i$ , and the two-electron integrals depend quartically on the  $C_{,i}$  coefficients. The stationary conditions with respect to these  $C_{,i}$  parameters must be solved iteratively because of this quartic dependence.

It is well known that minimization of a function ( $E$ ) of several non-linear parameters (the  $C_{i,j}$ ) is a difficult task that can suffer from poor convergence and may locate local rather than global minima. In an MCSCF wavefunction containing many CSFs, the energy is only weakly dependent on the orbitals that are weakly occupied (i.e., those that appear in CSFs with small  $C_I$  values); in contrast,  $E$  is strongly dependent on the  $C_{i,j}$  coefficients of those orbitals that appear in the CSFs with larger  $C_I$  values. One is therefore faced with minimizing a function of many variables (there may be as many  $C_{i,j}$  as the square of the number of orbital basis functions) that depends strongly on several of the variables and weakly on many others. This is a very difficult job.

For these reasons, in the MCSCF method, the number of CSFs is usually kept to a small to moderate number (e.g., a few to several hundred) chosen to describe essential correlations (i.e., configuration crossings, proper dissociation) and important dynamical correlations (those electron-pair correlations of angular, radial, left-right, etc. nature that arise when low-lying 'virtual' orbitals are present). In such a compact wavefunction, only spin-orbitals with reasonably large occupations (e.g., as characterized by the diagonal elements of the one-particle density matrix  $\rho_{i,j}$ ) appear. As a result, the energy functional is expressed in terms of variables on which it is strongly dependent, in which case the non-linear optimization process is less likely to be pathological.

Such a compact MCSCF wavefunction is designed to provide a good description of the set of strongly occupied spin-orbitals and of the CI amplitudes for CSFs in which only these spin-orbitals appear. It, of course, provides no information about the spin-orbitals that are not used to form the CSFs on which the MCSCF calculation is based. As a result, the MCSCF energy is invariant to a unitary transformation among these 'virtual' orbitals.

In addition to the references mentioned earlier in ACP and MTC, the following papers describe several of the advances that have been made in the MCSCF method, especially with respect to enhancing its rate and range of convergence: E. Dalggaard and P. Jørgensen, *J. Chem. Phys.* 69, 3833 (1978); H. J. Aa. Jensen, P. Jørgensen, and H. Ågren, *J. Chem. Phys.* 87, 457 (1987); B. H. Lengsfeld, III and B. Liu, *J. Chem. Phys.* 75, 478 (1981).

## B. The Configuration Interaction Method

In the CI method, one usually attempts to realize a high-level treatment of electron correlation. A set of orthonormal molecular orbitals are first obtained from an SCF or MCSCF calculation (usually involving a small to moderate list of CSFs). The LCAO-MO

coefficients of these orbitals are no longer considered as variational parameters in the subsequent CI calculation; only the  $C_I$  coefficients are to be further optimized.

The CI wavefunction

$$= \sum_J C_J \Psi_J$$

is most commonly constructed from CSFs  $\Psi_J$  that include:

1. All of the CSFs in the SCF (in which case only a single CSF is included) or MCSCF wavefunction that was used to generate the molecular orbitals  $\psi_i$ . This set of CSFs are referred to as spanning the 'reference space' of the subsequent CI calculation, and the particular combination of these CSFs used in this orbital optimization (i.e., the SCF or MCSCF wavefunction) is called the reference function.
2. CSFs that are generated by carrying out single, double, triple, etc. level 'excitations' (i.e., orbital replacements) relative to reference CSFs. CI wavefunctions limited to include contributions through various levels of excitation (e.g., single, double, etc.) are denoted S (singly excited), D (doubly), SD (singly and doubly), SDT (singly, doubly, and triply), and so on.

The orbitals from which electrons are removed and those into which electrons are excited can be restricted to focus attention on correlations among certain orbitals. For example, if excitations out of core electrons are excluded, one computes a total energy that contains no correlation corrections for these core orbitals. Often it is possible to so limit the nature of the orbital excitations to focus on the energetic quantities of interest (e.g., the CC bond breaking in ethane requires correlation of the  $\psi_{CC}$  orbital but the 1s Carbon core orbitals and the CH bond orbitals may be treated in a non-correlated manner).

Clearly, the number of CSFs included in the CI calculation can be far in excess of the number considered in typical MCSCF calculations; CI wavefunctions including 5,000 to 50,000 CSFs are routinely used, and functions with one to several million CSFs are within the realm of practicality (see, for example, J. Olsen, B. Roos, Poul Jørgensen, and H. J. Aa. Jensen, J. Chem. Phys. 89, 2185 (1988) and J. Olsen, P. Jørgensen, and J. Simons, Chem. Phys. Letters 169, 463 (1990)).

The need for such large CSF expansions should not come as a surprise once one considers that (i) each electron pair requires at least two CSFs (let us say it requires P of them, on average, a dominant one and P-1 others which are doubly excited) to form

polarized orbital pairs, (ii) there are of the order of  $N(N-1)/2 = X$  electron pairs in an atom or molecule containing  $N$  electrons, and (iii) that the number of terms in the CI wavefunction scales as  $P^X$ . So, for an  $H_2O$  molecule containing ten electrons, there would be  $P^{55}$  terms in the CI expansion. This is  $3.6 \times 10^{16}$  terms if  $P=2$  and  $1.7 \times 10^{26}$  terms if  $P=3$ . Undoubtedly, this is an over estimate of the number of CSFs needed to describe electron correlation in  $H_2O$ , but it demonstrates how rapidly the number of CSFs can grow with the number of electrons in the system.

The  $H_{I,J}$  matrices that arise in CI calculations are evaluated in terms of one- and two- electron integrals over the molecular orbitals using the equivalent of the Slater-Condon rules. For large CI calculations, the full  $H_{I,J}$  matrix is not actually evaluated and stored in the computer's memory (or on its disk); rather, so-called 'direct CI' methods (see the article by Roos and Siegbahn in MTC) are used to compute and immediately sum contributions to the sum  $\sum_I H_{I,J} C_J$  in terms of integrals, density matrix elements, and approximate values of the  $C_J$  amplitudes. Iterative methods (see, for example, E. R. Davidson, *J. Comput. Phys.* **17**, 87 (1975)), in which approximate values for the  $C_J$  coefficients and energy  $E$  are refined through sequential application of  $\sum_I H_{I,J}$  to the preceding estimate of the  $C_J$  vector, are employed to solve these large CI matrix eigenvalue problems.

### C. The MPPT/MBPT Method

In the MPPT/MBPT method, once the reference CSF is chosen and the SCF orbitals belonging to this CSF are determined, the wavefunction and energy  $E$  are determined in an order-by-order manner. This is one of the primary strengths of the MPPT/MBPT technique; it does not require one to make further (potentially arbitrary) choices once the basis set and dominant (SCF) configuration are specified. In contrast to the MCSCF and CI treatments, one need not make choices of CSFs to include in or exclude from . The MPPT/MBPT perturbation equations determine what CSFs must be included through any particular order.

For example, the first-order wavefunction correction  $\psi^1$  (i.e.,  $\psi = \psi^0 + \psi^1$  through first order) is given by:

$$\psi^1 = - \sum_{i < j, m < n} \langle i, j | g | m, n \rangle \langle i, j | g | m, n \rangle^{-1} [ \epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j ]^{-1} | i, j | m, n \rangle$$

$$= - \sum_{i < j, m < n} [ \langle i, j | g | m, n \rangle - \langle i, j | g | n, m \rangle ] [ \epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j ]^{-1} | i, j | m, n \rangle$$

where the SCF orbital energies are denoted  $\epsilon_k$  and  $\phi_{i,j}^{m,n}$  represents a CSF that is doubly excited relative to  $\phi_0$ . Thus, only doubly excited CSFs contribute to the first-order wavefunction; as a result, the energy  $E$  is given through second order as:

$$\begin{aligned}
 E &= \langle \phi_0 | H^0 | \phi_0 \rangle + \langle \phi_0 | H - H^0 | \phi_0 \rangle + \langle \phi_0 | H - H^0 | \phi_1 \rangle \\
 &= \langle \phi_0 | H | \phi_0 \rangle - \sum_{i<j, m<n} \langle \phi_0 | \phi_{i,j}^{m,n} | H - H^0 | \phi_0 \rangle^2 / [\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j] \\
 &= E_{\text{SCF}} - \sum_{i<j, m<n} \langle \phi_0 | \phi_{i,j}^{m,n} | H - H^0 | \phi_0 \rangle^2 / [\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j] \\
 &= E^0 + E^1 + E^2.
 \end{aligned}$$

These contributions have been expressed, using the SC rules, in terms of the two-electron integrals  $\langle i,j | g | m,n \rangle$  coupling the excited spin-orbitals to the spin-orbitals from which electrons were excited as well as the orbital energy differences  $[\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j]$  accompanying such excitations. In this form, it becomes clear that major contributions to the correlation energy of the pair of occupied orbitals  $i, j$  are made by double excitations into virtual orbitals  $m, n$  that have large coupling (i.e., large  $\langle i,j | g | m,n \rangle$  integrals) and small orbital energy gaps,  $[\epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j]$ .

In higher order corrections to the wavefunction and to the energy, contributions from CSFs that are singly, triply, etc. excited relative to  $\phi_0$  appear, and additional contributions from the doubly excited CSFs also enter. It is relatively common to carry MPPT/MBPT calculations (see the references given above in Chapter 19.I.3 where the contributions of the Pople and Bartlett groups to the development of MPPT/MBPT are documented) through to third order in the energy (whose evaluation can be shown to require only  $\epsilon_0$  and  $\epsilon_1$ ). The entire GAUSSIAN-8X series of programs, which have been used in thousands of important chemical studies, calculate  $E$  through third order in this manner.

In addition to being size-extensive and not requiring one to specify input beyond the basis set and the dominant CSF, the MPPT/MBPT approach is able to include the effect of all CSFs (that contribute to any given order) without having to find any eigenvalues of a matrix. This is an important advantage because matrix eigenvalue determination, which is necessary in MCSCF and CI calculations, requires computer time in proportion to the third power of the dimension of the  $H_{I,J}$  matrix. Despite all of these advantages, it is important to remember the primary disadvantages of the MPPT/MBPT approach; its energy is not an upper bound to the true energy and it may not be able to treat cases for which two or more

CSFs have equal or nearly equal amplitudes because it obtains the amplitudes of all but the dominant CSF from perturbation theory formulas that assume the perturbation is 'small'.

#### D. The Coupled-Cluster Method

The implementation of the CC method begins much as in the MPPT/MBPT case; one selects a reference CSF that is used in the SCF process to generate a set of spin-orbitals to be used in the subsequent correlated calculation. The set of working equations of the CC technique given above in Chapter 19.I.4 can be written explicitly by introducing the form of the so-called cluster operator  $T$ ,

$$T = \sum_{i,m} t_i^m m^+ i + \sum_{i,j,m,n} t_{i,j}^{m,n} m^+ n^+ j i + \dots,$$

where the combination of operators  $m^+ i$  denotes creation of an electron in virtual spin-orbital  $m$  and removal of an electron from occupied spin-orbital  $i$  to generate a single excitation. The operation  $m^+ n^+ j i$  therefore represents a double excitation from  $i j$  to  $m n$ . Expressing the cluster operator  $T$  in terms of the amplitudes  $t_i^m$ ,  $t_{i,j}^{m,n}$ , etc. for singly, doubly, etc. excited CSFs, and expanding the exponential operators in  $\exp(-T) H \exp(T)$  one obtains:

$$\langle i^m | H + [H,T] + 1/2 [[H,T],T] + 1/6 [[[H,T],T],T] + 1/24 [[[[H,T],T],T],T] | \rangle = 0;$$

$$\langle i,j^{m,n} | H + [H,T] + 1/2 [[H,T],T] + 1/6 [[[H,T],T],T] + 1/24 [[[[H,T],T],T],T] | \rangle = 0;$$

$$\langle i,j,k^{m,n,p} | H + [H,T] + 1/2 [[H,T],T] + 1/6 [[[H,T],T],T] + 1/24 [[[[H,T],T],T],T] | \rangle = 0,$$

and so on for higher order excited CSFs. It can be shown, because of the one- and two-electron operator nature of  $H$ , that the expansion of the exponential operators truncates exactly at the fourth power; that is terms such as  $[[[[[H,T],T],T],T],T]$  and higher commutators vanish identically (this is demonstrated in Chapter 4 of Second Quantization Based Methods in Quantum Chemistry, P. Jørgensen and J. Simons, Academic Press, New York (1981).

As a result, the exact CC equations are quartic equations for the  $t_i^m$ ,  $t_{i,j}^{m,n}$ , etc. amplitudes. Although it is a rather formidable task to evaluate all of the commutator matrix elements appearing in the above CC equations, it can be and has been done (the references given above to Purvis and Bartlett are especially relevant in this context). The result is to express each such matrix element, via the Slater-Condon rules, in terms of one- and two-electron integrals over the spin-orbitals used in determining  $\Psi$ , including those in  $\Psi$  itself and the 'virtual' orbitals not in  $\Psi$ .

In general, these quartic equations must then be solved in an iterative manner and are susceptible to convergence difficulties that are similar to those that arise in MCSCF-type calculations. In any such iterative process, it is important to start with an approximation (to the  $t$  amplitudes, in this case) which is reasonably close to the final converged result. Such an approximation is often achieved, for example, by neglecting all of the terms that are non-linear in the  $t$  amplitudes (because these amplitudes are assumed to be less than unity in magnitude). This leads, for the CC working equations obtained by projecting onto the doubly excited CSFs, to:

$$\langle i,j | g | m,n \rangle' + [ \epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j ] t_{i,j}^{m,n} + \sum_{i',j',m',n'} \langle i',j',m',n' | H - H^0 | i,j,m,n \rangle t_{i',j'}^{m',n'} = 0 ,$$

where the notation  $\langle i,j | g | m,n \rangle'$  is used to denote the two-electron integral difference  $\langle i,j | g | m,n \rangle - \langle i,j | g | n,m \rangle$ . If, in addition, the factors that couple different doubly excited CSFs are ignored (i.e., the sum over  $i',j',m',n'$ ), the equations for the  $t$  amplitudes reduce to the equations for the CSF amplitudes of the first-order MPPT/MBPT wavefunction:

$$t_{i,j}^{m,n} = - \langle i,j | g | m,n \rangle' / [ \epsilon_m - \epsilon_i + \epsilon_n - \epsilon_j ] .$$

As Bartlett and Pople have both demonstrated, there is, in fact, close relationship between the MPPT/MBPT and CC methods when the CC equations are solved iteratively starting with such an MPPT/MBPT-like initial 'guess' for these double-excitation amplitudes.

The CC method, as presented here, suffers from the same drawbacks as the MPPT/MBPT approach; its energy is not an upper bound and it may not be able to accurately describe wavefunctions which have two or more CSFs with approximately equal amplitude. Moreover, solution of the non-linear CC equations may be difficult and slowly (if at all) convergent. It has the same advantages as the MPPT/MBPT method; its energy is

size-extensive, it requires no large matrix eigenvalue solution, and its energy and wavefunction are determined once one specifies the basis and the dominant CSF.

### E. Density Functional Methods

These approaches provide alternatives to the conventional tools of quantum chemistry. The CI, MCSCF, MPPT/MBPT, and CC methods move beyond the single-configuration picture by adding to the wave function more configurations whose amplitudes they each determine in their own way. This can lead to a very large number of CSFs in the correlated wave function, and, as a result, a need for extraordinary computer resources.

The density functional approaches are different. Here one solves a set of orbital-level equations

$$\left[ -\frac{\hbar^2}{2m_e} \nabla^2 - \sum_A \frac{Z_A e^2}{|\mathbf{r} - \mathbf{R}_A|} + \int (\mathbf{r}') \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + U(\mathbf{r}) \right] \psi_i = \epsilon_i \psi_i$$

in which the orbitals  $\{\psi_i\}$  'feel' potentials due to the nuclear centers (having charges  $Z_A$ ), Coulombic interaction with the total electron density  $\rho(\mathbf{r}')$ , and a so-called exchange-correlation potential denoted  $U(\mathbf{r}')$ . The particular electronic state for which the calculation is being performed is specified by forming a corresponding density  $\rho(\mathbf{r}')$ . Before going further in describing how DFT calculations are carried out, let us examine the origins underlying this theory.

The so-called Hohenberg-Kohn theorem states that the ground-state electron density  $\rho(\mathbf{r})$  describing an N-electron system uniquely determines the potential  $V(\mathbf{r})$  in the Hamiltonian

$$H = \sum_j \left\{ -\frac{\hbar^2}{2m_e} \nabla_j^2 + V(\mathbf{r}_j) + \frac{e^2}{2} \sum_{k \neq j} \frac{1}{r_{j,k}} \right\},$$

and, because  $H$  determines the ground-state energy and wave function of the system, the ground-state density  $\rho(\mathbf{r})$  determines the ground-state properties of the system. The proof of this theorem proceeds as follows:

- a.  $\rho(\mathbf{r})$  determines  $N$  because  $\int \rho(\mathbf{r}) d^3r = N$ .
- b. Assume that there are two distinct potentials (aside from an additive constant that simply shifts the zero of total energy)  $V(\mathbf{r})$  and  $V'(\mathbf{r})$  which, when used in  $H$  and  $H'$ , respectively, to solve for a ground state produce  $E_0, \psi_0(\mathbf{r})$  and  $E_0', \psi_0'(\mathbf{r})$  that have the same one-electron density:  $|\psi_0|^2 dr_2 dr_3 \dots dr_N = \rho(\mathbf{r}) = |\psi_0'|^2 dr_2 dr_3 \dots dr_N$ .
- c. If we think of  $\psi_0'$  as trial variational wave function for the Hamiltonian  $H$ , we know that  $E_0 < \langle \psi_0' | H | \psi_0' \rangle = \langle \psi_0' | H' | \psi_0' \rangle + \int \rho(\mathbf{r}) [V(\mathbf{r}) - V'(\mathbf{r})] d^3r = E_0' + \int \rho(\mathbf{r}) [V(\mathbf{r}) - V'(\mathbf{r})] d^3r$ .
- d. Similarly, taking  $\psi_0$  as a trial function for the  $H'$  Hamiltonian, one finds that  $E_0' < E_0 + \int \rho(\mathbf{r}) [V'(\mathbf{r}) - V(\mathbf{r})] d^3r$ .
- e. Adding the equations in c and d gives  $E_0 + E_0' < E_0 + E_0'$ , a clear contradiction.

Hence, there cannot be two distinct potentials  $V$  and  $V'$  that give the same ground-state  $\rho(\mathbf{r})$ . So, the ground-state density  $\rho(\mathbf{r})$  uniquely determines  $N$  and  $V$ , and thus  $H$ , and therefore  $E_0$  and  $\psi_0$ . Furthermore, because  $\rho(\mathbf{r})$  determines all properties of the ground state, then  $\rho(\mathbf{r})$ , in principle, determines all such properties. This means that even the kinetic energy and the electron-electron interaction energy of the ground-state are determined by  $\rho(\mathbf{r})$ . It is easy to see that  $\int \rho(\mathbf{r}) V(\mathbf{r}) d^3r = V[\rho]$  gives the average value of the electron-nuclear (plus any additional one-electron additive potential) interaction in terms of the ground-state density  $\rho(\mathbf{r})$ , but how are the kinetic energy  $T[\rho]$  and the electron-electron interaction  $V_{ee}[\rho]$  energy expressed in terms of  $\rho$ ?

The main difficulty with DFT is that the Hohenberg-Kohn theorem shows that the ground-state values of  $T, V_{ee}, V$ , etc. are all unique functionals of the ground-state (i.e.,

that they can, in principle, be determined once  $\rho$  is given), but it does not tell us what these functional relations are.

To see how it might make sense that a property such as the kinetic energy, whose operator  $-\hbar^2/2m_e \nabla^2$  involves derivatives, can be related to the electron density, consider a simple system of  $N$  non-interacting electrons moving in a three-dimensional cubic “box” potential. The energy states of such electrons are known to be

$$E = (\hbar^2/2m_e L^2) (n_x^2 + n_y^2 + n_z^2),$$

where  $L$  is the length of the box along the three axes, and  $n_x$ ,  $n_y$ , and  $n_z$  are the quantum numbers describing the state. We can view  $n_x^2 + n_y^2 + n_z^2 = R^2$  as defining the squared radius of a sphere in three dimensions, and we realize that the density of quantum states in this space is one state per unit volume in the  $n_x$ ,  $n_y$ ,  $n_z$  space. Because  $n_x$ ,  $n_y$ , and  $n_z$  must be positive integers, the volume covering all states with energy less than or equal to a specified energy  $E = (\hbar^2/2m_e L^2) R^2$  is  $1/8$  the volume of the sphere of radius  $R$ :

$$V(E) = 1/8 (4\pi/3) R^3 = (\pi/6) (8m_e L^2 E/\hbar^2)^{3/2}.$$

Since there is one state per unit of such volume,  $V(E)$  is also the number of states with energy less than or equal to  $E$ , and is called the integrated density of states. The number of states  $g(E) dE$  with energy between  $E$  and  $E+dE$ , the density of states, is the derivative of  $V(E)$ :

$$g(E) = dV/dE = (\pi/4) (8m_e L^2/\hbar^2)^{3/2} E^{1/2}.$$

If we calculate the total energy for  $N$  electrons, with the states having energies up to the so-called Fermi energy (i.e., the energy of the highest occupied molecular orbital HOMO) doubly occupied, we obtain the ground-state energy:

$$E_0 = 2 \int_0^{E_F} g(E) E dE = (8\pi/5) (2m_e/\hbar^2)^{3/2} L^3 E_F^{5/2}.$$

The total number of electrons  $N$  can be expressed as

$$N = 2 \int_0^{E_F} g(E) dE = (8/3) (2m_e/h^2)^{3/2} L^3 E_F^{3/2},$$

which can be solved for  $E_F$  in terms of  $N$  to then express  $E_0$  in terms of  $N$  instead of  $E_F$ :

$$E_0 = (3h^2/10m_e) (3/8)^{2/3} L^3 (N/L^3)^{5/3}.$$

This gives the total energy, which is also the kinetic energy in this case because the potential energy is zero within the “box”, in terms of the electron density  $n(x,y,z) = (N/L^3)$ . It therefore may be plausible to express kinetic energies in terms of electron densities  $n(\mathbf{r})$ , but it is by no means clear how to do so for “real” atoms and molecules with electron-nuclear and electron-electron interactions operative.

In one of the earliest DFT models, the Thomas-Fermi theory, the kinetic energy of an atom or molecule is approximated using the above kind of treatment on a “local” level. That is, for each volume element in  $\mathbf{r}$  space, one assumes the expression given above to be valid, and then one integrates over all  $\mathbf{r}$  to compute the total kinetic energy:

$$T_{TF}[n] = (3h^2/10m_e) (3/8)^{2/3} \int [n(\mathbf{r})]^{5/3} d^3r = C_F \int [n(\mathbf{r})]^{5/3} d^3r,$$

where the last equality simply defines the  $C_F$  constant (which is 2.8712 in atomic units). Ignoring the correlation and exchange contributions to the total energy, this  $T$  is combined with the electron-nuclear  $V$  and Coulombic electron-electron potential energies to give the Thomas-Fermi total energy:

$$E_{0,TF}[n] = C_F \int [n(\mathbf{r})]^{5/3} d^3r + \int V(\mathbf{r}) n(\mathbf{r}) d^3r + e^2/2 \int \int n(\mathbf{r}) n(\mathbf{r}')/|\mathbf{r}-\mathbf{r}'| d^3r d^3r',$$

This expression is an example of how  $E_0$  is given as a local density functional approximation (LDA). The term local means that the energy is given as a functional (i.e., a function of  $n$ ) which depends only on  $n(\mathbf{r})$  at points in space but not on  $n(\mathbf{r})$  at more than one point in space.

Unfortunately, the Thomas-Fermi energy functional does not produce results that are of sufficiently high accuracy to be of great use in chemistry. What is missing in this

theory are a. the exchange energy and b. the correlation energy; moreover, the kinetic energy is treated only in the approximate manner described.

In the book by Parr and Yang, it is shown how Dirac was able to address the exchange energy for the 'uniform electron gas' (N Coulomb interacting electrons moving in a uniform positive background charge whose magnitude balances the charge of the N electrons). If the exact expression for the exchange energy of the uniform electron gas is applied on a local level, one obtains the commonly used Dirac local density approximation to the exchange energy:

$$E_{\text{ex,Dirac}}[\rho] = -C_x \int \rho(\mathbf{r})^{4/3} d^3r,$$

with  $C_x = (3/4) (3/\pi)^{1/3} = 0.7386$  in atomic units. Adding this exchange energy to the Thomas-Fermi total energy  $E_{0,\text{TF}}[\rho]$  gives the so-called Thomas-Fermi-Dirac (TFD) energy functional.

Because electron densities vary rather strongly spatially near the nuclei, corrections to the above approximations to  $T[\rho]$  and  $E_{\text{ex,Dirac}}$  are needed. One of the more commonly used so-called gradient-corrected approximations is that invented by Becke, and referred to as the Becke88 exchange functional:

$$E_{\text{ex}}(\text{Becke88}) = E_{\text{ex,Dirac}}[\rho] - \int x^2 \rho^{4/3} (1 + 6 \rho^{-1/3} |\nabla \rho| \sinh^{-1}(x))^{-1} d\mathbf{r},$$

where  $x = \alpha |\nabla \rho| \rho^{-1/3}$ , and  $\alpha$  is a parameter chosen so that the above exchange energy can best reproduce the known exchange energies of specific electronic states of the inert gas atoms (Becke finds  $\alpha$  to equal 0.0042). A common gradient correction to the earlier  $T[\rho]$  is called the Weizsacker correction and is given by

$$T_{\text{Weizsacker}} = (1/72)(\hbar/m_e) \int |\nabla \rho|^2 / \rho d\mathbf{r}.$$

Although the above discussion suggests how one might compute the ground-state energy once the ground-state density  $\rho(\mathbf{r})$  is given, one still needs to know how to obtain

. Kohn and Sham (KS) introduced a set of so-called KS orbitals obeying the following equation:

$$\left\{ -\frac{1}{2} \nabla^2 + V(\mathbf{r}) + \frac{e^2}{2} \int \frac{|\mathbf{r}'|}{|\mathbf{r}-\mathbf{r}'|} |\mathbf{d}\mathbf{r}' + U_{xc}(\mathbf{r}) \right\} \psi_j = \epsilon_j \psi_j,$$

where the so-called exchange-correlation potential  $U_{xc}(\mathbf{r}) = \delta E_{xc}[\rho] / \delta \rho(\mathbf{r})$  could be obtained by functional differentiation if the exchange-correlation energy functional  $E_{xc}[\rho]$  were known. KS also showed that the KS orbitals  $\{\psi_j\}$  could be used to compute the density by simply adding up the orbital densities multiplied by orbital occupancies  $n_j$ :

$$\rho(\mathbf{r}) = \sum_j n_j |\psi_j(\mathbf{r})|^2.$$

(here  $n_j = 0, 1, \text{ or } 2$  is the occupation number of the orbital  $\psi_j$  in the state being studied) and that the kinetic energy should be calculated as

$$T = \sum_j n_j \langle \psi_j(\mathbf{r}) | -\frac{1}{2} \nabla^2 | \psi_j(\mathbf{r}) \rangle.$$

The same investigations of the idealized 'uniform electron gas' that identified the Dirac exchange functional, found that the correlation energy (per electron) could also be written exactly as a function of the electron density of the system, but only in two limiting cases- the high-density limit (large  $\rho$ ) and the low-density limit. There still exists no exact expression for the correlation energy even for the uniform electron gas that is valid at arbitrary values of  $\rho$ . Therefore, much work has been devoted to creating efficient and accurate interpolation formulas connecting the low- and high- density uniform electron gas expressions. One such expression is

$$E_c[\rho] = \int \epsilon_c(\rho(\mathbf{r})) \mathbf{d}\mathbf{r},$$

where

$$\epsilon_c(\rho) = A/2 \{ \ln(x/X) + 2b/Q \tan^{-1}(Q/(2x+b)) - bx_0/X_0 [\ln((x-x_0)^2/X) \}$$

$$+2(b+2x_0)/Q \tan^{-1}(Q/(2x+b))]$$

is the correlation energy per electron. Here  $x = r_s^{1/2}$ ,  $X = x^2 + bx + c$ ,  $X_0 = x_0^2 + bx_0 + c$  and  $Q = (4c - b^2)^{1/2}$ ,  $A = 0.0621814$ ,  $x_0 = -0.409286$ ,  $b = 13.0720$ , and  $c = 42.7198$ . The parameter  $r_s$  is how the density enters since  $4/3 r_s^3$  is equal to  $1/\rho$ ; that is,  $r_s$  is the radius of a sphere whose volume is the effective volume occupied by one electron. A reasonable approximation to the full  $E_{xc}[\rho]$  would contain the Dirac (and perhaps gradient corrected) exchange functional plus the above  $E_c[\rho]$ , but there are many alternative approximations to the exchange-correlation energy functional. Currently, many workers are doing their best to “cook up” functionals for the correlation and exchange energies, but no one has yet invented functionals that are so reliable that most workers agree to use them.

To summarize, in implementing any DFT, one usually proceeds as follows:

1. An atomic orbital basis is chosen in terms of which the KS orbitals are to be expanded.
2. Some initial guess is made for the LCAO-KS expansion coefficients  $C_{j,a}$ :  $\psi_j = \sum_a C_{j,a} \phi_a$ .
3. The density is computed as  $\rho(\mathbf{r}) = \sum_j n_j |\psi_j(\mathbf{r})|^2$ . Often,  $\rho(\mathbf{r})$  is expanded in an atomic orbital basis, which need not be the same as the basis used for the  $\psi_j$ , and the expansion coefficients of  $\rho$  are computed in terms of those of the  $\psi_j$ . It is also common to use an atomic orbital basis to expand  $\rho^{1/3}(\mathbf{r})$  which, together with  $\rho$ , is needed to evaluate the exchange-correlation functional's contribution to  $E_0$ .
4. The current iteration's density is used in the KS equations to determine the Hamiltonian  $\{-1/2 \nabla^2 + V(\mathbf{r}) + e^2/2 \sum_{j'} \rho(\mathbf{r}')/|\mathbf{r}-\mathbf{r}'| - \mathbf{p} \cdot \mathbf{p}' + U_{xc}(\mathbf{r})\}$  whose “new” eigenfunctions  $\{\psi_j\}$  and eigenvalues  $\{\epsilon_j\}$  are found by solving the KS equations.
5. These new  $\psi_j$  are used to compute a new density, which, in turn, is used to solve a new set of KS equations. This process is continued until convergence is reached (i.e., until the  $\psi_j$  used to determine the current iteration's  $\rho$  are the same  $\psi_j$  that arise as solutions on the next iteration).
6. Once the converged  $\rho(\mathbf{r})$  is determined, the energy can be computed using the earlier expression

$$E[\psi] = \int \psi^* \left[ -\frac{1}{2} \nabla^2 + V(\mathbf{r}) \right] \psi \, d\mathbf{r} + \frac{e^2}{2} \iint \frac{|\psi(\mathbf{r})\psi(\mathbf{r}')|^2}{|\mathbf{r}-\mathbf{r}'|} \, d\mathbf{r} \, d\mathbf{r}' + E_{xc}[\psi].$$

In closing this section, it should once again be emphasized that this area is currently undergoing explosive growth and much scrutiny. As a result, it is nearly certain that many of the specific functionals discussed above will be replaced in the near future by improved and more rigorously justified versions. It is also likely that extensions of DFT to excited states (many workers are actively pursuing this) will be placed on more solid ground and made applicable to molecular systems. Because the computational effort involved in these approaches scales much less strongly with basis set size than for conventional (SCF, MCSCF, CI, etc.) methods, density functional methods offer great promise and are likely to contribute much to quantum chemistry in the next decade.

## Chapter 20

*Many physical properties of a molecule can be calculated as expectation values of a corresponding quantum mechanical operator. The evaluation of other properties can be formulated in terms of the "response" (i.e., derivative) of the electronic energy with respect to the application of an external field perturbation.*

### I. Calculations of Properties Other Than the Energy

There are, of course, properties other than the energy that are of interest to the practicing chemist. Dipole moments, polarizabilities, transition probabilities among states, and vibrational frequencies all come to mind. Other properties that are of importance involve operators whose quantum numbers or symmetry indices label the state of interest. Angular momentum and point group symmetries are examples of the latter properties; for these quantities the properties are precisely specified once the quantum number or symmetry label is given (e.g., for a  $^3P$  state, the average value of  $L^2$  is  $\langle ^3P | L^2 | ^3P \rangle = \hbar^2 1(1+1) = 2\hbar^2$ ).

Although it may be straightforward to specify what property is to be evaluated, often computational difficulties arise in carrying out the calculation. For some ab initio methods, these difficulties are less severe than for others. For example, to compute the electric dipole transition matrix element  $\langle \psi_2 | \mathbf{r} | \psi_1 \rangle$  between two states  $\psi_1$  and  $\psi_2$ , one must evaluate the integral involving the one-electron dipole operator  $\mathbf{r} = \sum_j e \mathbf{r}_j - \sum_a Z_a \mathbf{R}_a$ ; here the first sum runs over the  $N$  electrons and the second sum runs over the nuclei whose charges are denoted  $Z_a$ . To evaluate such transition matrix elements in terms of the Slater-Condon rules is relatively straightforward as long as  $\psi_1$  and  $\psi_2$  are expressed in terms of Slater determinants involving a single set of orthonormal spin-orbitals. If  $\psi_1$  and  $\psi_2$  have been obtained, for example, by carrying out separate MCSCF calculations on the two states in question, the energy optimized spin-orbitals for one state will not be the same as the optimal spin-orbitals for the second state. As a result, the determinants in  $\psi_1$  and those in  $\psi_2$  will involve spin-orbitals that are not orthonormal to one another. Thus, the SC rules can not immediately be applied. Instead, a transformation of the spin-orbitals of  $\psi_1$  and  $\psi_2$  to a single set of orthonormal functions must be carried out. This then expresses  $\psi_1$  and  $\psi_2$  in terms of new Slater determinants over this new set of orthonormal spin-orbitals, after which the SC rules can be exploited.

In contrast, if  $\psi_1$  and  $\psi_2$  are obtained by carrying out a CI calculation using a single set of orthonormal spin-orbitals (e.g., with  $\psi_1$  and  $\psi_2$  formed from two different

eigenvectors of the resulting secular matrix), the SC rules can immediately be used to evaluate the transition dipole integral.

### A. Formulation of Property Calculations as Responses

Essentially all experimentally measured properties can be thought of as arising through the response of the system to some externally applied perturbation or disturbance. In turn, the calculation of such properties can be formulated in terms of the response of the energy  $E$  or wavefunction to a perturbation. For example, molecular dipole moments  $\mu$  are measured, via electric-field deflection, in terms of the change in energy

$$E = \mu \cdot \mathbf{E} + 1/2 \mathbf{E} \cdot \alpha \cdot \mathbf{E} + 1/6 \mathbf{E} \cdot \beta \cdot \mathbf{E} \cdot \mathbf{E} + \dots$$

caused by the application of an external electric field  $\mathbf{E}$  which is spatially inhomogeneous, and thus exerts a force

$$\mathbf{F} = - \nabla E$$

on the molecule proportional to the dipole moment (good treatments of response properties for a wide variety of wavefunction types (i.e., SCF, MCSCF, MPPT/MBPT, etc.) are given in Second Quantization Based Methods in Quantum Chemistry, P. Jørgensen and J. Simons, Academic Press, New York (1981) and in Geometrical Derivatives of Energy Surfaces and Molecular Properties, P. Jørgensen and J. Simons, Eds., NATO ASI Series, Vol. 166, D. Reidel, Dordrecht (1985)).

To obtain expressions that permit properties other than the energy to be evaluated in terms of the state wavefunction, the following strategy is used:

1. The perturbation  $V = H - H^0$  appropriate to the particular property is identified. For dipole moments ( $\mu$ ), polarizabilities ( $\alpha$ ), and hyperpolarizabilities ( $\beta$ ),  $V$  is the interaction of the nuclei and electrons with the external electric field

$$V = - \sum_a Z_a e \mathbf{R}_a \cdot \mathbf{E} - \sum_j e \mathbf{r}_j \cdot \mathbf{E}.$$

For vibrational frequencies, one needs the derivatives of the energy  $E$  with respect to deformation of the bond lengths and angles of the molecule, so  $V$  is the sum of all changes in the electronic Hamiltonian that arise from displacements  $\mathbf{R}_a$  of the atomic centers

$$V = \sum_a (\mathbf{R}_a H) \cdot \mathbf{R}_a.$$

2. A power series expansion of the state energy  $E$ , computed in a manner consistent with how  $\psi$  is determined (i.e., as an expectation value for SCF, MCSCF, and CI wavefunctions or as  $\langle \psi | H | \psi \rangle$  for MPPT/MBPT or as  $\langle \exp(-T) H \exp(T) | \psi \rangle$  for CC wavefunctions), is carried out in powers of the perturbation  $V$ :

$$E = E^0 + E^{(1)} + E^{(2)} + E^{(3)} + \dots$$

In evaluating the terms in this expansion, the dependence of  $H = H^0 + V$  and of  $\psi$  (which is expressed as a solution of the SCF, MCSCF, ..., or CC equations for  $H$  not for  $H^0$ ) must be included.

3. The desired physical property must be extracted from the power series expansion of  $E$  in powers of  $V$ .

## B. The MCSCF Response Case

### 1. The Dipole Moment

To illustrate how the above developments are carried out and to demonstrate how the results express the desired quantities in terms of the original wavefunction, let us consider, for an MCSCF wavefunction, the response to an external electric field. In this case, the Hamiltonian is given as the conventional one- and two-electron operators  $H^0$  to which the above one-electron electric dipole perturbation  $V$  is added. The MCSCF wavefunction  $\psi$  and energy  $E$  are assumed to have been obtained via the MCSCF procedure with  $H = H^0 + V$ , where  $\epsilon$  can be thought of as a measure of the strength of the applied electric field.

The terms in the expansion of  $E(\epsilon)$  in powers of  $\epsilon$  :

$$E = E(\epsilon=0) + \left( \frac{dE}{d\epsilon} \right)_0 + \frac{1}{2} \epsilon^2 \left( \frac{d^2E}{d\epsilon^2} \right)_0 + \dots$$

are obtained by writing the total derivatives of the MCSCF energy functional with respect to  $\psi$  and evaluating these derivatives at  $\epsilon=0$

(which is indicated by the subscript  $(\cdot)_0$  on the above derivatives):

$$E(\lambda=0) = \langle \psi(\lambda=0) | H^0 | \psi(\lambda=0) \rangle = E^0,$$

$$\left(\frac{dE}{d\lambda}\right)_0 = \langle \psi(\lambda=0) | V | \psi(\lambda=0) \rangle + 2 \sum_J \left(\frac{C_J}{\lambda}\right)_0 \langle \psi(\lambda=0) | C_J H^0 | \psi(\lambda=0) \rangle$$

$$+ 2 \sum_{i,a} \left(\frac{C_{a,i}}{\lambda}\right)_0 \langle \psi(\lambda=0) | C_{a,i} H^0 | \psi(\lambda=0) \rangle$$

$$+ 2 \sum_{i,a} \left(\frac{C_{a,i}}{\lambda}\right)_0 \langle \psi(\lambda=0) | H^0 | C_{a,i} \psi(\lambda=0) \rangle,$$

and so on for higher order terms. The factors of 2 in the last three terms come through using the hermiticity of  $H^0$  to combine terms in which derivatives of  $\psi$  occur.

The first-order correction can be thought of as arising from the response of the wavefunction (as contained in its LCAO-MO and CI amplitudes and basis functions  $\psi$ ) plus the response of the Hamiltonian to the external field. Because the MCSCF energy functional has been made stationary with respect to variations in the  $C_J$  and  $C_{i,a}$  amplitudes, the second and third terms above vanish:

$$E/C_J = 2 \langle \psi(\lambda=0) | C_J H^0 | \psi(\lambda=0) \rangle = 0,$$

$$E/C_{a,i} = 2 \langle \psi(\lambda=0) | C_{a,i} H^0 | \psi(\lambda=0) \rangle = 0.$$

If, as is common, the atomic orbital bases used to carry out the MCSCF energy optimization are not explicitly dependent on the external field, the third term also vanishes because  $(\psi(\lambda=0) | C_{a,i} \psi(\lambda=0))_0 = 0$ . Thus for the MCSCF case, the first-order response is given as the average value of the perturbation over the wavefunction with  $\lambda=0$ :

$$\left(\frac{dE}{d\lambda}\right)_0 = \langle \psi(\lambda=0) | V | \psi(\lambda=0) \rangle.$$

For the external electric field case at hand, this result says that the field-dependence of the state energy will have a linear term equal to

$$\langle \psi(\lambda=0) | V | \psi(\lambda=0) \rangle = \langle \psi(\lambda=0) | \sum_a Z_a \mathbf{e} \cdot \mathbf{R}_a - \sum_j e \mathbf{r}_j \cdot \mathbf{e} | \psi(\lambda=0) \rangle,$$

where  $\mathbf{e}$  is a unit vector in the direction of the applied electric field (the magnitude of the field  $\lambda$  having already been removed in the power series expansion). Since the dipole

moment is determined experimentally as the energy's slope with respect to field strength, this means that the dipole moment is given as:

$$\mu = \langle \psi | \sum_a Z_a e \mathbf{R}_a - \sum_j e \mathbf{r}_j | \psi \rangle.$$

## 2. The Geometrical Force

These same techniques can be used to determine the response of the energy to displacements  $\mathbf{R}_a$  of the atomic centers. In such a case, the perturbation is

$$V = - \sum_a \mathbf{R}_a \cdot \sum_i Z_a e^2 \frac{\mathbf{r}_i - \mathbf{R}_a}{|\mathbf{r}_i - \mathbf{R}_a|^3}$$

$$= - \sum_a Z_a e^2 \mathbf{R}_a \cdot \sum_i (\mathbf{r}_i - \mathbf{R}_a) / |\mathbf{r}_i - \mathbf{R}_a|^3.$$

Here, the one-electron operator  $\sum_i (\mathbf{r}_i - \mathbf{R}_a) / |\mathbf{r}_i - \mathbf{R}_a|^3$  is referred to as 'the Hellmann-Feynman' force operator; it is the derivative of the Hamiltonian with respect to displacement of center-a in the x, y, or z direction.

The expressions given above for  $E(\lambda=0)$  and  $(dE/d\lambda)_0$  can once again be used, but with the Hellmann-Feynman form for  $V$ . Once again, for the MCSCF wavefunction, the variational optimization of the energy gives

$$\langle \psi | \sum_j C_j | H^0 | \psi(\lambda=0) \rangle = \langle \psi | \sum_{a,i} C_{a,i} | H^0 | \psi(\lambda=0) \rangle = 0.$$

However, because the atomic basis orbitals are attached to the centers, and because these centers are displaced in forming  $V$ , it is no longer true that  $(d\psi/d\lambda)_0 = 0$ ; the variation in the wavefunction caused by movement of the basis functions now contributes to the first-order energy response. As a result, one obtains

$$(dE/d\lambda)_0 = - \sum_a Z_a e^2 \mathbf{R}_a \cdot \langle \psi | \sum_i (\mathbf{r}_i - \mathbf{R}_a) / |\mathbf{r}_i - \mathbf{R}_a|^3 | \psi \rangle$$

$$+ 2 \sum_a \mathbf{R}_a \cdot \left( \frac{d\psi}{d\lambda} \right)_0 \langle \psi | \sum_i | H^0 | \psi(\lambda=0) \rangle.$$

The first contribution to the force

$$\mathbf{F}_a = - Z_a e^2 \langle \psi | \sum_i (\mathbf{r}_i - \mathbf{R}_a) / |\mathbf{r}_i - \mathbf{R}_a|^3 | \psi \rangle$$

$$+ 2 \langle \mathbf{R}_a \rangle_0 \langle \cdot \cdot | \mathbf{H}^0 | \cdot \cdot = 0 \rangle$$

along the x, y, and z directions for center-a involves the expectation value, with respect to the MCSCF wavefunction with  $\rho = 0$ , of the Hellmann-Feynman force operator. The second contribution gives the forces due to infinitesimal displacements of the basis functions on center-a.

The evaluation of the latter contributions can be carried out by first realizing that

$$\rho = \sum_J C_J | J_1 J_2 J_3 \dots J_n \dots J_N \rangle$$

with

$$j = \sum_{\mu} C_{\mu,j} \mu$$

involves the basis orbitals through the LCAO-MO expansion of the  $\rho_j$ s. So the derivatives of the basis orbitals contribute as follows:

$$\langle \mathbf{R}_a \rangle_0 \langle \cdot \cdot | = \sum_j \rho_j C_J C_{j,j} \langle \cdot \cdot | J_1 J_2 J_3 \dots \mathbf{R}_a \cdot \cdot J_N \rangle$$

Each of these factors can be viewed as combinations of CSFs with the same  $C_J$  and  $C_{j,j}$  coefficients as in  $\rho$  but with the  $j^{\text{th}}$  spin-orbital involving basis functions that have been differentiated with respect to displacement of center-a. It turns out that such derivatives of Gaussian basis orbitals can be carried out analytically (giving rise to new Gaussians with one higher and one lower l-quantum number).

When substituted into  $\langle \mathbf{R}_a \rangle_0 \langle \cdot \cdot | \mathbf{H}^0 | \cdot \cdot = 0 \rangle$ , these basis derivative terms yield

$$\langle \mathbf{R}_a \rangle_0 \langle \cdot \cdot | \mathbf{H}^0 | \cdot \cdot = 0 \rangle = \sum_j \rho_j C_J C_{j,j} \langle \cdot \cdot | J_1 J_2 J_3 \dots \mathbf{R}_a \cdot \cdot J_N | \mathbf{H}^0 | \cdot \cdot \rangle,$$

whose evaluation via the Slater-Condon rules is straightforward. It is simply the expectation value of  $\mathbf{H}^0$  with respect to  $\rho$  (with the same density matrix elements that arise

in the evaluation of  $\psi_a$ 's energy) but with the one- and two-electron integrals over the atomic basis orbitals involving one of these differentiated functions:

$$\begin{aligned} \langle \mu | g | \psi_a \rangle &= \mathbf{R}_a \langle \mu | g | \psi_a \rangle = \langle \mathbf{R}_a \mu | g | \psi_a \rangle \\ &+ \langle \mu | \mathbf{R}_a | g | \psi_a \rangle + \langle \mu | g | \mathbf{R}_a | \psi_a \rangle + \langle \mu | g | \mathbf{R}_a | \psi_a \rangle. \end{aligned}$$

In summary, the force  $\mathbf{F}_a$  felt by the nuclear framework due to a displacement of center-a along the x, y, or z axis is given as

$$\mathbf{F}_a = -Z_a e^2 \langle \psi_a | \sum_i (\mathbf{r}_i - \mathbf{R}_a) / |\mathbf{r}_i - \mathbf{R}_a|^3 | \psi_a \rangle + \langle \mathbf{R}_a | \langle \psi_a | \mathbf{H}^0 | \psi_a \rangle \rangle,$$

where the second term is the energy of  $\psi_a$  but with all atomic integrals replaced by integral derivatives:  $\langle \mu | g | \psi_a \rangle = \mathbf{R}_a \langle \mu | g | \psi_a \rangle$ .

### C. Responses for Other Types of Wavefunctions

It should be stressed that the MCSCF wavefunction yields especially compact expressions for responses of E with respect to an external perturbation because of the variational conditions

$$\langle \psi_a | \sum_J C_J | \mathbf{H}^0 | \psi_a \rangle = 0 \quad \langle \psi_a | \sum_{a,i} C_{a,i} | \mathbf{H}^0 | \psi_a \rangle = 0$$

that apply. The SCF case, which can be viewed as a special case of the MCSCF situation, also admits these simplifications. However, the CI, CC, and MPPT/MBPT cases involve additional factors that arise because the above variational conditions do not apply (in the CI case,  $\langle \psi_a | \sum_J C_J | \mathbf{H}^0 | \psi_a \rangle = 0$  still applies, but the orbital condition  $\langle \psi_a | \sum_{a,i} C_{a,i} | \mathbf{H}^0 | \psi_a \rangle = 0$  does not because the orbitals are not varied to make the CI energy functional stationary).

Within the CC, CI, and MPPT/MBPT methods, one must evaluate the so-called responses of the  $C_J$  and  $C_{a,i}$  coefficients  $(C_J / \lambda)_0$  and  $(C_{a,i} / \lambda)_0$  that appear in the full energy response as (see above)

$$2 \sum_J (C_J / \lambda)_0 \langle \psi_a | \sum_J C_J | \mathbf{H}^0 | \psi_a \rangle + 2 \sum_{i,a} (C_{a,i} / \lambda)_0 \langle \psi_a | \sum_{a,i} C_{a,i} | \mathbf{H}^0 | \psi_a \rangle.$$

To do so requires solving a set of response equations that are obtained by differentiating whatever

equations govern the  $C_I$  and  $C_{a,i}$  coefficients in the particular method (e.g., CI, CC, or MPPT/MBPT) with respect to the external perturbation. In the geometrical derivative case, this amounts to differentiating with respect to x, y, and z displacements of the atomic centers. These response equations are discussed in Geometrical Derivatives of Energy Surfaces and Molecular Properties, P. Jørgensen and J. Simons, Eds., NATO ASI Series, Vol. 166, D. Reidel, Dordrecht (1985). Their treatment is somewhat beyond the scope of this text, so they will not be dealt with further here.

## D. The Use of Geometrical Energy Derivatives

### 1. Gradients as Newtonian Forces

The first energy derivative is called the gradient  $\mathbf{g}$  and is the negative of the force  $\mathbf{F}$  (with components along the  $a^{\text{th}}$  center denoted  $\mathbf{F}_a$ ) experienced by the atomic centers  $\mathbf{F} = -\mathbf{g}$ . These forces, as discussed in Chapter 16, can be used to carry out classical trajectory simulations of molecular collisions or other motions of large organic and biological molecules for which a quantum treatment of the nuclear motion is prohibitive.

The second energy derivatives with respect to the x, y, and z directions of centers a and b (for example, the x, y component for centers a and b is  $H_{ax,by} = (\partial^2 E / \partial x_a \partial y_b)_0$ ) form the Hessian matrix  $\mathbf{H}$ . The elements of  $\mathbf{H}$  give the local curvatures of the energy surface along the 3N cartesian directions.

The gradient and Hessian can be used to systematically locate local minima (i.e., stable geometries) and transition states that connect one local minimum to another. At each of these stationary points, all forces and thus all elements of the gradient  $\mathbf{g}$  vanish. At a local minimum, the  $\mathbf{H}$  matrix has 5 or 6 zero eigenvalues corresponding to translational and rotational displacements of the molecule (5 for linear molecules; 6 for non-linear species) and  $3N-5$  or  $3N-6$  positive eigenvalues. At a transition state,  $\mathbf{H}$  has one negative eigenvalue, 5 or 6 zero eigenvalues, and  $3N-6$  or  $3N-7$  positive eigenvalues.

### 2. Transition State Rate Coefficients

The transition state theory of Eyring or its extensions due to Truhlar and co-workers (see, for example, D. G. Truhlar and B. C. Garrett, *Ann. Rev. Phys. Chem.* **35**, 159 (1984)) allow knowledge of the Hessian matrix at a transition state to be used to compute a rate coefficient  $k_{\text{rate}}$  appropriate to the chemical reaction for which the transition state applies.

More specifically, the geometry of the molecule at the transition state is used to compute a rotational partition function  $Q_{\text{rot}}^\ddagger$  in which the principal moments of inertia  $I_a$ ,  $I_b$ , and  $I_c$  (see Chapter 13) are those of the transition state (the  $\ddagger$  symbol is, by convention, used to label the transition state):

$$Q_{\text{rot}}^\ddagger = \prod_{n=a,b,c} \sqrt{\frac{8\pi^2 I_n kT}{h^2}},$$

where  $k$  is the Boltzmann constant and  $T$  is the temperature in  $^\circ\text{K}$ .

The eigenvalues  $\{\nu_i\}$  of the mass weighted Hessian matrix (see below) are used to compute, for each of the  $3N-7$  vibrations with real and positive  $\nu_i$  values, a vibrational partition function that is combined to produce a transition-state vibrational partition function:

$$Q_{\text{vib}}^\ddagger = \prod_{i=1}^{3N-7} \frac{\exp(-h\nu_i/2kT)}{1 - \exp(-h\nu_i/kT)}.$$

The electronic partition function of the transition state is expressed in terms of the activation energy (the energy of the transition state relative to the electronic energy of the reactants)  $E^\ddagger$  as:

$$Q_{\text{electronic}}^\ddagger = g^\ddagger \exp(-E^\ddagger/kT)$$

where  $g^\ddagger$  is the degeneracy of the electronic state at the transition state geometry.

In the original Eyring version of transition state theory (TST), the rate coefficient  $k_{\text{rate}}$  is then given by:

$$k_{\text{rate}} = \frac{kT}{h} g^\ddagger \exp(-E^\ddagger/kT) \frac{Q_{\text{rot}}^\ddagger Q_{\text{vib}}^\ddagger}{Q_{\text{reactants}}},$$

where  $Q_{\text{reactants}}$  is the conventional partition function for the reactant materials.

For example, in a bimolecular reaction such as:



the reactant partition function

$$Q_{\text{reactants}} = Q_{\text{F}} Q_{\text{H}_2}$$

is written in terms of the translational and electronic (the degeneracy of the  $^2\text{P}$  state produces the 2 (3) overall degeneracy factor) partition functions of the F atom

$$Q_{\text{F}} = \frac{2}{h^2} \frac{m_{\text{F}} k T}{h^2}^{3/2} \cdot 2 \quad (3)$$

and the translational, electronic, rotational, and vibrational partition functions of the  $\text{H}_2$  molecule

$$Q_{\text{H}_2} = \frac{2}{h^2} \frac{m_{\text{H}_2} k T}{h^2}^{3/2} \frac{8}{2h^2} \frac{\exp(-\frac{h}{2} \frac{H_2}{2kT})}{1 - \exp(-\frac{h}{H_2} / kT)}$$

The factor of 2 in the denominator of the  $\text{H}_2$  molecule's rotational partition function is the "symmetry number" that must be inserted because of the identity of the two H nuclei.

The overall rate coefficient  $k_{\text{rate}}$  (with units  $\text{sec}^{-1}$  because this is a rate per collision pair) can thus be expressed entirely in terms of energetic, geometrical, and vibrational information about the reactants and the transition state. Even within the extensions to Eyring's original model, such is the case. The primary difference in the more modern theories is that the transition state is identified not as the point on the potential energy surface at which the gradient vanishes and there is one negative Hessian eigenvalue. Instead, a so-called variational transition state (see the above reference by Truhlar and Garrett) is identified. The geometry, energy, and local vibrational frequencies of this transition state are then used to compute, must like outlined above,  $k_{\text{rate}}$ .

### 3. Harmonic Vibrational Frequencies

It is possible (see, for example, J. Nichols, H. L. Taylor, P. Schmidt, and J. Simons, *J. Chem. Phys.* 92, 340 (1990) and references therein) to remove from  $\mathbf{H}$  the zero eigenvalues that correspond to rotation and translation and to thereby produce a Hessian matrix whose eigenvalues correspond only to internal motions of the system. After doing so, the number of negative eigenvalues of  $\mathbf{H}$  can be used to characterize the nature of the

stationary point (local minimum or transition state), and  $\mathbf{H}$  can be used to evaluate the local harmonic vibrational frequencies of the system.

The relationship between  $\mathbf{H}$  and vibrational frequencies can be made clear by recalling the classical equations of motion in the Lagrangian formulation:

$$d/dt( L/ \dot{q}_j ) - ( L/ q_j ) = 0,$$

where  $q_j$  denotes, in our case, the  $3N$  cartesian coordinates of the  $N$  atoms, and  $\dot{q}_j$  is the velocity of the corresponding coordinate. Expressing the Lagrangian  $L$  as kinetic energy minus potential energy and writing the potential energy as a local quadratic expansion about a point where  $\mathbf{g}$  vanishes, gives

$$L = 1/2 \sum_j m_j \dot{q}_j^2 - E(0) - 1/2 \sum_{j,k} q_j H_{j,k} q_k .$$

Here,  $E(0)$  is the energy at the stationary point,  $m_j$  is the mass of the atom to which  $q_j$  applies, and the  $H_{j,k}$  are the elements of  $\mathbf{H}$  along the  $x$ ,  $y$ , and  $z$  directions of the various atomic centers.

Applying the Lagrangian equations to this form for  $L$  gives the equations of motion of the  $q_j$  coordinates:

$$m_j \ddot{q}_j = - \sum_k H_{j,k} q_k .$$

To find solutions that correspond to local harmonic motion, one assumes that the coordinates  $q_j$  oscillate in time according to

$$q_j(t) = q_j \cos( \omega t ) .$$

Substituting this form for  $q_j(t)$  into the equations of motion gives

$$m_j \omega^2 q_j = - \sum_k H_{j,k} q_k .$$

Defining

$$q_j' = q_j (m_j)^{1/2}$$

and introducing this into the above equation of motion yields

$$\ddot{q}_j' = - \sum_k H'_{j,k} q_k' ,$$

where

$$H'_{j,k} = H_{j,k} (m_j m_k)^{-1/2}$$

is the so-called mass-weighted Hessian matrix.

The squares of the desired harmonic vibrational frequencies  $\omega_j^2$  are thus given as eigenvalues of the mass-weighted Hessian  $\mathbf{H}'$ :

$$\mathbf{H}' \mathbf{q}' = - \omega^2 \mathbf{q}'$$

The corresponding eigenvector,  $\{q'_j\}$  gives, when multiplied by  $m_j^{-1/2}$ , the atomic displacements that accompany that particular harmonic vibration. At a transition state, one of the  $\omega^2$  will be negative and  $3N-6$  or  $3N-7$  will be positive.

#### 4. Reaction Path Following

The Hessian and gradient can also be used to trace out 'streambeds' connecting local minima to transition states. In doing so, one utilizes a local harmonic description of the potential energy surface

$$E(\mathbf{x}) = E(\mathbf{0}) + \mathbf{x} \cdot \mathbf{g} + 1/2 \mathbf{x} \cdot \mathbf{H} \cdot \mathbf{x} + \dots,$$

where  $\mathbf{x}$  represents the (small) step away from the point  $\mathbf{x}=\mathbf{0}$  at which the gradient  $\mathbf{g}$  and Hessian  $\mathbf{H}$  have been evaluated. By expressing  $\mathbf{x}$  and  $\mathbf{g}$  in terms of the eigenvectors  $\mathbf{v}$  of  $\mathbf{H}$

$$\mathbf{H} \mathbf{v} = - \omega^2 \mathbf{v} ,$$

$$\mathbf{x} = \sum_i \langle \mathbf{v}_i | \mathbf{x} \rangle \mathbf{v}_i = \sum_i x_i \mathbf{v}_i ,$$

$$\mathbf{g} = \sum_i \langle \mathbf{v}_i | \mathbf{g} \rangle \mathbf{v}_i = \sum_i g_i \mathbf{v}_i ,$$

the energy change  $E(\mathbf{x}) - E(\mathbf{0})$  can be expressed in terms of a sum of independent changes along the eigendirections:

$$E(\mathbf{x}) - E(\mathbf{0}) = \left[ x_i g_i + \frac{1}{2} x_i^2 \right] + \dots$$

Depending on the signs of  $g_i$  and of  $\omega_i$ , various choices for the displacements  $x_i$  will produce increases or decreases in energy:

1. If  $\omega_i$  is positive, then a step  $x_i$  'along'  $g_i$  (i.e., one with  $x_i g_i$  positive) will generate an energy increase. A step 'opposed to'  $g_i$  will generate an energy decrease if it is short enough that  $x_i g_i$  is larger in magnitude than  $\frac{1}{2} x_i^2 \omega_i$ , otherwise the energy will increase.
2. If  $\omega_i$  is negative, a step opposed to  $g_i$  will generate an energy decrease. A step along  $g_i$  will give an energy increase if it is short enough for  $x_i g_i$  to be larger in magnitude than  $\frac{1}{2} x_i^2 \omega_i$ , otherwise the energy will decrease.

Thus, to proceed downhill in all directions (such as one wants to do when searching for local minima), one chooses each  $x_i$  in opposition to  $g_i$  and of small enough length to guarantee that the magnitude of  $x_i g_i$  exceeds that of  $\frac{1}{2} x_i^2 \omega_i$  for those modes with  $\omega_i > 0$ . To proceed uphill along a mode with  $\omega_i < 0$  and downhill along all other modes with  $\omega_i > 0$ , one chooses  $x_i$  'along'  $g_i$  with  $x_i$  short enough to guarantee that  $x_i g_i$  is larger in magnitude than  $\frac{1}{2} x_i^2 \omega_i$ , and one chooses the other  $x_j$  opposed to  $g_j$  and short enough that  $x_j g_j$  is larger in magnitude than  $\frac{1}{2} x_j^2 \omega_j$ .

Such considerations have allowed the development of highly efficient potential energy surface 'walking' algorithms (see, for example, J. Nichols, H. L. Taylor, P. Schmidt, and J. Simons, *J. Chem. Phys.* 92, 340 (1990) and references therein) designed to trace out streambeds and to locate and characterize, via the local harmonic frequencies, minima and transition states. These algorithms form essential components of most modern ab initio, semi-empirical, and empirical computational chemistry software packages.

## II. Ab Initio, Semi-Empirical and Empirical Force Field Methods

### A. *Ab Initio* Methods

Most of the techniques described in this Chapter are of the ab initio type. This means that they attempt to compute electronic state energies and other physical properties, as functions of the positions of the nuclei, from first principles without the use or knowledge of experimental input. Although perturbation theory or the variational method may be used to generate the working equations of a particular method, and although finite atomic orbital basis sets are nearly always utilized, these approximations do not involve 'fitting' to known experimental data. They represent approximations that can be systematically improved as the level of treatment is enhanced.

### B. Semi-Empirical and Fully Empirical Methods

Semi-empirical methods, such as those outlined in Appendix F, use experimental data or the results of ab initio calculations to determine some of the matrix elements or integrals needed to carry out their procedures. Totally empirical methods attempt to describe the internal electronic energy of a system as a function of geometrical degrees of freedom (e.g., bond lengths and angles) in terms of analytical 'force fields' whose parameters have been determined to 'fit' known experimental data on some class of compounds. Examples of such parameterized force fields were presented in Section III. A of Chapter 16.

### C. Strengths and Weaknesses

Each of these tools has advantages and limitations. Ab initio methods involve intensive computation and therefore tend to be limited, for practical reasons of computer time, to smaller atoms, molecules, radicals, and ions. Their CPU time needs usually vary with basis set size ( $M$ ) as at least  $M^4$ ; correlated methods require time proportional to at least  $M^5$  because they involve transformation of the atomic-orbital-based two-electron integrals to the molecular orbital basis. As computers continue to advance in power and memory size, and as theoretical methods and algorithms continue to improve, ab initio techniques will be applied to larger and more complex species. When dealing with systems in which qualitatively new electronic environments and/or new bonding types arise, or excited electronic states that are unusual, ab initio methods are essential. Semi-empirical or empirical methods would be of little use on systems whose electronic properties have not been included in the data base used to construct the parameters of such models.

On the other hand, to determine the stable geometries of large molecules that are made of conventional chemical units (e.g., CC, CH, CO, etc. bonds and steric and torsional interactions among same), fully empirical force-field methods are usually quite reliable and computationally very fast. Stable geometries and the relative energetic stabilities of various conformers of large macromolecules and biopolymers can routinely be predicted using such tools if the system contains only conventional bonding and common chemical building blocks. These empirical potentials usually do not contain sufficient flexibility (i.e., their parameters and input data do not include enough knowledge) to address processes that involve rearrangement of the electronic configurations. For example, they can not treat:

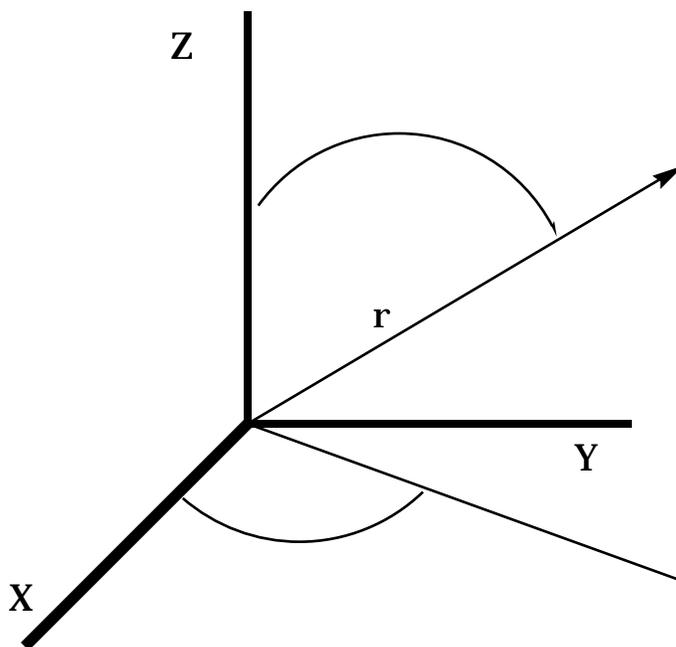
1. Electronic transitions, because knowledge of the optical oscillator strengths and of the energies of excited states is absent in most such methods;
2. Concerted chemical reactions involving simultaneous bond breaking and forming, because to do so would require the force-field parameters to evolve from those of the reactant bonding to those for the product bonding as the reaction proceeds;
3. Molecular properties such as dipole moment and polarizability, although in certain fully empirical models, bond dipoles and lone-pair contributions have been incorporated (although again only for conventional chemical bonding situations).

Semi-empirical techniques share some of the strengths and weaknesses of ab initio and of fully empirical methods. They treat at least the valence electrons explicitly, so they are able to address questions that are inherently electronic such as electronic transitions, dipole moments, polarizability, and bond breaking and forming. Some of the integrals involving the Hamiltonian operator and the atomic basis orbitals are performed ab initio; others are obtained by fitting to experimental data. The computational needs of semi-empirical methods lie between those of the ab initio methods and the force-field techniques. As with the empirical methods, they should never be employed when qualitatively new electronic bonding situations are encountered because the data base upon which their parameters were determined contain, by assumption, no similar bonding cases.

## Section 1 Exercises, Problems, and Solutions

### Review Exercises

1. Transform (using the coordinate system provided below) the following functions accordingly:



a. from cartesian to spherical polar coordinates

$$3x + y - 4z = 12$$

b. from cartesian to cylindrical coordinates

$$y^2 + z^2 = 9$$

c. from spherical polar to cartesian coordinates

$$r = 2 \sin \theta \cos \phi$$

2. Perform a separation of variables and indicate the general solution for the following expressions:

a.  $9x + 16y \frac{y}{x} = 0$

b.  $2y + \frac{y}{x} + 6 = 0$

3. Find the eigenvalues and corresponding eigenvectors of the following matrices:

a.  $\begin{bmatrix} -1 & 2 \\ 2 & 2 \end{bmatrix}$

b.  $\begin{bmatrix} -2 & 0 & 0 \\ 0 & -1 & 2 \\ 0 & 2 & 2 \end{bmatrix}$

4. For the hermitian matrix in review exercise 3a show that the eigenfunctions can be normalized and that they are orthogonal.
5. For the hermitian matrix in review exercise 3b show that the pair of degenerate eigenvalues can be made to have orthonormal eigenfunctions.
6. Solve the following second order linear differential equation subject to the specified "boundary conditions":

$$\frac{d^2x}{dt^2} + k^2x(t) = 0, \text{ where } x(t=0) = L, \text{ and } \frac{dx(t=0)}{dt} = 0.$$

### Exercises

1. Replace the following classical mechanical expressions with their corresponding quantum mechanical operators.

- K.E. =  $\frac{mv^2}{2}$  in three-dimensional space.
- $\mathbf{p} = m\mathbf{v}$ , a three-dimensional cartesian vector.
- y-component of angular momentum:  $L_y = zp_x - xp_z$ .

2. Transform the following operators into the specified coordinates:

- $L_x = \frac{\hbar}{i} \left( y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right)$  from cartesian to spherical polar coordinates.
- $L_z = \frac{\hbar}{i} \frac{\partial}{\partial \phi}$  from spherical polar to cartesian coordinates.

3. Match the eigenfunctions in column B to their operators in column A. What is the eigenvalue for each eigenfunction?

<u>Column A</u>	<u>Column B</u>
i. $(1-x^2) \frac{d^2}{dx^2} - x \frac{d}{dx}$	$4x^4 - 12x^2 + 3$
ii. $\frac{d^2}{dx^2}$	$5x^4$
iii. $x \frac{d}{dx}$	$e^{3x} + e^{-3x}$
iv. $\frac{d^2}{dx^2} - 2x \frac{d}{dx}$	$x^2 - 4x + 2$
v. $x \frac{d^2}{dx^2} + (1-x) \frac{d}{dx}$	$4x^3 - 3x$

4. Show that the following operators are hermitian.

- $\mathbf{P}_x$
- $\mathbf{L}_x$

5. For the following basis of functions ( ${}_2p_{-1}$ ,  ${}_2p_0$ , and  ${}_2p_{+1}$ ), construct the matrix representation of the  $\mathbf{L}_x$  operator (use the ladder operator representation of  $\mathbf{L}_x$ ). Verify that

the matrix is hermitian. Find the eigenvalues and corresponding eigenvectors. Normalize the eigenfunctions and verify that they are orthogonal.

$$2p_{-1} = \frac{1}{8} \frac{Z}{a^{1/2}} \frac{Z}{a}^{5/2} r e^{-zr/2a} \sin e^{-i}$$

$$2p_0 = \frac{1}{1/2} \frac{Z}{2a} \frac{Z}{a}^{5/2} r e^{-zr/2a} \cos$$

$$2p_1 = \frac{1}{8} \frac{Z}{1/2} \frac{Z}{a} \frac{Z}{a}^{5/2} r e^{-zr/2a} \sin e^i$$

6. Using the set of eigenstates (with corresponding eigenvalues) from the preceding problem, determine the probability for observing a z-component of angular momentum equal to  $1\hbar$  if the state is given by the  $L_x$  eigenstate with  $0\hbar$   $L_x$  eigenvalue.

7. Use the following definitions of the angular momentum operators:

$$L_x = \frac{\hbar}{i} \left( y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right), \quad L_y = \frac{\hbar}{i} \left( z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right),$$

$$L_z = \frac{\hbar}{i} \left( x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right), \quad \text{and } L^2 = L_x^2 + L_y^2 + L_z^2,$$

and the relationships:

$[x, p_x] = i\hbar$ ,  $[y, p_y] = i\hbar$ , and  $[z, p_z] = i\hbar$ ,  
to demonstrate the following operator identities:

- $[L_x, L_y] = i\hbar L_z$ ,
- $[L_y, L_z] = i\hbar L_x$ ,
- $[L_z, L_x] = i\hbar L_y$ ,
- $[L_x, L^2] = 0$ ,
- $[L_y, L^2] = 0$ ,
- $[L_z, L^2] = 0$ .

8. In exercise 7 above you determined whether or not many of the angular momentum operators commute. Now, examine the operators below along with an appropriate given function. Determine if the given function is simultaneously an eigenfunction of both operators. Is this what you expected?

a.  $L_z, L^2$ , with function:  $Y_0^0(\theta, \phi) = \frac{1}{\sqrt{4}}$ .

b.  $L_x, L_z$ , with function:  $Y_0^0(\theta, \phi) = \frac{1}{\sqrt{4}}$ .

c.  $L_z, L^2$ , with function:  $Y_1^0(\theta, \phi) = \sqrt{\frac{3}{4}} \cos \theta$ .

d.  $L_x, L_z$ , with function:  $Y_1^0(\theta, \phi) = \sqrt{\frac{3}{4}} \cos \theta$ .

9. For a "particle in a box" constrained along two axes, the wavefunction  $\psi(x,y)$  as given in the text was :

$$\psi(x,y) = \frac{1}{2L_x} \frac{1}{2} \frac{1}{2L_y} \frac{1}{2} e^{\frac{in_x x}{L_x}} - e^{-\frac{in_x x}{L_x}} e^{\frac{in_y y}{L_y}} - e^{-\frac{in_y y}{L_y}},$$

with  $n_x$  and  $n_y = 1,2,3, \dots$ . Show that this wavefunction is normalized.

10. Using the same wavefunction,  $\psi(x,y)$ , given in exercise 9 show that the expectation value of  $\mathbf{p}_x$  vanishes.

11. Calculate the expectation value of the  $\mathbf{x}^2$  operator for the first two states of the harmonic oscillator. Use the  $v=0$  and  $v=1$  harmonic oscillator wavefunctions given below

$$\psi_0 = \frac{1}{\pi^{1/4}} e^{-x^2/2} \quad \text{and} \quad \psi_1 = \frac{\sqrt{2}}{\pi^{1/4}} x e^{-x^2/2}$$

which are normalized such that  $\int_{-\infty}^{\infty} \psi(x)^2 dx = 1$ . Remember that  $\int_{-\infty}^{\infty} x e^{-x^2/2} dx = 0$  and  $\int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx = \sqrt{2\pi}$ .

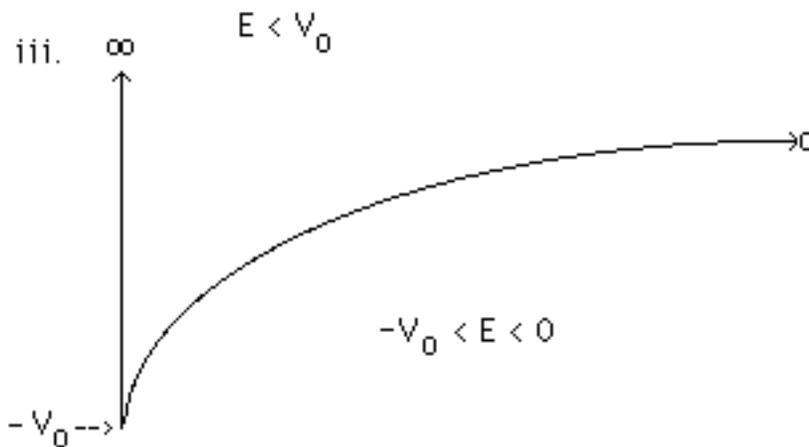
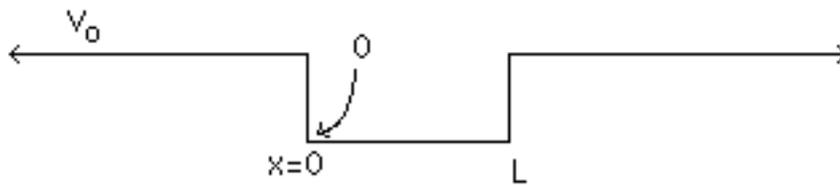
12. For each of the one-dimensional potential energy graphs shown below, determine:  
 a. whether you expect symmetry to lead to a separation into odd and even solutions,  
 b. whether you expect the energy will be quantized, continuous, or both, and  
 c. the boundary conditions that apply at each boundary (merely stating that

and/or  $\psi$  is continuous is all that is necessary).

i.



ii.



13. Consider a particle of mass  $m$  moving in the potential:

$$\begin{array}{lll}
 V(x) = \infty & \text{for} & x < 0 & \text{Region I} \\
 V(x) = 0 & \text{for} & 0 < x < L & \text{Region II} \\
 V(x) = V_0 \ (V_0 > 0) & \text{for} & x > L & \text{Region III}
 \end{array}$$

a. Write the general solution to the Schrödinger equation for the regions I, II, III, assuming a solution with energy  $E < V_0$  (i.e. a bound state).

b. Write down the wavefunction matching conditions at the interface between regions I and II and between II and III.

c. Write down the boundary conditions on  $\psi$  for  $x \rightarrow \pm \infty$ .

d. Use your answers to a. - c. to obtain an algebraic equation which must be satisfied for the bound state energies,  $E$ .

e. Demonstrate that in the limit  $V \rightarrow \infty$ , the equation you obtained for the bound state energies in d. gives the energies of a particle in an infinite box;  $E_n = \frac{n^2 \hbar^2 \pi^2}{2mL^2}$ ;  $n = 1, 2, 3, \dots$

### Problems

1. A particle of mass  $m$  moves in a one-dimensional box of length  $L$ , with boundaries at  $x = 0$  and  $x = L$ . Thus,  $V(x) = 0$  for  $0 < x < L$ , and  $V(x) = \infty$  elsewhere. The normalized eigenfunctions of the Hamiltonian for this system are given by  $\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right)$ , with

$E_n = \frac{n^2 \pi^2 \hbar^2}{2mL^2}$ , where the quantum number  $n$  can take on the values  $n=1, 2, 3, \dots$

a. Assuming that the particle is in an eigenstate,  $\psi_n(x)$ , calculate the probability that the particle is found somewhere in the region  $0 < x < \frac{L}{4}$ . Show how this probability depends on  $n$ .

b. For what value of  $n$  is there the largest probability of finding the particle in  $0 < x < \frac{L}{4}$ ?

c. Now assume that  $\psi$  is a superposition of two eigenstates,  $\psi = a \psi_n + b \psi_m$ , at time  $t = 0$ . What is  $\psi$  at time  $t$ ? What energy expectation value does  $\psi$  have at time  $t$  and how does this relate to its value at  $t = 0$ ?

d. For an experimental measurement which is capable of distinguishing systems in state  $\psi_n$  from those in  $\psi_m$ , what fraction of a large number of systems each described by  $\psi$  will be observed to be in  $\psi_n$ ? What energies will these experimental measurements find and with what probabilities?

e. For those systems originally in  $\psi = a \psi_n + b \psi_m$  which were observed to be in  $\psi_n$  at time  $t$ , what state ( $\psi_n$ ,  $\psi_m$ , or whatever) will they be found in if a second experimental measurement is made at a time  $t'$  later than  $t$ ?

f. Suppose by some method (which need not concern us at this time) the system has been prepared in a nonstationary state (that is, it is not an eigenfunction of  $\mathbf{H}$ ). At the time of a measurement of the particle's energy, this state is specified by the normalized

wavefunction  $\psi = \frac{30}{L^5} x(L-x)$  for  $0 < x < L$ , and  $\psi = 0$  elsewhere. What is the

probability that a measurement of the energy of the particle will give the value  $E_n = \frac{n^2 \pi^2 \hbar^2}{2mL^2}$  for any given value of  $n$ ?

g. What is the expectation value of  $\mathbf{H}$ , i.e. the average energy of the system, for the wavefunction  $\psi$  given in part f?

2. Show that for a system in a non-stationary state,

$$= \sum_j C_j e^{-iE_j t/\hbar}$$
 , the average value of the energy does not vary with time but the expectation values of other properties do vary with time.

3. A particle is confined to a one-dimensional box of length  $L$  having infinitely high walls and is in its lowest quantum state. Calculate:  $\langle x \rangle$ ,  $\langle x^2 \rangle$ ,  $\langle p \rangle$ , and  $\langle p^2 \rangle$ . Using the definition  $\Delta A = (\langle A^2 \rangle - \langle A \rangle^2)^{1/2}$ , to define the uncertainty,  $\Delta A$ , calculate  $\Delta x$  and  $\Delta p$ . Verify the Heisenberg uncertainty principle that  $\Delta x \Delta p \geq \hbar/2$ .

4. It has been claimed that as the quantum number  $n$  increases, the motion of a particle in a box becomes more classical. In this problem you will have an opportunity to convince yourself of this fact.

a. For a particle of mass  $m$  moving in a one-dimensional box of length  $L$ , with ends of the box located at  $x = 0$  and  $x = L$ , the classical probability density can be shown to be independent of  $x$  and given by  $P(x)dx = \frac{dx}{L}$  regardless of the energy of the particle. Using this probability density, evaluate the probability that the particle will be found within the interval from  $x = 0$  to  $x = \frac{L}{4}$ .

b. Now consider the quantum mechanical particle-in-a-box system. Evaluate the probability of finding the particle in the interval from  $x = 0$  to  $x = \frac{L}{4}$  for the system in its  $n$ th quantum state.

c. Take the limit of the result you obtained in part b as  $n \rightarrow \infty$ . How does your result compare to the classical result you obtained in part a?

5. According to the rules of quantum mechanics as we have developed them, if  $\psi$  is the state function, and  $\psi_n$  are the eigenfunctions of a linear, Hermitian operator,  $\mathbf{A}$ , with eigenvalues  $a_n$ ,  $\mathbf{A} \psi_n = a_n \psi_n$ , then we can expand  $\psi$  in terms of the complete set of eigenfunctions of  $\mathbf{A}$  according to  $\psi = \sum_n c_n \psi_n$ , where  $c_n = \int \psi_n^* \psi dx$ . Furthermore, the

probability of making a measurement of the property corresponding to  $\mathbf{A}$  and obtaining a value  $a_n$  is given by  $|c_n|^2$ , provided both  $\psi$  and  $\psi_n$  are properly normalized. Thus,  $P(a_n) = |c_n|^2$ . These rules are perfectly valid for operators which take on a discrete set of eigenvalues, but must be generalized for operators which can have a continuum of eigenvalues. An example of this latter type of operator is the momentum operator,  $\mathbf{p}_x$ ,

which has eigenfunctions given by  $\psi_p(x) = A e^{ipx/\hbar}$  where  $p$  is the eigenvalue of the  $\mathbf{p}_x$  operator and  $A$  is a normalization constant. Here  $p$  can take on any value, so we have a continuous spectrum of eigenvalues of  $\mathbf{p}_x$ . The obvious generalization to the equation for

is to convert the sum over discrete states to an integral over the continuous spectrum of states:

$$\psi(x) = \int_{-\infty}^{+\infty} C(p) \psi_p(x) dp = \int_{-\infty}^{+\infty} C(p) A e^{ipx/\hbar} dp$$

The interpretation of  $C(p)$  is now the desired generalization of the equation for the probability  $P(p)dp = C(p)^2 dp$ . This equation states that the probability of measuring the momentum and finding it in the range from  $p$  to  $p+dp$  is given by  $C(p)^2 dp$ . Accordingly, the probability of measuring  $p$  and finding it in the range from  $p_1$  to  $p_2$  is given by

$P(p)dp = C(p)^* C(p) dp$ .  $C(p)$  is thus the probability amplitude for finding the particle with momentum between  $p$  and  $p+dp$ . This is the momentum representation of the

wavefunction. Clearly we must require  $C(p)$  to be normalized, so that  $\int C(p)^* C(p) dp = 1$ .

With this restriction we can derive the normalization constant  $A = \frac{1}{\sqrt{2\pi\hbar}}$ , giving a direct

relationship between the wavefunction in coordinate space,  $\psi(x)$ , and the wavefunction in momentum space,  $C(p)$ :

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int C(p) e^{ipx/\hbar} dp,$$

and by the fourier integral theorem:

$$C(p) = \frac{1}{\sqrt{2\pi\hbar}} \int \psi(x) e^{-ipx/\hbar} dx.$$

Lets use these ideas to solve some problems focusing our attention on the harmonic oscillator; a particle of mass  $m$  moving in a one-dimensional potential described by  $V(x) = \frac{kx^2}{2}$ .

a. Write down the Schrödinger equation in the coordinate representation.

b. Now lets proceed by attempting to write the Schrödinger equation in the momentum representation. Identifying the kinetic energy operator  $\mathbf{T}$ , in the momentum

representation is quite straightforward  $\mathbf{T} = \frac{\mathbf{p}^2}{2m}$ .

**Error!** Writing the potential,  $V(x)$ , in the momentum representation is not quite as straightforward. The relationship between position and momentum is realized in their

commutation relation  $[\mathbf{x}, \mathbf{p}] = i\hbar$ , or  $(\mathbf{x}\mathbf{p} - \mathbf{p}\mathbf{x}) = i\hbar$

This commutation relation is easily verified in the coordinate representation leaving  $\mathbf{x}$  untouched ( $\mathbf{x} = x \cdot$ ) and using the above definition for  $\mathbf{p}$ . In the momentum representation we want to leave  $\mathbf{p}$  untouched ( $\mathbf{p} = p \cdot$ ) and define the operator  $\mathbf{x}$  in such a manner that the commutation relation is still satisfied. Write the operator  $\mathbf{x}$  in the momentum representation. Write the full Hamiltonian in the momentum representation and hence the Schrödinger equation in the momentum representation.

c. Verify that as given below is an eigenfunction of the Hamiltonian in the coordinate representation. What is the energy of the system when it is in this state?

Determine the normalization constant C, and write down the normalized ground state wavefunction in coordinate space.

$$\psi(x) = C \exp\left(-\sqrt{mk} \frac{x^2}{2\hbar}\right).$$

d. Now consider  $\psi$  in the momentum representation. Assuming that an eigenfunction of the Hamiltonian may be found of the form  $\psi(p) = C \exp(-p^2)$ , substitute this form of  $\psi$  into the Schrödinger equation in the momentum representation to find the value of  $k$  which makes this an eigenfunction of  $\mathbf{H}$  having the same energy as  $\psi(x)$  had. Show that this  $\psi(p)$  is the proper Fourier transform of  $\psi(x)$ . The following integral may be useful:

$$\int_{-\infty}^{+\infty} e^{-x^2} \cos bx \, dx = \sqrt{\frac{\pi}{1+b^2}} e^{-b^2/4}.$$

Since this Hamiltonian has no degenerate states, you may conclude that  $\psi(x)$  and  $\psi(p)$  represent the same state of the system if they have the same energy.

6. The energy states and wavefunctions for a particle in a 3-dimensional box whose lengths are  $L_1$ ,  $L_2$ , and  $L_3$  are given by

$$E(n_1, n_2, n_3) = \frac{\hbar^2}{8m} \left( \frac{n_1^2}{L_1^2} + \frac{n_2^2}{L_2^2} + \frac{n_3^2}{L_3^2} \right) \quad \text{and}$$

$$\psi(n_1, n_2, n_3) = \frac{2}{L_1} \frac{1}{2} \frac{2}{L_2} \frac{1}{2} \frac{2}{L_3} \frac{1}{2} \sin \frac{n_1 x}{L_1} \sin \frac{n_2 y}{L_2} \sin \frac{n_3 z}{L_3}.$$

These wavefunctions and energy levels are sometimes used to model the motion of electrons in a central metal atom (or ion) which is surrounded by six ligands.

a. Show that the lowest energy level is nondegenerate and the second energy level is triply degenerate if  $L_1 = L_2 = L_3$ . What values of  $n_1$ ,  $n_2$ , and  $n_3$  characterize the states belonging to the triply degenerate level?

b. For a box of volume  $V = L_1 L_2 L_3$ , show that for three electrons in the box (two in the nondegenerate lowest "orbital", and one in the next), a lower total energy will result if the box undergoes a rectangular distortion ( $L_1 = L_2 \neq L_3$ ) which preserves the total volume than if the box remains undistorted (hint: if  $V$  is fixed and  $L_1 = L_2$ , then  $L_3 = \frac{V}{L_1^2}$  and  $L_1$  is the only "variable").

c. Show that the degree of distortion (ratio of  $L_3$  to  $L_1$ ) which will minimize the total energy is  $L_3 = \sqrt{2} L_1$ . How does this problem relate to Jahn-Teller distortions? Why (in terms of the property of the central atom or ion) do we do the calculation with fixed volume?

d. By how much (in eV) will distortion lower the energy (from its value for a cube,  $L_1 = L_2 = L_3$ ) if  $V = 8 \text{ \AA}^3$  and  $\frac{\hbar^2}{8m} = 6.01 \times 10^{-27} \text{ erg cm}^2$ .  $1 \text{ eV} = 1.6 \times 10^{-12} \text{ erg}$

7. The wavefunction  $\psi = A e^{-a|x|}$  is an exact eigenfunction of some one-dimensional Schrödinger equation in which  $x$  varies from  $-\infty$  to  $+\infty$ . The value of  $a$  is:  $a = (2 \text{ \AA})^{-1}$ . For

now, the potential  $V(x)$  in the Hamiltonian ( $\mathbf{H} = -\frac{\hbar}{2m} \frac{d^2}{dx^2} + V(x)$ ) for which  $\psi(x)$  is an eigenfunction is unknown.

a. Find a value of  $A$  which makes  $\psi(x)$  normalized. Is this value unique? What units does  $\psi(x)$  have?

b. Sketch the wavefunction for positive and negative values of  $x$ , being careful to show the behavior of its slope near  $x = 0$ . Recall that  $|x|$  is defined as:

$$|x| = \begin{cases} x & \text{if } x > 0 \\ -x & \text{if } x < 0 \end{cases}$$

c. Show that the derivative of  $\psi(x)$  undergoes a discontinuity of magnitude  $2(a)^{3/2}$  as  $x$  goes through  $x = 0$ . What does this fact tell you about the potential  $V(x)$ ?

d. Calculate the expectation value of  $|x|$  for the above normalized wavefunction (obtain a numerical value and give its units). What does this expectation value give a measure of?

e. The potential  $V(x)$  appearing in the Schrödinger equation for which  $\psi = Ae^{-a|x|}$  is an exact solution is given by  $V(x) = \frac{\hbar^2 a}{m} \psi(x)$ . Using this potential, compute the

expectation value of the Hamiltonian ( $\mathbf{H} = -\frac{\hbar}{2m} \frac{d^2}{dx^2} + V(x)$ ) for your normalized wavefunction. Is  $V(x)$  an attractive or repulsive potential? Does your wavefunction correspond to a bound state? Is  $\langle H \rangle$  negative or positive? What does the sign of  $\langle H \rangle$  tell you? To obtain a numerical value for  $\langle H \rangle$  use  $\frac{\hbar^2}{2m} = 6.06 \times 10^{-28} \text{ erg cm}^2$  and  $1\text{eV} = 1.6 \times 10^{-12} \text{ erg}$ .

f. Transform the wavefunction,  $\psi = Ae^{-a|x|}$ , from coordinate space to momentum space.

g. What is the ratio of the probability of observing a momentum equal to  $2a\hbar$  to the probability of observing a momentum equal to  $-a\hbar$ ?

8. The  $\pi$ -orbitals of benzene,  $C_6H_6$ , may be modeled very crudely using the wavefunctions and energies of a particle on a ring. Lets first treat the particle on a ring problem and then extend it to the benzene system.

a. Suppose that a particle of mass  $m$  is constrained to move on a circle (of radius  $r$ ) in the  $xy$  plane. Further assume that the particle's potential energy is constant (zero is a good choice). Write down the Schrödinger equation in the normal cartesian coordinate representation. Transform this Schrödinger equation to cylindrical coordinates where  $x = r \cos \theta$ ,  $y = r \sin \theta$ , and  $z = z$  ( $z = 0$  in this case).

Taking  $r$  to be held constant, write down the general solution,  $\psi(\theta)$ , to this Schrödinger equation. The "boundary" conditions for this problem require that  $\psi(\theta) = \psi(\theta + 2\pi)$ . Apply this boundary condition to the general solution. This results in the quantization of the energy levels of this system. Write down the final expression for the normalized wavefunction and quantized energies. What is the physical significance of these quantum

numbers which can have both positive and negative values? Draw an energy diagram representing the first five energy levels.

b. Treat the six  $\pi$ -electrons of benzene as particles free to move on a ring of radius  $1.40 \text{ \AA}$ , and calculate the energy of the lowest electronic transition. Make sure the Pauli principle is satisfied! What wavelength does this transition correspond to? Suggest some reasons why this differs from the wavelength of the lowest observed transition in benzene, which is  $2600 \text{ \AA}$ .

9. A diatomic molecule constrained to rotate on a flat surface can be modeled as a planar rigid rotor (with eigenfunctions,  $\psi(\phi)$ , analogous to those of the particle on a ring) with fixed bond length  $r$ . At  $t = 0$ , the rotational (orientational) probability distribution is

observed to be described by a wavefunction  $\psi(\phi, 0) = \sqrt{\frac{4}{3}} \cos^2 \phi$ . What values, and with

what probabilities, of the rotational angular momentum,  $L_z = -i\hbar \frac{\partial}{\partial \phi}$ , could be observed in this

system? Explain whether these probabilities would be time dependent as  $\psi(\phi, 0)$  evolves into  $\psi(\phi, t)$ .

10. A particle of mass  $m$  moves in a potential given by

$$V(x, y, z) = \frac{k}{2}(x^2 + y^2 + z^2) = \frac{kr^2}{2}.$$

a. Write down the time-independent Schrödinger equation for this system.

b. Make the substitution  $\psi(x, y, z) = X(x)Y(y)Z(z)$  and separate the variables for this system.

c. What are the solutions to the resulting equations for  $X(x)$ ,  $Y(y)$ , and  $Z(z)$ ?

d. What is the general expression for the quantized energy levels of this system, in terms of the quantum numbers  $n_x$ ,  $n_y$ , and  $n_z$ , which correspond to  $X(x)$ ,  $Y(y)$ , and  $Z(z)$ ?

e. What is the degree of degeneracy of a state of energy

$$E = 5.5\hbar\sqrt{\frac{k}{m}}$$

for this system?

f. An alternative solution may be found by making the substitution  $\psi(r, \theta, \phi) = F(r)G(\theta, \phi)$ . In this substitution, what are the solutions for  $G(\theta, \phi)$ ?

g. Write down the differential equation for  $F(r)$  which is obtained when the substitution  $\psi(r, \theta, \phi) = F(r)G(\theta, \phi)$  is made. Do not solve this equation.

11. Consider an  $N_2$  molecule, in the ground vibrational level of the ground electronic state, which is bombarded by  $100 \text{ eV}$  electrons. This leads to ionization of the  $N_2$  molecule to form  $N_2^+$ . In this problem we will attempt to calculate the vibrational distribution of the newly-formed  $N_2^+$  ions, using a somewhat simplified approach.

a. Calculate (according to classical mechanics) the velocity (in  $\text{cm/sec}$ ) of a  $100 \text{ eV}$  electron, ignoring any relativistic effects. Also calculate the amount of time required for a  $100 \text{ eV}$  electron to pass an  $N_2$  molecule, which you may estimate as having a length of  $2 \text{ \AA}$ .

b. The radial Schrödinger equation for a diatomic molecule treating vibration as a harmonic oscillator can be written as:

$$-\frac{\hbar^2}{2\mu r^2} \frac{d}{dr} \left( r^2 \frac{dF}{dr} \right) + \frac{k}{2}(r - r_e)^2 F = E F$$

Substituting  $F(r) = \frac{F(r)}{r}$ , this equation can be rewritten as:

$$-\frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} F(r) + \frac{k}{2}(r - r_e)^2 F(r) = E F(r)$$

The vibrational Hamiltonian for the ground electronic state of the  $N_2$  molecule within this approximation is given by:

$$H(N_2) = -\frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} + \frac{k_{N_2}}{2}(r - r_{N_2})^2$$

where  $r_{N_2}$  and  $k_{N_2}$  have been measured experimentally to be:

$$r_{N_2} = 1.09769 \text{ \AA}; k_{N_2} = 2.294 \times 10^6 \frac{\text{g}}{\text{sec}^2}$$

The vibrational Hamiltonian for the  $N_2^+$  ion, however, is given by:

$$H(N_2^+) = -\frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} + \frac{k_{N_2^+}}{2}(r - r_{N_2^+})^2$$

where  $r_{N_2^+}$  and  $k_{N_2^+}$  have been measured experimentally to be:

$$r_{N_2^+} = 1.11642 \text{ \AA}; k_{N_2^+} = 2.009 \times 10^6 \frac{\text{g}}{\text{sec}^2}$$

In both systems the reduced mass is  $\mu = 1.1624 \times 10^{-23} \text{ g}$ . Use the above information to write out the ground state vibrational wavefunctions of the  $N_2$  and  $N_2^+$  molecules, giving explicit values for any constants which appear in them. Note: For this problem use the "normal" expression for the ground state wavefunction of a harmonic oscillator. You need not solve the differential equation for this system.

c. During the time scale of the ionization event (which you calculated in part a), the vibrational wavefunction of the  $N_2$  molecule has effectively no time to change. As a result, the newly-formed  $N_2^+$  ion finds itself in a vibrational state which is not an eigenfunction of the new vibrational Hamiltonian,  $H(N_2^+)$ . Assuming that the  $N_2$  molecule was originally in its  $v=0$  vibrational state, calculate the probability that the  $N_2^+$  ion will be produced in its  $v=0$  vibrational state.

12. The force constant,  $k$ , of the C-O bond in carbon monoxide is  $1.87 \times 10^6 \text{ g/sec}^2$ .

Assume that the vibrational motion of CO is purely harmonic and use the reduced mass  $\mu = 6.857 \text{ amu}$ .

a. Calculate the spacing between vibrational energy levels in this molecule, in units of ergs and  $\text{cm}^{-1}$ .

b. Calculate the uncertainty in the internuclear distance in this molecule, assuming it is in its ground vibrational level. Use the ground state vibrational wavefunction ( $v=0$ ), and calculate  $\langle x \rangle$ ,  $\langle x^2 \rangle$ , and  $\Delta x = (\langle x^2 \rangle - \langle x \rangle^2)^{1/2}$ .

c. Under what circumstances (i.e. large or small values of  $k$ ; large or small values of  $\mu$ ) is the uncertainty in internuclear distance large? Can you think of any relationship between this observation and the fact that helium remains a liquid down to absolute zero?

13. Suppose you are given a trial wavefunction of the form:

$$= \frac{Z_e^3}{a_0^3} \exp \frac{-Z_e r_1}{a_0} \exp \frac{-Z_e r_2}{a_0}$$

to represent the electronic structure of a two-electron ion of nuclear charge  $Z$  and suppose that you were also lucky enough to be given the variational integral,  $W$ , (instead of asking you to derive it!):

$$W = Z_e^2 - 2ZZ_e + \frac{5}{8} Z_e \frac{e^2}{a_0} .$$

a. Find the optimum value of the variational parameter  $Z_e$  for an arbitrary nuclear charge  $Z$  by setting  $\frac{dW}{dZ_e} = 0$  . Find both the optimal value of  $Z_e$  and the resulting value of  $W$ .

b. The total energies of some two-electron atoms and ions have been experimentally determined to be:

$Z = 1$	H <sup>-</sup>	-14.35 eV
$Z = 2$	He	-78.98 eV
$Z = 3$	Li <sup>+</sup>	-198.02 eV
$Z = 4$	Be <sup>+2</sup>	-371.5 eV
$Z = 5$	B <sup>+3</sup>	-599.3 eV
$Z = 6$	C <sup>+4</sup>	-881.6 eV
$Z = 7$	N <sup>+5</sup>	-1218.3 eV
$Z = 8$	O <sup>+6</sup>	-1609.5 eV

Using your optimized expression for  $W$ , calculate the estimated total energy of each of these atoms and ions. Also calculate the percent error in your estimate for each ion. What physical reason explains the decrease in percentage error as  $Z$  increases?

c. In 1928, when quantum mechanics was quite young, it was not known whether the isolated, gas-phase hydride ion, H<sup>-</sup>, was stable with respect to dissociation into a hydrogen atom and an electron. Compare your estimated total energy for H<sup>-</sup> to the ground state energy of a hydrogen atom and an isolated electron (system energy = -13.60 eV), and show that this simple variational calculation erroneously predicts H<sup>-</sup> to be unstable. (More complicated variational treatments give a ground state energy of H<sup>-</sup> of -14.35 eV, in agreement with experiment.)

14. A particle of mass  $m$  moves in a one-dimensional potential given by  $\mathbf{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + a|x|$  , where the absolute value function is defined by  $|x| = x$  if  $x \geq 0$  and  $|x| = -x$  if  $x < 0$ .

a. Use the normalized trial wavefunction  $\psi = \frac{1}{\sqrt{2b}} e^{-bx^2}$  to estimate the energy of the ground state of this system, using the variational principle to evaluate  $W(b)$ .

b. Optimize  $b$  to obtain the best approximation to the ground state energy of this system, using a trial function of the form  $\psi = \frac{2}{3} \frac{x}{a} - \frac{1}{3} \frac{x^2}{a^2}$ , as given above. The numerically calculated exact ground state energy is  $0.808616 \frac{\hbar^2}{m^3} \frac{1}{a^3}$ . What is the percent error in your value?

15. The harmonic oscillator is specified by the Hamiltonian:

$$H = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} kx^2.$$

Suppose the ground state solution to this problem were unknown, and that you wish to approximate it using the variational theorem. Choose as your trial wavefunction,

$$\psi = \begin{cases} \sqrt{\frac{15}{16}} \frac{1}{a^2} (a^2 - x^2) & \text{for } -a < x < a \\ 0 & \text{for } |x| > a \end{cases}$$

where  $a$  is an arbitrary parameter which specifies the range of the wavefunction. Note that  $\psi$  is properly normalized as given.

a. Calculate  $\int_{-\infty}^{\infty} \psi^* H \psi dx$  and show it to be given by:

$$\int_{-\infty}^{\infty} \psi^* H \psi dx = \frac{5}{4} \frac{\hbar^2}{ma^2} + \frac{ka^2}{14}.$$

b. Calculate  $\int_{-\infty}^{\infty} \psi^* H \psi dx$  for  $a = b \frac{\hbar^2}{km}^{\frac{1}{4}}$  with  $b = 0.2, 0.4, 0.6, 0.8, 1.0, 1.5, 2.0,$

2.5, 3.0, 4.0, and 5.0, and plot the result.

c. To find the best approximation to the true wavefunction and its energy, find the

minimum of  $\int_{-\infty}^{\infty} \psi^* H \psi dx$  by setting  $\frac{d}{da} \int_{-\infty}^{\infty} \psi^* H \psi dx = 0$  and solving for  $a$ . Substitute this value

into the expression for

$\int_{-\infty}^{\infty} \psi^* H \psi dx$  given in part a. to obtain the best approximation for the energy of the ground

state of the harmonic oscillator.

d. What is the percent error in your calculated energy of part c. ?

16. Einstein told us that the (relativistic) expression for the energy of a particle having rest mass  $m$  and momentum  $p$  is  $E^2 = m^2c^4 + p^2c^2$ .

a. Derive an expression for the relativistic kinetic energy operator which contains terms correct through one higher order than the "ordinary"  $E = mc^2 + \frac{p^2}{2m}$

b. Using the first order correction as a perturbation, compute the first-order perturbation theory estimate of the energy for the 1s level of a hydrogen-like atom (general Z). Show the Z dependence of the result.

$$\text{Note: } \langle r \rangle_{1s} = \frac{Z}{a} \int_0^\infty r^3 \frac{1}{2} e^{-\frac{Zr}{a}} dr \quad \text{and} \quad E_{1s} = -\frac{Z^2 m e^4}{2 \hbar^2}$$

c. For what value of Z does this first-order relativistic correction amount to 10% of the unperturbed (non-relativistic) 1s energy?

17. Consider an electron constrained to move on the surface of a sphere of radius r. The Hamiltonian for such motion consists of a kinetic energy term only  $\mathbf{H}_0 = \frac{\mathbf{L}^2}{2m_e r_0^2}$ , where  $\mathbf{L}$  is the orbital angular momentum operator involving derivatives with respect to the spherical polar coordinates  $(\theta, \phi)$ .  $\mathbf{H}_0$  has the complete set of eigenfunctions  $Y_{lm}^{(0)} = Y_{l,m}(\theta, \phi)$ .

a. Compute the zeroth order energy levels of this system.

b. A uniform electric field is applied along the z-axis, introducing a perturbation  $V = -e z = -e r_0 \cos \theta$ , where  $r_0$  is the strength of the field. Evaluate the correction to the energy of the lowest level through second order in perturbation theory, using the identity

$$\cos \theta Y_{l,m}(\theta, \phi) = \sqrt{\frac{(l+m+1)(l-m+1)}{(2l+1)(2l+3)}} Y_{l+1,m}(\theta, \phi) + \sqrt{\frac{(l+m)(l-m)}{(2l+1)(2l-1)}} Y_{l-1,m}(\theta, \phi).$$

Note that this identity enables you to utilize the orthonormality of the spherical harmonics.

c. The electric polarizability  $\alpha$  gives the response of a molecule to an externally applied electric field, and is defined by  $\alpha = -\frac{2E}{E_0} = 0$  where E is the energy in the presence

of the field and  $E_0$  is the strength of the field. Calculate  $\alpha$  for this system.

d. Use this problem as a model to estimate the polarizability of a hydrogen atom, where  $r_0 = a_0 = 0.529 \text{ \AA}$ , and a cesium atom, which has a single 6s electron with  $r_0 = 2.60 \text{ \AA}$ . The corresponding experimental values are  $\alpha_{\text{H}} = 0.6668 \text{ \AA}^3$  and  $\alpha_{\text{Cs}} = 59.6 \text{ \AA}^3$ .

18. An electron moving in a conjugated bond framework can be viewed as a particle in a box. An externally applied electric field of strength  $\mathcal{E}$  interacts with the electron in a fashion described by the perturbation  $V = e x - \frac{L}{2}$ , where x is the position of the electron in the box, e is the electron's charge, and L is the length of the box.

a. Compute the first order correction to the energy of the n=1 state and the first order wavefunction for the n=1 state. In the wavefunction calculation, you need only compute the contribution to  $\psi_1^{(1)}$  made by  $\psi_2^{(0)}$ . Make a rough (no calculation needed) sketch of  $\psi_1^{(0)} + \psi_1^{(1)}$  as a function of x and physically interpret the graph.

b. Using your answer to part a. compute the induced dipole moment caused by the polarization of the electron density due to the electric field effect  $\mu_{\text{induced}} = -e \int x \psi_1^{(1)*} \psi_1^{(0)} dx$ . You may neglect the term proportional to  $L^2$ ; merely obtain the term linear in  $L$ .

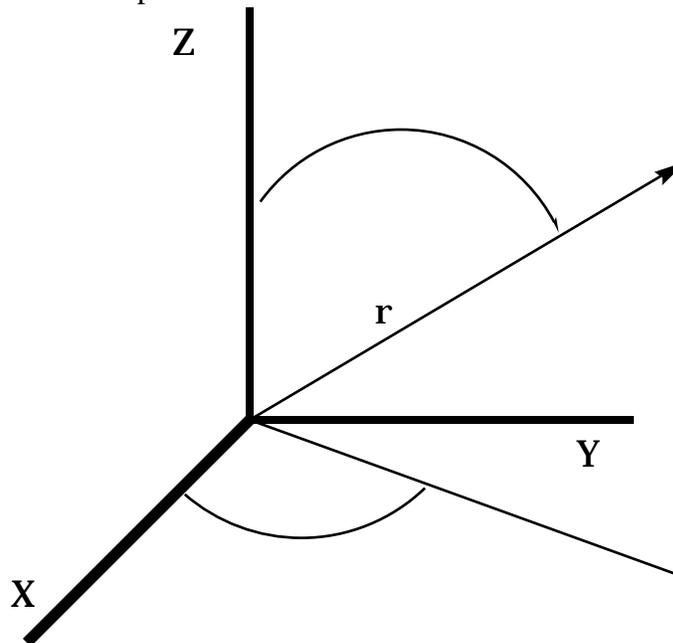
c. Compute the polarizability,  $\alpha$ , of the electron in the  $n=1$  state of the box, and explain physically why  $\alpha$  should depend as it does upon the length of the box  $L$ .

Remember that  $\alpha = \frac{\mu}{E} = 0$ .

## Solutions

### Review Exercises

1. The general relationships are as follows:



$$x = r \sin \theta \cos \phi$$

$$y = r \sin \theta \sin \phi$$

$$z = r \cos \theta$$

$$r^2 = x^2 + y^2 + z^2$$

$$\sin \theta = \frac{\sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2 + z^2}}$$

$$\cos \theta = \frac{z}{\sqrt{x^2 + y^2 + z^2}}$$

$$\tan \phi = \frac{y}{x}$$

a.  $3x + y - 4z = 12$

$$3(r \sin \theta \cos \phi) + r \sin \theta \sin \phi - 4(r \cos \theta) = 12$$

$$r(3 \sin \theta \cos \phi + \sin \theta \sin \phi - 4 \cos \theta) = 12$$

b.  $x = r \cos \theta$   $r^2 = x^2 + y^2$

$$y = r \sin \theta$$

$$\tan \theta = \frac{y}{x}$$

$$z = z$$

$$y^2 + z^2 = 9$$

$$r^2 \sin^2 \theta + z^2 = 9$$

c.  $r = 2 \sin \theta \cos \theta$

$$r = 2 \frac{x}{r}$$

$$r^2 = 2x$$

$$x^2 + y^2 + z^2 = 2x$$

$$x^2 - 2x + y^2 + z^2 = 0$$

$$x^2 - 2x + 1 + y^2 + z^2 = 1$$

$$(x - 1)^2 + y^2 + z^2 = 1$$

2. a.  $9x + 16y \frac{y}{x} = 0$

$$16y dy = -9x dx$$

$$\frac{16}{2} y^2 = -\frac{9}{2} x^2 + c$$

$$16y^2 = -9x^2 + c'$$

$$\frac{y^2}{9} + \frac{x^2}{16} = c'' \text{ (general equation for an ellipse)}$$

b.  $2y + \frac{y}{x} + 6 = 0$

$$2y + 6 = -\frac{dy}{dx}$$

$$y + 3 = -\frac{dy}{2dx}$$

$$-2dx = \frac{dy}{y + 3}$$

$$-2x = \ln(y + 3) + c$$

$$c'e^{-2x} = y + 3$$

$$y = c'e^{-2x} - 3$$

3. a. First determine the eigenvalues:

$$\det \begin{pmatrix} -1 - \lambda & 2 \\ 2 & 2 - \lambda \end{pmatrix} = 0$$

$$(-1 - \lambda)(2 - \lambda) - 2^2 = 0$$

$$-2 + \lambda - 2\lambda + \lambda^2 - 4 = 0$$

$$\lambda^2 - \lambda - 6 = 0$$

$$(\lambda - 3)(\lambda + 2) = 0$$

$$\lambda = 3 \text{ or } \lambda = -2.$$

Next, determine the eigenvectors. First, the eigenvector associated with eigenvalue -2:

$$\begin{pmatrix} -1 & 2 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} C_{11} \\ C_{21} \end{pmatrix} = -2 \begin{pmatrix} C_{11} \\ C_{21} \end{pmatrix}$$

$$\begin{aligned}
 -C_{11} + 2C_{21} &= -2C_{11} \\
 C_{11} &= -2C_{21} \quad (\text{Note: The second row offers no new information, e.g. } 2C_{11} \\
 + 2C_{21} &= -2C_{21})
 \end{aligned}$$

$$C_{11}^2 + C_{21}^2 = 1 \quad (\text{from normalization})$$

$$(-2C_{21})^2 + C_{21}^2 = 1$$

$$4C_{21}^2 + C_{21}^2 = 1$$

$$5C_{21}^2 = 1$$

$$C_{21}^2 = 0.2$$

$$C_{21} = \sqrt{0.2}, \text{ and therefore } C_{11} = -2\sqrt{0.2}.$$

For the eigenvector associated with eigenvalue 3:

$$\begin{array}{ccc|c}
 -1 & 2 & C_{12} & \\
 2 & 2 & C_{22} & \\
 \hline
 & & & = 3 \begin{array}{c} C_{12} \\ C_{22} \end{array}
 \end{array}$$

$$-C_{12} + 2C_{22} = 3C_{12}$$

$$-4C_{12} = -2C_{22}$$

$$C_{12} = 0.5C_{22} \quad (\text{again the second row offers no new information})$$

$$C_{12}^2 + C_{22}^2 = 1 \quad (\text{from normalization})$$

$$(0.5C_{22})^2 + C_{22}^2 = 1$$

$$0.25C_{22}^2 + C_{22}^2 = 1$$

$$1.25C_{22}^2 = 1$$

$$C_{22}^2 = 0.8$$

$$C_{22} = \sqrt{0.8} = 2\sqrt{0.2}, \text{ and therefore } C_{12} = \sqrt{0.2}.$$

Therefore the eigenvector matrix becomes:

$$\begin{array}{cc}
 -2\sqrt{0.2} & \sqrt{0.2} \\
 \sqrt{0.2} & 2\sqrt{0.2}
 \end{array}$$

b. First determine the eigenvalues:

$$\det \begin{bmatrix} -2 & - & 0 & 0 \\ 0 & -1 & - & 2 \\ 0 & 2 & 2 & - \end{bmatrix} = 0$$

$$\det \begin{bmatrix} -2 & - & 2 \\ - & - & 2 \end{bmatrix} \det \begin{bmatrix} -1 & - & 2 \\ 2 & 2 & - \end{bmatrix} = 0$$

From 3a, the solutions then become -2, -2, and 3. Next, determine the eigenvectors. First the eigenvector associated with eigenvalue 3 (the third root):

$$\begin{array}{ccc|c}
 -2 & 0 & 0 & C_{11} \\
 0 & -1 & 2 & C_{21} \\
 0 & 2 & 2 & C_{31} \\
 \hline
 & & & = 3 \begin{array}{c} C_{11} \\ C_{21} \\ C_{31} \end{array}
 \end{array}$$

$$-2C_{13} = 3C_{13} \quad (\text{row one})$$

$$C_{13} = 0$$

$$-C_{23} + 2C_{33} = 3C_{23} \quad (\text{row two})$$

$$2C_{33} = 4C_{23}$$

$$C_{33} = 2C_{23} \quad (\text{again the third row offers no new information})$$

$$C_{13}^2 + C_{23}^2 + C_{33}^2 = 1 \quad (\text{from normalization})$$

$$0 + C_{23}^2 + (2C_{23})^2 = 1$$

$$5C_{23}^2 = 1$$

$$C_{23} = \sqrt{0.2}, \text{ and therefore } C_{33} = 2\sqrt{0.2}.$$

Next, find the pair of eigenvectors associated with the degenerate eigenvalue of -2. First, root one eigenvector one:

$$-2C_{11} = -2C_{11} \text{ (no new information from row one)}$$

$$-C_{21} + 2C_{31} = -2C_{21} \text{ (row two)}$$

$$C_{21} = -2C_{31} \text{ (again the third row offers no new information)}$$

$$C_{11}^2 + C_{21}^2 + C_{31}^2 = 1 \text{ (from normalization)}$$

$$C_{11}^2 + (-2C_{31})^2 + C_{31}^2 = 1$$

$$C_{11}^2 + 5C_{31}^2 = 1$$

$$C_{11} =$$

$$\sqrt{1 - 5C_{31}^2} \text{ (Note: There are now two equations with three unknowns.)}$$

Second, root two eigenvector two:

$$-2C_{12} = -2C_{12} \text{ (no new information from row one)}$$

$$-C_{22} + 2C_{32} = -2C_{22} \text{ (row two)}$$

$$C_{22} = -2C_{32} \text{ (again the third row offers no new information)}$$

$$C_{12}^2 + C_{22}^2 + C_{32}^2 = 1 \text{ (from normalization)}$$

$$C_{12}^2 + (-2C_{32})^2 + C_{32}^2 = 1$$

$$C_{12}^2 + 5C_{32}^2 = 1$$

$$C_{12} =$$

$$\sqrt{1 - 5C_{32}^2} \text{ (Note: Again there are now two equations with three unknowns)}$$

$$C_{11}C_{12} + C_{21}C_{22} + C_{31}C_{32} = 0 \text{ (from orthogonalization)}$$

Now there are five equations with six unknowns.

Arbitrarily choose  $C_{11} = 0$

$$C_{11} = 0 = \sqrt{1 - 5C_{31}^2}$$

$$5C_{31}^2 = 1$$

$$C_{31} = \sqrt{0.2}$$

$$C_{21} = -2\sqrt{0.2}$$

$$C_{11}C_{12} + C_{21}C_{22} + C_{31}C_{32} = 0 \text{ (from orthogonalization)}$$

$$0 + -2\sqrt{0.2}(-2C_{32}) + \sqrt{0.2} C_{32} = 0$$

$$5C_{32} = 0$$

$$C_{32} = 0, C_{22} = 0, \text{ and } C_{12} = 1$$

Therefore the eigenvector matrix becomes:

$$\begin{matrix} 0 & 1 & 0 \\ -2\sqrt{0.2} & 0 & \sqrt{0.2} \\ \sqrt{0.2} & 0 & 2\sqrt{0.2} \end{matrix}$$

4. Show:  $\langle 1 | 1 \rangle = 1$ ,  $\langle 2 | 2 \rangle = 1$ , and  $\langle 1 | 2 \rangle = 0$

$$\langle 1 | 1 \rangle = 1$$

$$(-2\sqrt{0.2})^2 + (\sqrt{0.2})^2 = 1$$

$$4(0.2) + 0.2 = 1$$

$$\begin{aligned}
0.8 + 0.2 & \stackrel{?}{=} 1 \\
1 & = 1 \\
\langle 2 | 2 \rangle & \stackrel{?}{=} 1 \\
(\sqrt{0.2})^2 + (2\sqrt{0.2})^2 & \stackrel{?}{=} 1 \\
0.2 + 4(0.2) & \stackrel{?}{=} 1 \\
0.2 + 0.8 & \stackrel{?}{=} 1 \\
1 & = 1 \\
\langle 1 | 2 \rangle = \langle 2 | 1 \rangle & \stackrel{?}{=} 0 \\
-2\sqrt{0.2}\sqrt{0.2} + \sqrt{0.2}2\sqrt{0.2} & \stackrel{?}{=} 0 \\
-2(0.2) + 2(0.2) & \stackrel{?}{=} 0 \\
-0.4 + 0.4 & \stackrel{?}{=} 0 \\
0 & = 0
\end{aligned}$$

5. Show (for the degenerate eigenvalue;  $\lambda = -2$ ):  $\langle 1 | 1 \rangle = 1$ ,  $\langle 2 | 2 \rangle = 1$ , and  $\langle 1 | 2 \rangle = 0$

$$\begin{aligned}
\langle 1 | 1 \rangle & \stackrel{?}{=} 1 \\
0 + (-2\sqrt{0.2})^2 + (\sqrt{0.2})^2 & \stackrel{?}{=} 1 \\
4(0.2) + 0.2 & \stackrel{?}{=} 1 \\
0.8 + 0.2 & \stackrel{?}{=} 1 \\
1 & = 1 \\
\langle 2 | 2 \rangle & \stackrel{?}{=} 1 \\
1^2 + 0 + 0 & \stackrel{?}{=} 1 \\
1 & = 1 \\
\langle 1 | 2 \rangle = \langle 2 | 1 \rangle & \stackrel{?}{=} 0 \\
(0)(1) + (-2\sqrt{0.2})(0) + (\sqrt{0.2})(0) & \stackrel{?}{=} 0 \\
0 & = 0
\end{aligned}$$

6. Suppose the solution is of the form  $x(t) = e^{-\lambda t}$ , with  $\lambda$  unknown. Inserting this trial solution into the differential equation results in the following:

$$\begin{aligned}
\frac{d^2}{dt^2} e^{-\lambda t} + k^2 e^{-\lambda t} & = 0 \\
\lambda^2 e^{-\lambda t} + k^2 e^{-\lambda t} & = 0 \\
(\lambda^2 + k^2) x(t) & = 0 \\
(\lambda^2 + k^2) & = 0 \\
\lambda^2 & = -k^2
\end{aligned}$$

$$= \sqrt{-k^2}$$

$$= \pm ik$$

Solutions are of the form  $e^{ikt}$ ,  $e^{-ikt}$ , or a combination of both:  $x(t) = C_1 e^{ikt} + C_2 e^{-ikt}$ .

Euler's formula also states that:  $e^{\pm i\theta} = \cos \theta \pm i \sin \theta$ , so the previous equation for  $x(t)$  can also be written as:

$$x(t) = C_1 \{ \cos(kt) + i \sin(kt) \} + C_2 \{ \cos(kt) - i \sin(kt) \}$$

$$x(t) = (C_1 + C_2) \cos(kt) + (C_1 - C_2) i \sin(kt), \text{ or alternatively}$$

$$x(t) = C_3 \cos(kt) + C_4 \sin(kt).$$

We can determine these coefficients by making use of the "boundary conditions".

$$\text{at } t = 0, x(0) = L$$

$$x(0) = C_3 \cos(0) + C_4 \sin(0) = L$$

$$C_3 = L$$

$$\text{at } t = 0, \frac{dx(0)}{dt} = 0$$

$$\frac{d}{dt} x(t) = \frac{d}{dt} (C_3 \cos(kt) + C_4 \sin(kt))$$

$$\frac{d}{dt} x(t) = -C_3 k \sin(kt) + C_4 k \cos(kt)$$

$$\frac{d}{dt} x(0) = 0 = -C_3 k \sin(0) + C_4 k \cos(0)$$

$$C_4 k = 0$$

$$C_4 = 0$$

The solution is of the form:  $x(t) = L \cos(kt)$

### Exercises

1. a. 
$$\text{K.E.} = \frac{mv^2}{2} = \frac{m}{m} \frac{mv^2}{2} = \frac{(mv)^2}{2m} = \frac{p^2}{2m}$$

$$\text{K.E.} = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2)$$

$$\text{K.E.} = \frac{1}{2m} \left( \frac{\hbar}{i} \frac{\partial}{\partial x} \right)^2 + \frac{\hbar}{i} \frac{\partial}{\partial y} \right)^2 + \left( \frac{\hbar}{i} \frac{\partial}{\partial z} \right)^2$$

$$\text{K.E.} = \frac{-\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right)$$

b. 
$$\mathbf{p} = m\mathbf{v} = \mathbf{i}p_x + \mathbf{j}p_y + \mathbf{k}p_z$$

$$\mathbf{p} = \mathbf{i} \frac{\hbar}{i} \frac{\partial}{\partial x} + \mathbf{j} \frac{\hbar}{i} \frac{\partial}{\partial y} + \mathbf{k} \frac{\hbar}{i} \frac{\partial}{\partial z}$$

where  $\mathbf{i}$ ,  $\mathbf{j}$ , and  $\mathbf{k}$  are unit vectors along the x, y, and z axes.

c. 
$$L_y = zp_x - xp_z$$

$$L_y = z \frac{\hbar}{i} \frac{\partial}{\partial x} - x \frac{\hbar}{i} \frac{\partial}{\partial z}$$

2. First derive the general formulas for  $\frac{r}{x}$ ,  $\frac{r}{y}$ ,  $\frac{r}{z}$  in terms of  $r$ ,  $\theta$ , and  $\phi$ , and  $\frac{r}{r}$ ,  $\frac{\theta}{r}$ ,  $\frac{\phi}{r}$ ,

and  $\frac{r}{r}$  in terms of  $x, y$ , and  $z$ . The general relationships are as follows:

$$\begin{aligned} x &= r \sin \theta \cos \phi & r^2 &= x^2 + y^2 + z^2 \\ y &= r \sin \theta \sin \phi & \sin \theta &= \frac{\sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2 + z^2}} \\ z &= r \cos \theta & \cos \theta &= \frac{z}{\sqrt{x^2 + y^2 + z^2}} \\ & & \tan \theta &= \frac{y}{x} \end{aligned}$$

First  $\frac{r}{x}$ ,  $\frac{r}{y}$ , and  $\frac{r}{z}$  from the chain rule:

$$\begin{aligned} \frac{r}{x} &= \frac{r}{x} \frac{r}{y,z} \frac{r}{r} + \frac{r}{x} \frac{r}{y,z} \frac{r}{r} + \frac{r}{x} \frac{r}{y,z} \frac{r}{r}, \\ \frac{r}{y} &= \frac{r}{y} \frac{r}{x,z} \frac{r}{r} + \frac{r}{y} \frac{r}{x,z} \frac{r}{r} + \frac{r}{y} \frac{r}{x,z} \frac{r}{r}, \\ \frac{r}{z} &= \frac{r}{z} \frac{r}{x,y} \frac{r}{r} + \frac{r}{z} \frac{r}{x,y} \frac{r}{r} + \frac{r}{z} \frac{r}{x,y} \frac{r}{r}. \end{aligned}$$

Evaluation of the many "coefficients" gives the following:

$$\begin{aligned} \frac{r}{x} \frac{r}{y,z} &= \sin \theta \cos \phi, \quad \frac{r}{x} \frac{r}{y,z} = \frac{\cos \theta \cos \phi}{r}, \quad \frac{r}{x} \frac{r}{y,z} = -\frac{\sin \theta}{r \sin \theta}, \\ \frac{r}{y} \frac{r}{x,z} &= \sin \theta \sin \phi, \quad \frac{r}{y} \frac{r}{x,z} = \frac{\cos \theta \sin \phi}{r}, \quad \frac{r}{y} \frac{r}{x,z} = \frac{\cos \theta}{r \sin \theta}, \\ \frac{r}{z} \frac{r}{x,y} &= \cos \theta, \quad \frac{r}{z} \frac{r}{x,y} = -\frac{\sin \theta}{r}, \quad \text{and} \quad \frac{r}{z} \frac{r}{x,y} = 0. \end{aligned}$$

Upon substitution of these "coefficients":

$$\begin{aligned} \frac{r}{x} &= \sin \theta \cos \phi \frac{r}{r} + \frac{\cos \theta \cos \phi}{r} \frac{r}{r} - \frac{\sin \theta}{r \sin \theta} \frac{r}{r}, \\ \frac{r}{y} &= \sin \theta \sin \phi \frac{r}{r} + \frac{\cos \theta \sin \phi}{r} \frac{r}{r} + \frac{\cos \theta}{r \sin \theta} \frac{r}{r}, \quad \text{and} \\ \frac{r}{z} &= \cos \theta \frac{r}{r} - \frac{\sin \theta}{r} \frac{r}{r} + 0 \frac{r}{r}. \end{aligned}$$

Next  $\frac{r}{r}$ ,  $\frac{\theta}{r}$ , and  $\frac{\phi}{r}$  from the chain rule:

$$\frac{r}{r} = \frac{x}{r}, \quad \frac{\theta}{r} = \frac{y}{r}, \quad \frac{\phi}{r} = \frac{z}{r}, \quad \frac{r}{z}$$

$$\frac{x}{r} = \frac{x}{r}, \quad \frac{y}{r} = \frac{y}{r}, \quad \frac{z}{r} = \frac{z}{r}, \quad \text{and}$$

$$\frac{z}{r} = \frac{z}{r}, \quad \frac{x}{r} = \frac{x}{r}, \quad \frac{y}{r} = \frac{y}{r}, \quad \frac{z}{r} = \frac{z}{r}.$$

Again evaluation of the the many "coefficients" results in:

$$\frac{x}{r} = \frac{x}{\sqrt{x^2 + y^2 + z^2}}, \quad \frac{y}{r} = \frac{y}{\sqrt{x^2 + y^2 + z^2}},$$

$$\frac{z}{r} = \frac{z}{\sqrt{x^2 + y^2 + z^2}}, \quad \frac{x}{r} = \frac{xz}{\sqrt{x^2 + y^2}}, \quad \frac{y}{r} = \frac{yz}{\sqrt{x^2 + y^2}},$$

$$\frac{z}{r} = -\sqrt{x^2 + y^2}, \quad \frac{x}{r} = -y, \quad \frac{y}{r} = x, \quad \text{and} \quad \frac{z}{r} = 0$$

Upon substitution of these "coefficients":

$$\frac{x}{r} = \frac{x}{\sqrt{x^2 + y^2 + z^2}} \frac{1}{x} + \frac{y}{\sqrt{x^2 + y^2 + z^2}} \frac{1}{y}$$

$$+ \frac{z}{\sqrt{x^2 + y^2 + z^2}} \frac{1}{z}$$

$$\frac{z}{r} = \frac{xz}{\sqrt{x^2 + y^2}} \frac{1}{x} + \frac{yz}{\sqrt{x^2 + y^2}} \frac{1}{y} - \sqrt{x^2 + y^2} \frac{1}{z}$$

$$\frac{z}{r} = -y \frac{1}{x} + x \frac{1}{y} + 0 \frac{1}{z}.$$

Note, these many "coefficients" are the elements which make up the Jacobian matrix used whenever one wishes to transform a function from one coordinate representation to another. One very familiar result should be in transforming the volume element  $dx dy dz$  to  $r^2 \sin \theta dr d\theta d\phi$ . For example:

$$f(x,y,z) dx dy dz =$$

$$f(x(r, \theta, \phi), y(r, \theta, \phi), z(r, \theta, \phi)) \begin{pmatrix} \frac{x}{r} & \frac{x}{r} & \frac{x}{r} \\ \frac{y}{r} & \frac{y}{r} & \frac{y}{r} \\ \frac{z}{r} & \frac{z}{r} & \frac{z}{r} \end{pmatrix} dr d\theta d\phi$$

$$a. \quad \mathbf{L}_x = \frac{\hbar}{i} \left( y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right)$$

$$\mathbf{L}_x = \frac{\hbar}{i} \left( r \sin \theta \sin \phi \cos \theta - \frac{\sin \theta}{r} \right)$$

$$-\frac{\hbar}{i} r \cos \theta \sin \theta \frac{\partial}{\partial r} + \frac{\cos \theta \sin \theta}{r} \frac{\partial}{\partial \theta} + \frac{\cos \theta}{r \sin \theta} \frac{\partial}{\partial \phi}$$

$$L_x = -\frac{\hbar}{i} \sin \theta \frac{\partial}{\partial \phi} + \cot \theta \cos \theta \frac{\partial}{\partial \theta}$$

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

$$L_z = \frac{\hbar}{i} \left( -y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} \right)$$

	B	B'	B''
i.	$4x^4 - 12x^2 + 3$	$16x^3 - 24x$	$48x^2 - 24$
ii.	$5x^4$	$20x^3$	$60x^2$
iii.	$e^{3x} + e^{-3x}$	$3(e^{3x} - e^{-3x})$	$9(e^{3x} + e^{-3x})$
iv.	$x^2 - 4x + 2$	$2x - 4$	$2$
v.	$4x^3 - 3x$	$12x^2 - 3$	$24x$

B(v.) is an eigenfunction of A(i.):

$$(1-x^2) \frac{d^2}{dx^2} B(v.) - x \frac{d}{dx} B(v.) =$$

$$(1-x^2)(24x) - x(12x^2 - 3) \\ 24x - 24x^3 - 12x^3 + 3x \\ -36x^3 + 27x$$

$$-9(4x^3 - 3x) \text{ (eigenvalue is -9)}$$

B(iii.) is an eigenfunction of A(ii.):

$$\frac{d^2}{dx^2} B(iii.) =$$

$$9(e^{3x} + e^{-3x}) \text{ (eigenvalue is 9)}$$

B(ii.) is an eigenfunction of A(iii.):

$$x \frac{d}{dx} B(ii.) =$$

$$x(20x^3) \\ 20x^4$$

$$4(5x^4) \text{ (eigenvalue is 4)}$$

B(i.) is an eigenfunction of A(vi.):

$$\frac{d^2}{dx^2} B(i.) - 2x \frac{d}{dx} B(i.) =$$

$$(48x^2 - 24) - 2x(16x^3 - 24x) \\ 48x^2 - 24 - 32x^4 + 48x^2 \\ -32x^4 + 96x^2 - 24$$

$$-8(4x^4 - 12x^2 + 3) \text{ (eigenvalue is -8)}$$

B(iv.) is an eigenfunction of A(v.):

$$\begin{aligned}
& x \frac{d^2}{dx^2} + (1-x) \frac{d}{dx} \quad B(\text{iv.}) = \\
& x(2) + (1-x)(2x-4) \\
& 2x + 2x - 4 - 2x^2 + 4x \\
& -2x^2 + 8x - 4 \\
& -2(x^2 - 4x + 2) \quad (\text{eigenvalue is } -2)
\end{aligned}$$

4. Show that:  $f^* A g = g(Af)^*$

a. Suppose  $f$  and  $g$  are functions of  $x$  and evaluate the integral on the left hand side by "integration by parts":

$$f(x)^* \left(-i\hbar \frac{d}{dx}\right) g(x) dx$$

$$\text{let } dv = \frac{d}{dx} g(x) dx \quad \text{and} \quad u = -i\hbar f(x)^*$$

$$v = g(x) \quad du = -i\hbar \frac{d}{dx} f(x)^* dx$$

$$\text{Now, } udv = uv - vdu,$$

so:

$$f(x)^* \left(-i\hbar \frac{d}{dx}\right) g(x) dx = -i\hbar f(x)^* g(x) + i\hbar g(x) \frac{d}{dx} f(x)^* dx.$$

Note that in, principle, it is impossible to prove hermiticity unless you are given knowledge of the type of function on which the operator is acting. Hermiticity requires (as can be seen in this example) that the term  $-i\hbar f(x)^* g(x)$  vanish when evaluated at the integral limits. This, in general, will occur for the "well behaved" functions (e.g., in bound state quantum chemistry, the wavefunctions will vanish as the distances among particles approaches infinity). So, in proving the hermiticity of an operator, one must be careful to specify the behavior of the functions on which the operator is considered to act. This means that an operator may be hermitian for one class of functions and non-hermitian for another class of functions. If we assume that  $f$  and  $g$  vanish at the boundaries, then we have

$$f(x)^* \left(-i\hbar \frac{d}{dx}\right) g(x) dx = g(x) \left(-i\hbar \frac{d}{dx}\right) f(x)^* dx$$

b. Suppose  $f$  and  $g$  are functions of  $y$  and  $z$  and evaluate the integral on the left hand side by "integration by parts" as in the previous exercise:

$$f(y,z)^* \left(-i\hbar y \frac{d}{dz} - z \frac{d}{dy}\right) g(y,z) dy dz$$

$$= f(y,z)^* \left(-i\hbar y \frac{d}{dz}\right) g(y,z) dy dz - f(y,z)^* \left(-i\hbar z \frac{d}{dy}\right) g(y,z) dy dz$$

For the first integral,  $f(z)^* \left(-i\hbar y \frac{d}{dz}\right) g(z) dz,$

$$\text{let } dv = \frac{1}{z} g(z) dz \quad u = -i\hbar y f(z)^*$$

$$v = g(z) \quad du = -i\hbar y \frac{1}{z} f(z)^* dz$$

so:

$$\begin{aligned} f(z)^* (-i\hbar y \frac{1}{z}) g(z) dz &= -i\hbar y f(z)^* g(z) + i\hbar y \frac{1}{z} g(z) f(z)^* dz \\ &= g(z) -i\hbar y \frac{1}{z} f(z)^* dz . \end{aligned}$$

For the second integral,  $f(y)^* -i\hbar z \frac{1}{y} g(y) dy$ ,

$$\text{let } dv = \frac{1}{y} g(y) dy \quad u = -i\hbar z f(y)^*$$

$$v = g(y) \quad du = -i\hbar z \frac{1}{y} f(y)^* dy$$

so:

$$\begin{aligned} f(y)^* (-i\hbar z \frac{1}{y}) g(y) dy &= -i\hbar z f(y)^* g(y) + i\hbar z \frac{1}{y} g(y) f(y)^* dy \\ &= g(y) -i\hbar z \frac{1}{y} f(y)^* dy \end{aligned}$$

$$\begin{aligned} f(y,z)^* -i\hbar y \frac{1}{z} - z \frac{1}{y} g(y,z) dy dz \\ &= g(z) -i\hbar y \frac{1}{z} f(z)^* dz - g(y) -i\hbar z \frac{1}{y} f(y)^* dy \\ &= g(y,z) -i\hbar y \frac{1}{z} - z \frac{1}{y} f(y,z)^* dy dz . \end{aligned}$$

Again we have had to assume that the functions  $f$  and  $g$  vanish at the boundary.

$$\begin{aligned} 5. \quad \mathbf{L}_+ &= \mathbf{L}_x + i\mathbf{L}_y \\ \mathbf{L}_- &= \mathbf{L}_x - i\mathbf{L}_y, \text{ so} \\ \mathbf{L}_+ + \mathbf{L}_- &= 2\mathbf{L}_x, \text{ or } \mathbf{L}_x = \frac{1}{2}(\mathbf{L}_+ + \mathbf{L}_-) \\ \mathbf{L}_+ Y_{1,m} &= \sqrt{1(1+1) - m(m+1)} \hbar Y_{1,m+1} \\ \mathbf{L}_- Y_{1,m} &= \sqrt{1(1+1) - m(m-1)} \hbar Y_{1,m-1} \end{aligned}$$

Using these relationships:

$$\mathbf{L}_- Y_{1,-1} = 0, \mathbf{L}_- Y_{1,0} = \sqrt{2}\hbar Y_{1,-1}, \mathbf{L}_- Y_{1,1} = \sqrt{2}\hbar Y_{1,0}$$

$L_+ |2p-1\rangle = \sqrt{2\hbar} |2p_0\rangle$ ,  $L_+ |2p_0\rangle = \sqrt{2\hbar} |2p+1\rangle$ ,  $L_+ |2p+1\rangle = 0$ , and the following  $L_x$  matrix elements can be evaluated:

$$L_x(1,1) = \langle 2p-1 | \frac{1}{2}(L_+ + L_-) | 2p-1 \rangle = 0$$

$$L_x(1,2) = \langle 2p-1 | \frac{1}{2}(L_+ + L_-) | 2p_0 \rangle = \frac{\sqrt{2}}{2} \hbar$$

$$L_x(1,3) = \langle 2p-1 | \frac{1}{2}(L_+ + L_-) | 2p+1 \rangle = 0$$

$$L_x(2,1) = \langle 2p_0 | \frac{1}{2}(L_+ + L_-) | 2p-1 \rangle = \frac{\sqrt{2}}{2} \hbar$$

$$L_x(2,2) = \langle 2p_0 | \frac{1}{2}(L_+ + L_-) | 2p_0 \rangle = 0$$

$$L_x(2,3) = \langle 2p_0 | \frac{1}{2}(L_+ + L_-) | 2p+1 \rangle = \frac{\sqrt{2}}{2} \hbar$$

$$L_x(3,1) = \langle 2p+1 | \frac{1}{2}(L_+ + L_-) | 2p-1 \rangle = 0$$

$$L_x(3,2) = \langle 2p+1 | \frac{1}{2}(L_+ + L_-) | 2p_0 \rangle = \frac{\sqrt{2}}{2} \hbar$$

$$L_x(3,3) = 0$$

$$0 \quad \frac{\sqrt{2}}{2} \hbar \quad 0$$

This matrix:  $\begin{pmatrix} \frac{\sqrt{2}}{2} \hbar & 0 & \frac{\sqrt{2}}{2} \hbar \\ 0 & \frac{\sqrt{2}}{2} \hbar & 0 \\ 0 & \frac{\sqrt{2}}{2} \hbar & 0 \end{pmatrix}$ , can now be diagonalized:

$$0 \quad \frac{\sqrt{2}}{2} \hbar \quad 0$$

$$0 - \quad \frac{\sqrt{2}}{2} \hbar \quad 0$$

$$\frac{\sqrt{2}}{2} \hbar \quad 0 - \quad \frac{\sqrt{2}}{2} \hbar = 0$$

$$0 \quad \frac{\sqrt{2}}{2} \hbar \quad 0 -$$

$$0 - \quad \frac{\sqrt{2}}{2} \hbar \quad \frac{\sqrt{2}}{2} \hbar \quad \frac{\sqrt{2}}{2} \hbar = 0$$

$$\frac{\sqrt{2}}{2} \hbar \quad 0 - \quad 0 \quad 0 -$$

Expanding these determinants yields:

$$\left( \frac{\sqrt{2}}{2} \hbar - \frac{\hbar^2}{2} \right) \left( \frac{\sqrt{2}}{2} \hbar - \frac{\sqrt{2}}{2} \hbar \right) \frac{\sqrt{2}}{2} \hbar = 0$$

$$- (\hbar^2 - \hbar^2) = 0$$

$$- (\hbar - \hbar)(\hbar + \hbar) = 0$$

with roots:  $0, \hbar$ , and  $-\hbar$

Next, determine the corresponding eigenvectors:

For  $\omega = 0$ :

$$0 \quad \frac{\sqrt{2}}{2}\hbar \quad 0 \quad C_{11} \quad C_{11}$$

$$\frac{\sqrt{2}}{2}\hbar \quad 0 \quad \frac{\sqrt{2}}{2}\hbar \quad C_{21} = 0 \quad C_{21}$$

$$0 \quad \frac{\sqrt{2}}{2}\hbar \quad 0 \quad C_{31} \quad C_{31}$$

$$\frac{\sqrt{2}}{2}\hbar C_{21} = 0 \text{ (row one)}$$

$$C_{21} = 0$$

$$\frac{\sqrt{2}}{2}\hbar C_{11} + \frac{\sqrt{2}}{2}\hbar C_{31} = 0 \text{ (row two)}$$

$$C_{11} + C_{31} = 0$$

$$C_{11} = -C_{31}$$

$$C_{11}^2 + C_{21}^2 + C_{31}^2 = 1 \text{ (normalization)}$$

$$C_{11}^2 + (-C_{11})^2 = 1$$

$$2C_{11}^2 = 1$$

$$C_{11} = \frac{1}{\sqrt{2}}, C_{21} = 0, \text{ and } C_{31} = -\frac{1}{\sqrt{2}}$$

For  $\omega = 1\hbar$ :

$$0 \quad \frac{\sqrt{2}}{2}\hbar \quad 0 \quad C_{12} \quad C_{12}$$

$$\frac{\sqrt{2}}{2}\hbar \quad 0 \quad \frac{\sqrt{2}}{2}\hbar \quad C_{22} = 1\hbar \quad C_{22}$$

$$0 \quad \frac{\sqrt{2}}{2}\hbar \quad 0 \quad C_{32} \quad C_{32}$$

$$\frac{\sqrt{2}}{2}\hbar C_{22} = \hbar C_{12} \text{ (row one)}$$

$$C_{12} = \frac{\sqrt{2}}{2} C_{22}$$

$$\frac{\sqrt{2}}{2}\hbar C_{12} + \frac{\sqrt{2}}{2}\hbar C_{32} = \hbar C_{22} \text{ (row two)}$$

$$\frac{\sqrt{2}}{2} \frac{\sqrt{2}}{2} C_{22} + \frac{\sqrt{2}}{2} C_{32} = C_{22}$$

$$\frac{1}{2} C_{22} + \frac{\sqrt{2}}{2} C_{32} = C_{22}$$

$$\frac{\sqrt{2}}{2} C_{32} = \frac{1}{2} C_{22}$$

$$C_{32} = \frac{\sqrt{2}}{2} C_{22}$$

$$C_{12}^2 + C_{22}^2 + C_{32}^2 = 1 \text{ (normalization)}$$

$$\frac{\sqrt{2}}{2} C_{22}^2 + C_{22}^2 + \frac{\sqrt{2}}{2} C_{22}^2 = 1$$

$$\frac{1}{2} C_{22}^2 + C_{22}^2 + \frac{1}{2} C_{22}^2 = 1$$

$$2C_{22}^2 = 1$$

$$C_{22} = \frac{\sqrt{2}}{2}$$

$$C_{12} = \frac{1}{2}, C_{22} = \frac{\sqrt{2}}{2}, \text{ and } C_{32} = \frac{1}{2}$$

For  $\omega = -1\hbar$  :

$$0 \quad \frac{\sqrt{2}}{2}\hbar \quad 0 \quad C_{13} \quad C_{13}$$

$$\frac{\sqrt{2}}{2}\hbar \quad 0 \quad \frac{\sqrt{2}}{2}\hbar \quad C_{23} = -1\hbar \quad C_{23}$$

$$0 \quad \frac{\sqrt{2}}{2}\hbar \quad 0 \quad C_{33} \quad C_{33}$$

$$\frac{\sqrt{2}}{2} \hbar C_{23} = -\hbar C_{13} \text{ (row one)}$$

$$C_{13} = -\frac{\sqrt{2}}{2} C_{23}$$

$$\frac{\sqrt{2}}{2} \hbar C_{13} + \frac{\sqrt{2}}{2} \hbar C_{33} = -\hbar C_{23} \text{ (row two)}$$

$$\frac{\sqrt{2}}{2} \left(-\frac{\sqrt{2}}{2} C_{23}\right) + \frac{\sqrt{2}}{2} C_{33} = -C_{23}$$

$$-\frac{1}{2} C_{23} + \frac{\sqrt{2}}{2} C_{33} = -C_{23}$$

$$\frac{\sqrt{2}}{2} C_{33} = -\frac{1}{2} C_{23}$$

$$C_{33} = -\frac{\sqrt{2}}{2} C_{23}$$

$$C_{13}^2 + C_{23}^2 + C_{33}^2 = 1 \text{ (normalization)}$$

$$-\frac{\sqrt{2}}{2} C_{23}^2 + C_{23}^2 + -\frac{\sqrt{2}}{2} C_{23}^2 = 1$$

$$\frac{1}{2} C_{23}^2 + C_{23}^2 + \frac{1}{2} C_{23}^2 = 1$$

$$2C_{23}^2 = 1$$

$$C_{23} = \frac{\sqrt{2}}{2}$$

$$C_{13} = -\frac{1}{2}, C_{23} = \frac{\sqrt{2}}{2}, \text{ and } C_{33} = -\frac{1}{2}$$

Show:  $\langle 1|1\rangle = 1$ ,  $\langle 2|2\rangle = 1$ ,  $\langle 3|3\rangle = 1$ ,  $\langle 1|2\rangle = 0$ ,  $\langle 1|3\rangle = 0$ , and  $\langle 2|3\rangle = 0$ .

$$\langle 1|1\rangle = 1$$

$$\frac{\sqrt{2}}{2}^2 + 0 + \frac{-\sqrt{2}}{2}^2 = 1$$

$$\frac{1}{2} + \frac{1}{2} = 1$$

$$1 = 1$$

$$\langle 2|2\rangle = 1$$

$$\frac{1}{2}^2 + \frac{\sqrt{2}}{2}^2 + \frac{1}{2}^2 = 1$$

$$\frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$$

$$1 = 1$$

$$\langle 3|3\rangle = 1$$

$$-\frac{1}{2}^2 + \frac{\sqrt{2}}{2}^2 + \frac{-1}{2}^2 = 1$$

$$\frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1$$

$$1 = 1$$

$$\langle 1|2\rangle = \langle 2|1\rangle = 0$$

$$\frac{\sqrt{2}}{2} \frac{1}{2} + (0) \frac{\sqrt{2}}{2} + \frac{-\sqrt{2}}{2} \frac{1}{2} = 0$$

$$\frac{\sqrt{2}}{4} - \frac{\sqrt{2}}{4} = 0$$

$$0 = 0$$

$$\langle 1|3\rangle = \langle 3|1\rangle = 0$$

$$\frac{\sqrt{2}}{2} \frac{-1}{2} + (0) \frac{\sqrt{2}}{2} + \frac{-\sqrt{2}}{2} \frac{-1}{2} = 0$$

$$-\frac{\sqrt{2}}{4} + \frac{\sqrt{2}}{4} = 0$$

$$0 = 0$$

$$\langle 2|3\rangle = \langle 3|2\rangle = 0$$

$$\frac{1}{2} \frac{-1}{2} + \frac{\sqrt{2}}{2} \frac{\sqrt{2}}{2} + \frac{1}{2} \frac{-1}{2} = 0$$

$$\frac{1}{4} + \frac{1}{2} + \frac{1}{4} \stackrel{?}{=} 0$$

$$0 = 0$$

$$6. \quad P_{2p+1} = \left\langle 2p+1 \quad 0\hbar \right\rangle_{L_x}^2$$

$$\frac{0\hbar}{L_x} = \frac{1}{\sqrt{2}} \quad 2p-1 \quad \frac{1}{\sqrt{2}} \quad 2p+1$$

$$P_{2p+1} = \frac{1}{\sqrt{2}} \left\langle 2p+1 \quad 2p+1 \right\rangle^2 = \frac{1}{2} \quad (\text{or } 50\%)$$

7. It is useful here to use some of the general commutator relations found in Appendix C.V.

$$a. \quad [L_x, L_y] = [yp_z - zp_y, zp_x - xp_z]$$

$$= [yp_z, zp_x] - [yp_z, xp_z] - [zp_y, zp_x] + [zp_y, xp_z]$$

$$= [y, z]p_x p_z + z[y, p_x]p_z + y[p_z, z]p_x + yz[p_z, p_x]$$

$$- [y, x]p_z p_z - x[y, p_z]p_z - y[p_z, x]p_z - yx[p_z, p_z]$$

$$- [z, z]p_x p_y - z[z, p_x]p_y - z[p_y, z]p_x - zz[p_y, p_x]$$

$$+ [z, x]p_z p_y + x[z, p_z]p_y + z[p_y, x]p_z + zx[p_y, p_z]$$

As can be easily ascertained, the only non-zero terms are:

$$[L_x, L_y] = y[p_z, z]p_x + x[z, p_z]p_y$$

$$= y(-i\hbar)p_x + x(i\hbar)p_y$$

$$= i\hbar(-yp_x + xp_y)$$

$$= i\hbar L_z$$

$$b. \quad [L_y, L_z] = [zp_x - xp_z, xp_y - yp_x]$$

$$= [zp_x, xp_y] - [zp_x, yp_x] - [xp_z, xp_y] + [xp_z, yp_x]$$

$$= [z, x]p_y p_x + x[z, p_y]p_x + z[p_x, x]p_y + zx[p_x, p_z]$$

$$- [z, y]p_x p_x - y[z, p_x]p_x - z[p_x, y]p_x - zy[p_x, p_x]$$

$$- [x, x]p_y p_z - x[x, p_y]p_z - x[p_z, x]p_y - xx[p_z, p_y]$$

$$+ [x, y]p_x p_z + y[x, p_x]p_z + x[p_z, y]p_x + xy[p_z, p_x]$$

Again, as can be easily ascertained, the only non-zero terms are:

$$[L_y, L_z] = z[p_x, x]p_y + y[x, p_x]p_z$$

$$= z(-i\hbar)p_y + y(i\hbar)p_z$$

$$= i\hbar(-zp_y + yp_z)$$

$$= i\hbar L_x$$

$$c. \quad [L_z, L_x] = [xp_y - yp_x, yp_z - zp_y]$$

$$= [xp_y, yp_z] - [xp_y, zp_y] - [yp_x, yp_z] + [yp_x, zp_y]$$

$$= [x, y]p_z p_y + y[x, p_z]p_y + x[p_y, y]p_z + xy[p_y, p_z]$$

$$- [x, z]p_y p_y - z[x, p_y]p_y - x[p_y, z]p_y - xz[p_y, p_y]$$

$$- [y, y]p_z p_x - y[y, p_z]p_x - y[p_x, y]p_z - yy[p_x, p_z]$$

$$+ [y, z]p_y p_x + z[y, p_y]p_x + y[p_x, z]p_y + yz[p_x, p_y]$$

Again, as can be easily ascertained, the only non-zero terms are:

$$[L_z, L_x] = x[p_y, y]p_z + z[y, p_y]p_x$$

$$\begin{aligned}
&= x(-i\hbar)p_z + z(i\hbar)p_x \\
&= i\hbar(-xp_z + zp_x) \\
&= i\hbar L_y
\end{aligned}$$

$$\begin{aligned}
\text{d. } [L_x, L^2] &= [L_x, L_x^2 + L_y^2 + L_z^2] \\
&= [L_x, L_x^2] + [L_x, L_y^2] + [L_x, L_z^2] \\
&= [L_x, L_y^2] + [L_x, L_z^2] \\
&= [L_x, L_y]L_y + L_y[L_x, L_y] + [L_x, L_z]L_z + L_z[L_x, L_z] \\
&= (i\hbar L_z)L_y + L_y(i\hbar L_z) + (-i\hbar L_y)L_z + L_z(-i\hbar L_y) \\
&= (i\hbar)(L_zL_y + L_yL_z - L_yL_z - L_zL_y) \\
&= (i\hbar)([L_z, L_y] + [L_y, L_z]) = 0
\end{aligned}$$

$$\begin{aligned}
\text{e. } [L_y, L^2] &= [L_y, L_x^2 + L_y^2 + L_z^2] \\
&= [L_y, L_x^2] + [L_y, L_y^2] + [L_y, L_z^2] \\
&= [L_y, L_x^2] + [L_y, L_z^2] \\
&= [L_y, L_x]L_x + L_x[L_y, L_x] + [L_y, L_z]L_z + L_z[L_y, L_z] \\
&= (-i\hbar L_z)L_x + L_x(-i\hbar L_z) + (i\hbar L_x)L_z + L_z(i\hbar L_x) \\
&= (i\hbar)(-L_zL_x - L_xL_z + L_xL_z + L_zL_x) \\
&= (i\hbar)([L_x, L_z] + [L_z, L_x]) = 0
\end{aligned}$$

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

8. Use the general angular momentum relationships:

$$J^2|j, m\rangle = \hbar^2(j(j+1))|j, m\rangle$$

$$J_z|j, m\rangle = \hbar m|j, m\rangle,$$

and the information used in exercise 5, namely that:

$$L_x = \frac{1}{2}(L_+ + L_-)$$

$$L_+ Y_{l,m} = \sqrt{l(l+1) - m(m+1)} \hbar Y_{l,m+1}$$

$$L_- Y_{l,m} = \sqrt{l(l+1) - m(m-1)} \hbar Y_{l,m-1}$$

Given that:

$$Y_{0,0}(\theta, \phi) = \frac{1}{\sqrt{4}} = |0,0\rangle$$

$$Y_{1,0}(\theta, \phi) = \sqrt{\frac{3}{4}} \cos \theta = |1,0\rangle.$$

$$\begin{aligned} \text{a. } \quad L_z|0,0\rangle &= 0 \\ L^2|0,0\rangle &= 0 \end{aligned}$$

Since  $L^2$  and  $L_z$  commute you would expect  $|0,0\rangle$  to be simultaneous eigenfunctions of both.

$$\begin{aligned} \text{b. } \quad L_x|0,0\rangle &= 0 \\ L_z|0,0\rangle &= 0 \end{aligned}$$

$L_x$  and  $L_z$  do not commute. It is unexpected to find a simultaneous eigenfunction ( $|0,0\rangle$ ) of both ... for sure these operators do not have the same full set of eigenfunctions.

$$\begin{aligned} \text{c. } \quad L_z|1,0\rangle &= 0 \\ L^2|1,0\rangle &= 2\hbar^2|1,0\rangle \end{aligned}$$

Again since  $L^2$  and  $L_z$  commute you would expect  $|1,0\rangle$  to be simultaneous eigenfunctions of both.

$$\begin{aligned} \text{d. } \quad L_x|1,0\rangle &= \frac{\sqrt{2}}{2} \hbar |1,-1\rangle + \frac{\sqrt{2}}{2} \hbar |1,1\rangle \\ L_z|1,0\rangle &= 0 \end{aligned}$$

Again,  $L_x$  and  $L_z$  do not commute. Therefore it is expected to find differing sets of eigenfunctions for both.

9. For:

$$\begin{aligned} \psi(x,y) &= \frac{1}{2L_x} \frac{1}{2} \frac{1}{2L_y} \frac{1}{2} e^{in_x x/L_x} - e^{-in_x x/L_x} e^{in_y y/L_y} - e^{-in_y y/L_y} \\ \langle \psi(x,y) | \psi(x,y) \rangle &= ? \end{aligned}$$

Let:  $a_x = \frac{n_x}{L_x}$ , and  $a_y = \frac{n_y}{L_y}$  and using Euler's formula, expand the exponentials into Sin and Cos terms.

$$\begin{aligned} \psi(x,y) &= \frac{1}{2L_x} \frac{1}{2} \frac{1}{2L_y} \frac{1}{2} [\cos(a_x x) + i\sin(a_x x) - \cos(a_x x) + \\ &\quad i\sin(a_x x)] [\cos(a_y y) + i\sin(a_y y) - \cos(a_y y) + i\sin(a_y y)] \end{aligned}$$

$$\psi(x,y) = \frac{1}{2L_x} \frac{1}{2} \frac{1}{2L_y} \frac{1}{2} 2i\sin(a_x x) 2i\sin(a_y y)$$

$$\psi(x,y) = -\frac{2}{L_x} \frac{1}{2} \frac{2}{L_y} \frac{1}{2} \sin(a_x x) \sin(a_y y)$$

$$\begin{aligned} \langle \psi(x,y) | \psi(x,y) \rangle &= -\frac{2}{L_x} \frac{1}{2} \frac{2}{L_y} \frac{1}{2} \sin^2(a_x x) \sin^2(a_y y) \int dx dy \\ &= \frac{2}{L_x} \frac{2}{L_y} \int \sin^2(a_x x) \sin^2(a_y y) dx dy \end{aligned}$$

Using the integral:

L

$$\int_0^L \sin^2 \frac{n\pi x}{L} dx = \frac{L}{2},$$

$$\langle (x,y) | (x,y) \rangle = \frac{2}{L_x} \frac{2}{L_y} \frac{L_x}{2} \frac{L_y}{2} = 1$$

10.

$$\begin{aligned} \langle (x,y) | p_x | (x,y) \rangle &= \frac{2}{L_y} \int_0^{L_y} \sin^2(a_y y) dy \frac{2}{L_x} \int_0^{L_x} \sin(a_x x) \left(-i\hbar \frac{\partial}{\partial x}\right) \sin(a_x x) dx \\ &= \frac{-i\hbar 2a_x}{L_x} \int_0^{L_x} \sin(a_x x) \cos(a_x x) dx \end{aligned}$$

But the integral:

$$\int_0^{L_x} \cos(a_x x) \sin(a_x x) dx = 0,$$

$$\langle (x,y) | p_x | (x,y) \rangle = 0$$

$$11. \langle 0 | x^2 | 0 \rangle = -\frac{1}{2} \int_0^{\infty} e^{-x^2/2} (x^2) e^{-x^2/2} dx$$

$$= \frac{1}{2} \int_0^{\infty} x^2 e^{-x^2} dx$$

Using the integral:

$$\int_0^{\infty} x^{2n} e^{-x^2} dx = \frac{1 \cdot 3 \cdots (2n-1)}{2^{n+1}} \frac{1}{2^{n+1}},$$

$$\langle 0 | x^2 | 0 \rangle = -\frac{1}{2} \frac{1}{2^2} \frac{1}{3} \frac{1}{2}$$

$$\langle 0 | x^2 | 0 \rangle = \frac{1}{2}$$

$$\langle 1 | x^2 | 1 \rangle = \frac{4}{3} \frac{1}{2} \int_0^{\infty} x e^{-x^2/2} (x^2) x e^{-x^2/2} dx$$

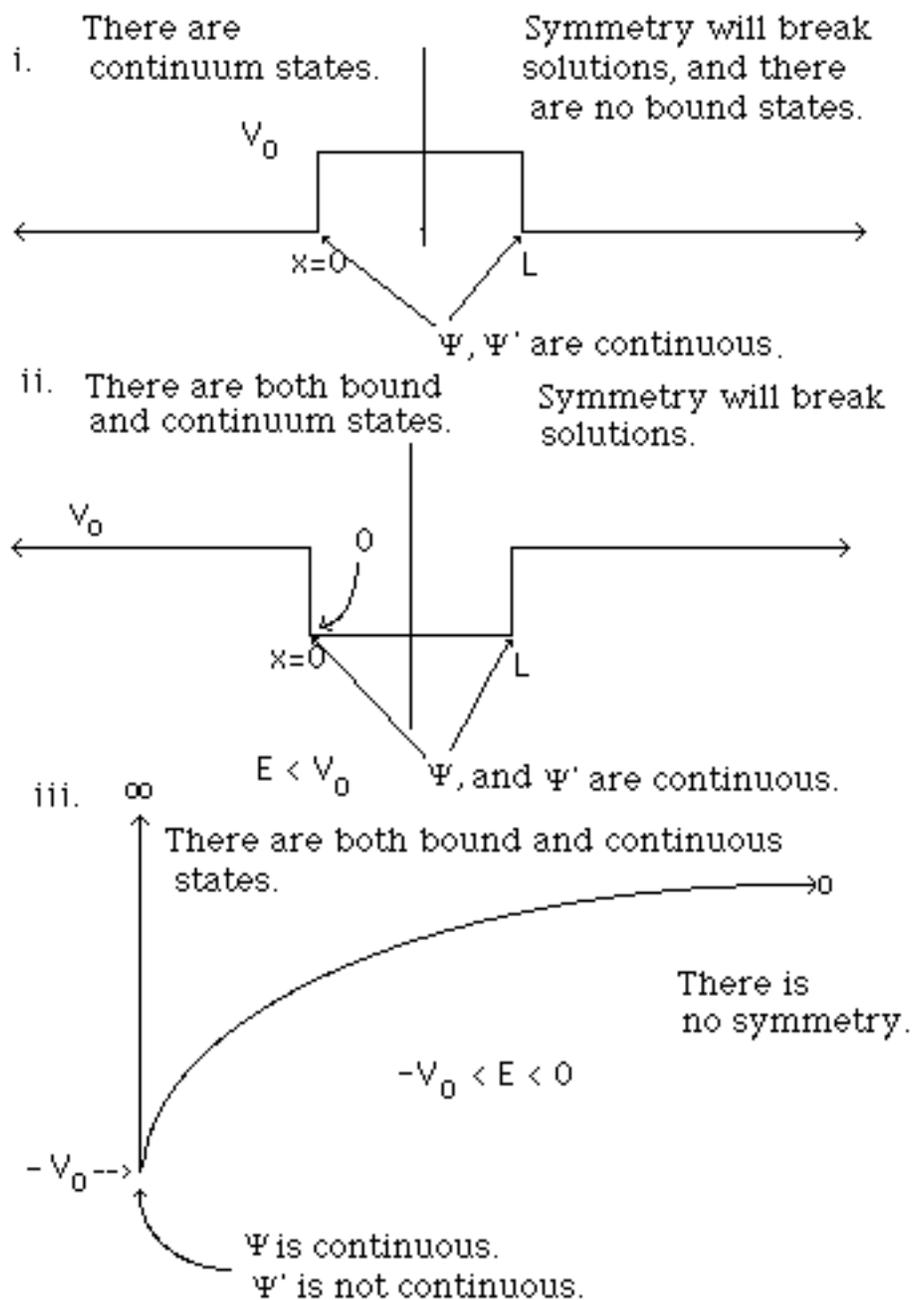
$$= \frac{4}{2} \int_0^{\infty} x^3 e^{-x^2/2} dx$$

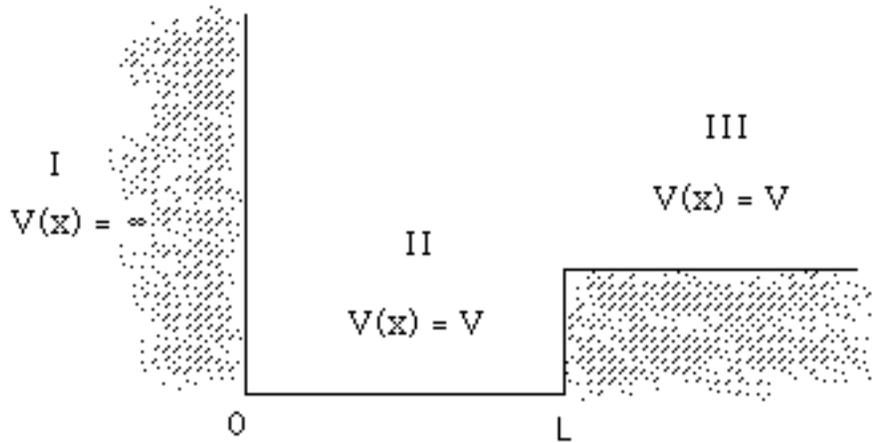
Using the previously defined integral:

$$\langle |x^2| \rangle = \frac{4}{2} \int_0^{\infty} x^3 e^{-x^2/2} dx = \frac{3}{2^3} \frac{1}{5}$$

$$\langle |x^2| \rangle = \frac{3}{2}$$

12.





a.  $I(x) = 0$

$$II(x) = Ae^{i\sqrt{2mE/\hbar^2} x} + Be^{-i\sqrt{2mE/\hbar^2} x}$$

$$III(x) = A'e^{i\sqrt{2m(V-E)/\hbar^2} x} + B'e^{-i\sqrt{2m(V-E)/\hbar^2} x}$$

b. I II

$$I(0) = II(0)$$

$$I(0) = 0 = II(0) = Ae^{i\sqrt{2mE/\hbar^2} (0)} + Be^{-i\sqrt{2mE/\hbar^2} (0)}$$

$$0 = A + B$$

$$B = -A$$

$$I'(0) = II'(0) \quad (\text{this gives no useful information since } I'(x) \text{ does not exist at } x = 0)$$

II III

$$II(L) = III(L)$$

$$Ae^{i\sqrt{2mE/\hbar^2} L} + Be^{-i\sqrt{2mE/\hbar^2} L} = A'e^{i\sqrt{2m(V-E)/\hbar^2} L}$$

$$+ B'e^{-i\sqrt{2m(V-E)/\hbar^2} L}$$

$$II'(L) = III'(L)$$

$$A(i\sqrt{2mE/\hbar^2})e^{i\sqrt{2mE/\hbar^2} L} - B(i\sqrt{2mE/\hbar^2})e^{-i\sqrt{2mE/\hbar^2} L}$$

$$= A'(i\sqrt{2m(V-E)/\hbar^2})e^{i\sqrt{2m(V-E)/\hbar^2} L}$$

$$- B'(i\sqrt{2m(V-E)/\hbar^2})e^{-i\sqrt{2m(V-E)/\hbar^2} L}$$

c. as  $x \rightarrow -\infty$ ,  $I(x) = 0$

$$\text{as } x \rightarrow +\infty, III(x) = 0 \quad A' = 0$$

d. Rewrite the equations for  $\psi(0) = \psi(L)$ ,  $\psi'(0) = \psi'(L)$ , and  $\psi''(L) = \psi''(0)$  using the information in 13c:

$$B = -A \text{ (eqn. 1)}$$

$$Ae^{i\sqrt{2mE/\hbar^2} L} + Be^{-i\sqrt{2mE/\hbar^2} L} = B'e^{-i\sqrt{2m(V-E)/\hbar^2} L} \text{ (eqn. 2)}$$

$$A(i\sqrt{2mE/\hbar^2})e^{i\sqrt{2mE/\hbar^2} L} - B(i\sqrt{2mE/\hbar^2})e^{-i\sqrt{2mE/\hbar^2} L} = -B'(i\sqrt{2m(V-E)/\hbar^2})e^{-i\sqrt{2m(V-E)/\hbar^2} L} \text{ (eqn. 3)}$$

substituting (eqn. 1) into (eqn. 2):

$$Ae^{i\sqrt{2mE/\hbar^2} L} - Ae^{-i\sqrt{2mE/\hbar^2} L} = B'e^{-i\sqrt{2m(V-E)/\hbar^2} L} \\ A(\cos(\sqrt{2mE/\hbar^2} L) + i\sin(\sqrt{2mE/\hbar^2} L)) - A(\cos(\sqrt{2mE/\hbar^2} L) - i\sin(\sqrt{2mE/\hbar^2} L)) = B'e^{-i\sqrt{2m(V-E)/\hbar^2} L}$$

$$2Ai\sin(\sqrt{2mE/\hbar^2} L) = B'e^{-i\sqrt{2m(V-E)/\hbar^2} L}$$

$$\sin(\sqrt{2mE/\hbar^2} L) = \frac{B'}{2Ai} e^{-i\sqrt{2m(V-E)/\hbar^2} L} \text{ (eqn. 4)}$$

substituting (eqn. 1) into (eqn. 3):

$$A(i\sqrt{2mE/\hbar^2})e^{i\sqrt{2mE/\hbar^2} L} + A(i\sqrt{2mE/\hbar^2})e^{-i\sqrt{2mE/\hbar^2} L} = -B'(i\sqrt{2m(V-E)/\hbar^2})e^{-i\sqrt{2m(V-E)/\hbar^2} L}$$

$$A(i\sqrt{2mE/\hbar^2})(\cos(\sqrt{2mE/\hbar^2} L) + i\sin(\sqrt{2mE/\hbar^2} L)) + A(i\sqrt{2mE/\hbar^2})(\cos(\sqrt{2mE/\hbar^2} L) - i\sin(\sqrt{2mE/\hbar^2} L)) = -B'(i\sqrt{2m(V-E)/\hbar^2})e^{-i\sqrt{2m(V-E)/\hbar^2} L}$$

$$2Ai\sqrt{2mE/\hbar^2} \cos(\sqrt{2mE/\hbar^2} L) = -B'i\sqrt{2m(V-E)/\hbar^2} e^{-i\sqrt{2m(V-E)/\hbar^2} L}$$

$$2Ai\sqrt{2mE/\hbar^2} \cos(\sqrt{2mE/\hbar^2} L) = -B'i\sqrt{2m(V-E)/\hbar^2} e^{-i\sqrt{2m(V-E)/\hbar^2} L}$$

$$\cos(\sqrt{2mE/\hbar^2} L) = -\frac{B'i\sqrt{2m(V-E)/\hbar^2}}{2Ai\sqrt{2mE/\hbar^2}} e^{-i\sqrt{2m(V-E)/\hbar^2} L}$$

$$\cos(\sqrt{2mE/\hbar^2} L) = -\frac{B'\sqrt{V-E}}{2A\sqrt{E}} e^{-i\sqrt{2m(V-E)/\hbar^2} L} \text{ (eqn. 5)}$$

Dividing (eqn. 4) by (eqn. 5):

$$\frac{\sin(\sqrt{2mE/\hbar^2} L)}{\cos(\sqrt{2mE/\hbar^2} L)} = \frac{B' - 2A\sqrt{E} e^{-i\sqrt{2m(V-E)/\hbar^2} L}}{2Ai B' \sqrt{V-E} e^{-i\sqrt{2m(V-E)/\hbar^2} L}}$$

$$\tan(\sqrt{2mE/\hbar^2} L) = -\frac{E}{V-E}^{1/2}$$

e. As  $V \rightarrow \infty$ ,  $\tan(\sqrt{2mE/\hbar^2} L) = 0$

So,  $\sqrt{2mE/\hbar^2} L = n$

$$E_n = \frac{n^2 \hbar^2}{2mL^2}$$

### Problems

1. a.  $\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right)$   
 $P_n(x) dx = |\psi_n(x)|^2 dx$

The probability that the particle lies in the interval  $0 \leq x \leq \frac{L}{4}$  is given by:

$$P_n = \int_0^{\frac{L}{4}} P_n(x) dx = \frac{2}{L} \int_0^{\frac{L}{4}} \sin^2\left(\frac{n\pi x}{L}\right) dx$$

This integral can be integrated to give (using integral equation 10 with  $u = \frac{n\pi x}{L}$ ):

$$P_n = \frac{L}{n} \int_0^{\frac{n\pi}{4}} \sin^2 u \, d\left(\frac{n\pi x}{L}\right)$$

$$P_n = \frac{L}{n} \int_0^{\frac{n\pi}{4}} \sin^2 u \, du$$

$$P_n = \frac{2}{n} - \frac{1}{4} \sin\left(\frac{2n\pi}{4}\right) + \frac{n\pi}{2} \left(\frac{1}{4}\right)$$

$$= \frac{2}{n} - \frac{1}{4} \sin\left(\frac{2n\pi}{4}\right) + \frac{n\pi}{(2)(4)}$$

$$= \frac{1}{4} - \frac{1}{2n} \sin \frac{n}{2}$$

b. If  $n$  is even,  $\sin \frac{n}{2} = 0$  and  $P_n = \frac{1}{4}$ .

If  $n$  is odd and  $n = 1, 5, 9, 13, \dots$   $\sin \frac{n}{2} = 1$

$$\text{and } P_n = \frac{1}{4} - \frac{1}{2n}$$

If  $n$  is odd and  $n = 3, 7, 11, 15, \dots$   $\sin \frac{n}{2} = -1$

$$\text{and } P_n = \frac{1}{4} + \frac{1}{2n}$$

The higher  $P_n$  is when  $n=3$ . Then  $P_n = \frac{1}{4} + \frac{1}{6}$

$$P_n = \frac{1}{4} + \frac{1}{6} = 0.303$$

$$\begin{aligned} \text{c. } \psi(t) &= e^{-\frac{iHt}{\hbar}} [a \psi_n + b \psi_m] = a \psi_n e^{-\frac{iE_n t}{\hbar}} + b \psi_m e^{-\frac{iE_m t}{\hbar}} \\ H \psi &= a \psi_n E_n e^{-\frac{iE_n t}{\hbar}} + b \psi_m E_m e^{-\frac{iE_m t}{\hbar}} \\ \langle H \rangle &= |a|^2 E_n + |b|^2 E_m + a^* b e^{\frac{i(E_n - E_m)t}{\hbar}} \langle \psi_n | H | \psi_m \rangle \\ &\quad + b^* a e^{-\frac{i(E_m - E_n)t}{\hbar}} \langle \psi_m | H | \psi_n \rangle \end{aligned}$$

Since  $\langle \psi_n | H | \psi_m \rangle$  and  $\langle \psi_m | H | \psi_n \rangle$  are zero,

$$\langle H \rangle = |a|^2 E_n + |b|^2 E_m \text{ (note the time independence)}$$

d. The fraction of systems observed in  $\psi_n$  is  $|a|^2$ . The possible energies measured are  $E_n$  and  $E_m$ . The probabilities of measuring each of these energies is  $|a|^2$  and  $|b|^2$ .

e. Once the system is observed in  $\psi_n$ , it stays in  $\psi_n$ .

$$\text{f. } P(E_n) = \int_0^L |\psi_n|^2 dx = |c_n|^2$$

$$\begin{aligned} c_n &= \int_0^L \sqrt{\frac{2}{L}} \sin \frac{n x}{L} \sqrt{\frac{30}{L^5}} x(L-x) dx \\ &= \sqrt{\frac{60}{L^6}} \int_0^L x(L-x) \sin \frac{n x}{L} dx \end{aligned}$$

$$= \sqrt{\frac{60}{L^6}} \int_0^L x \sin \frac{n x}{L} dx - \int_0^L x^2 \sin \frac{n x}{L} dx$$

These integrals can be evaluated from integral equations 14 and 16 to give:

$$c_n = \sqrt{\frac{60}{L^6}} \left[ \frac{L^2}{n^2} \sin \frac{n x}{L} - \frac{Lx}{n} \cos \frac{n x}{L} \right]_0^L$$

$$- \sqrt{\frac{60}{L^6}} \left[ \frac{2xL^2}{n^2} \sin \frac{n x}{L} - \frac{n^2 - 2x^2}{L^2} - 2 \frac{L^3}{n^3} \cos \frac{n x}{L} \right]_0^L$$

$$c_n = \sqrt{\frac{60}{L^6}} \left\{ \frac{L^3}{n^2} (\sin(n) - \sin(0)) \right.$$

$$\left. - \frac{L^2}{n} (L \cos(n) - 0 \cos(0)) \right.$$

$$\left. - \left( \frac{2L^2}{n^2} (L \sin(n) - 0 \sin(0)) \right. \right.$$

$$\left. - (n^2 - 2) \frac{L^3}{n^3} \cos(n) \right.$$

$$\left. + \frac{n^2 - 2}{L^2} - 2 \frac{L^3}{n^3} \cos(0) \right\}$$

$$c_n = L^{-3} \sqrt{60} \left\{ -\frac{L^3}{n} \cos(n) + (n^2 - 2) \frac{L^3}{n^3} \cos(n) \right.$$

$$\left. + \frac{2L^3}{n^3} \right\}$$

$$c_n = \sqrt{60} \left[ -\frac{1}{n} (-1)^n + (n^2 - 2) \frac{1}{n^3} (-1)^n + \frac{2}{n^3} \right]$$

$$c_n = \sqrt{60} \left[ \frac{-1}{n} + \frac{1}{n} - \frac{2}{n^3} (-1)^n + \frac{2}{n^3} \right]$$

$$c_n = \frac{2\sqrt{60}}{n^3} (1 - (-1)^n)$$

$$|c_n|^2 = \frac{4(60)}{n^6} (1 - (-1)^n)^2$$

If n is even then  $c_n = 0$

$$\text{If n is odd then } c_n = \frac{(4)(60)(4)}{n^6} = \frac{960}{n^6}$$

The probability of making a measurement of the energy and obtaining one of the eigenvalues, given by:

$$E_n = \frac{n^2 \hbar^2}{2mL^2} \text{ is:}$$

$$P(E_n) = 0 \text{ if n is even}$$

$$P(E_n) = \frac{960}{n^6} \text{ if } n \text{ is odd}$$

$$\begin{aligned} \text{g. } \langle |H| \rangle &= \int_0^L \frac{30}{L^5} \frac{1}{2} x(L-x) \frac{-\hbar^2}{2m} \frac{d^2}{dx^2} \frac{30}{L^5} \frac{1}{2} x(L-x) dx \\ &= \frac{30}{L^5} \frac{-\hbar^2}{2m} \int_0^L x(L-x) \frac{d^2}{dx^2} (xL-x^2) dx \\ &= \frac{-15\hbar^2}{mL^5} \int_0^L x(L-x)(-2) dx \\ &= \frac{30\hbar^2}{mL^5} \int_0^L xL-x^2 dx \\ &= \frac{30\hbar^2}{mL^5} \left[ L \frac{x^2}{2} - \frac{x^3}{3} \right]_0^L \\ &= \frac{30\hbar^2}{mL^5} \left[ \frac{L^3}{2} - \frac{L^3}{3} \right] \\ &= \frac{30\hbar^2}{mL^5} \frac{1}{2} \frac{1}{3} \\ &= \frac{30\hbar^2}{6mL^5} = \frac{5\hbar^2}{mL^2} \end{aligned}$$

$$2. \langle |H| \rangle = \sum_{ij} C_i^* e^{\frac{iE_i t}{\hbar}} \langle i|H|j \rangle e^{-\frac{iE_j t}{\hbar}} C_j$$

$$\text{Since } \langle i|H|j \rangle = E_j \delta_{ij}$$

$$\langle |H| \rangle = \sum_j C_j^* C_j E_j e^{\frac{i(E_j - E_j)t}{\hbar}}$$

$$\langle |H| \rangle = \sum_j C_j^* C_j E_j \text{ (not time dependent)}$$

For other properties:

$$\langle |A| \rangle = \sum_{ij} C_i^* e^{\frac{iE_i t}{\hbar}} \langle i|A|j \rangle e^{-\frac{iE_j t}{\hbar}} C_j$$

but,  $\langle i|A|j \rangle$  does not necessarily =  $a_j \delta_{ij}$ .

This is only true if  $[A, H] = 0$ .

$$\langle |A\rangle = \sum_{ij} C_i^* C_j e^{\frac{i(E_i - E_j)t}{\hbar}} \langle i|A\rangle |j\rangle$$

Therefore, in general, other properties are time dependent.

3. For a particle in a box in its lowest quantum state:

$$\begin{aligned} &= \sqrt{\frac{2}{L}} \sin \frac{x}{L} \\ \langle x \rangle &= \int_0^L \psi^* x \psi dx \\ &= \frac{2}{L} \int_0^L x \sin^2 \frac{x}{L} dx \end{aligned}$$

Using integral equation 18:

$$\begin{aligned} &= \frac{2}{L} \left[ \frac{x^2}{4} - \frac{xL}{4} \sin \frac{2x}{L} - \frac{L^2}{8} \cos \frac{2x}{L} \right]_0^L \\ &= \frac{2}{L} \left[ \frac{L^2}{4} - \frac{L^2}{8} (\cos(2)) - \cos(0) \right] \\ &= \frac{2}{L} \frac{L^2}{4} \\ &= \frac{L}{2} \\ \langle x^2 \rangle &= \int_0^L \psi^* x^2 \psi dx \\ &= \frac{2}{L} \int_0^L x^2 \sin^2 \frac{x}{L} dx \end{aligned}$$

Using integral equation 19:

$$\begin{aligned} &= \frac{2}{L} \left[ \frac{x^3}{6} - \frac{x^2 L}{4} - \frac{L^3}{8} \sin \frac{2x}{L} - \frac{xL^2}{4} \cos \frac{2x}{L} \right]_0^L \\ &= \frac{2}{L} \left[ \frac{L^3}{6} - \frac{L^2}{4} (L \cos(2)) - (0) \cos(0) \right] \\ &= \frac{2}{L} \left[ \frac{L^3}{6} - \frac{L^3}{4} \right] \\ &= \frac{L^2}{3} - \frac{L^2}{2} \end{aligned}$$

$$\begin{aligned}
\langle p \rangle &= \int_0^L \psi^* p \psi \, dx \\
&= \int_0^L \frac{2}{L} \sin \frac{x}{L} \frac{\hbar}{i} \frac{d}{dx} \sin \frac{x}{L} \, dx \\
&= \frac{2\hbar}{L^2 i} \int_0^L \sin \frac{x}{L} \cos \frac{x}{L} \, dx \\
&= \frac{2\hbar}{L i} \int_0^L \sin \frac{x}{L} \cos \frac{x}{L} \, d \frac{x}{L}
\end{aligned}$$

Using integral equation 15 (with  $u = \frac{x}{L}$ ):

$$= \frac{2\hbar}{L i} \left[ -\frac{1}{2} \cos^2 \left( \frac{x}{L} \right) \right]_0^L = 0$$

$$\begin{aligned}
\langle p^2 \rangle &= \int_0^L \psi^* p^2 \psi \, dx \\
&= \int_0^L \frac{2}{L} \sin \frac{x}{L} (-\hbar^2) \frac{d^2}{dx^2} \sin \frac{x}{L} \, dx \\
&= \frac{2}{L^3} \int_0^L \sin^2 \frac{x}{L} \, dx \\
&= \frac{2}{L^2} \int_0^L \sin^2 \frac{x}{L} \, d \frac{x}{L}
\end{aligned}$$

Using integral equation 10 (with  $u = \frac{x}{L}$ ):

$$\begin{aligned}
&= \frac{2}{L^2} \left[ \frac{1}{4} \sin(2u) + \frac{u}{2} \right]_0^L \\
&= \frac{2}{L^2} \left[ \frac{1}{4} \sin(2) + \frac{1}{2} \right] = \frac{2\hbar^2}{L^2}
\end{aligned}$$

$$\begin{aligned} x &= \langle x^2 \rangle - \langle x \rangle^2 \frac{1}{2} \\ &= \frac{L^2}{3} - \frac{L^2}{2} \frac{1}{2} - \frac{L^2}{4} \frac{1}{2} \\ &= L \frac{1}{12} - \frac{1}{2} \frac{1}{2} \end{aligned}$$

$$\begin{aligned} p &= \langle p^2 \rangle - \langle p \rangle^2 \frac{1}{2} \\ &= \frac{2\hbar^2}{L^2} - 0 \frac{1}{2} = \frac{\hbar}{L} \end{aligned}$$

$$\begin{aligned} x \cdot p &= \hbar \frac{1}{12} - \frac{1}{2} \frac{1}{2} \frac{1}{2} \\ &= \frac{\hbar}{2} \frac{4}{12} - \frac{4}{2} \frac{1}{2} \\ &= \frac{\hbar}{2} \frac{2}{3} - 2 \frac{1}{2} \end{aligned}$$

Finally,  $\frac{\hbar}{2} \frac{2}{3} - 2 \frac{1}{2} > \frac{\hbar}{2} \frac{(3)^2}{3} - 2 \frac{1}{2} = \frac{\hbar}{2}$

$$x \cdot p > \frac{\hbar}{2}$$

4. a.  $\int_0^{L/4} P(x) dx = \int_0^{L/4} \frac{1}{L} dx = \frac{1}{L} x \Big|_0^{L/4}$   
 $= \frac{1}{L} \frac{L}{4} = \frac{1}{4} = 25\%$

$$P_{\text{classical}} = \frac{1}{4} \quad (\text{for interval } 0 - L/4)$$

b. This was accomplished in problem 1a. to give:

$$P_n = \frac{1}{4} - \frac{1}{2} \frac{1}{n} \sin \frac{n}{2}$$

(for interval 0 - L/4)

c.  $\lim_n P_{\text{quantum}} = \lim_n \left( \frac{1}{4} - \frac{1}{2} \frac{1}{n} \sin \frac{n}{2} \right)$

$$\lim_n P_{\text{quantum}} = \frac{1}{4}$$

Therefore as n becomes large the classical limit is approached.

5. a. The Schrödinger equation for a Harmonic Oscillator in 1-dimensional coordinate representation,  $\mathbf{H} \psi(x) = E_x \psi(x)$ , with the Hamiltonian defined as:  $\mathbf{H} = \frac{-\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} kx^2$ , becomes:

$$\frac{-\hbar^2}{2m} \frac{d^2}{dx^2} \psi(x) + \frac{1}{2} kx^2 \psi(x) = E_x \psi(x).$$

b. The transformation of the kinetic energy term to the momentum representation is trivial:  $\mathbf{T} = \frac{\mathbf{p}_x^2}{2m}$ . In order to maintain the commutation relation  $[\mathbf{x}, \mathbf{p}_x] = i\hbar$  and keep the  $\mathbf{p}$  operator unchanged the coordinate operator must become  $\mathbf{x} = i\hbar \frac{d}{dp_x}$ . The Schrödinger equation for a Harmonic Oscillator in 1-dimensional momentum representation,  $\mathbf{H} \psi(p_x) = E_{p_x} \psi(p_x)$ , with the Hamiltonian defined as:  $\mathbf{H} = \frac{1}{2m} p_x^2 - \frac{k\hbar^2}{2} \frac{d^2}{dp_x^2}$ , becomes:

$$\frac{1}{2m} p_x^2 \psi(p_x) - \frac{k\hbar^2}{2} \frac{d^2}{dp_x^2} \psi(p_x) = E_{p_x} \psi(p_x).$$

c. For the wavefunction  $\psi(x) = C \exp\left(-\sqrt{mk} \frac{x^2}{2\hbar}\right)$ ,

let  $a = \frac{\sqrt{mk}}{2\hbar}$ , and hence  $\psi(x) = C \exp(-ax^2)$ . Evaluating the derivatives of this expression gives:

$$\begin{aligned} \frac{d}{dx} \psi(x) &= \frac{d}{dx} C \exp(-ax^2) = -2axC \exp(-ax^2) \\ \frac{d^2}{dx^2} \psi(x) &= \frac{d^2}{dx^2} C \exp(-ax^2) = \frac{d}{dx} (-2axC \exp(-ax^2)) \\ &= (-2axC) (-2ax \exp(-ax^2)) + (-2aC) (\exp(-ax^2)) \\ &= (4a^2x^2 - 2a) C \exp(-ax^2). \end{aligned}$$

$\mathbf{H} \psi(x) = E_x \psi(x)$  then becomes:

$$\mathbf{H} \psi(x) = \frac{-\hbar^2}{2m} (4a^2x^2 - 2a) + \frac{1}{2} kx^2 \psi(x).$$

Clearly the energy (eigenvalue) expression must be independent of  $x$  and the two terms containing  $x^2$  terms must cancel upon insertion of  $a$ :

$$\begin{aligned} E_x &= \frac{-\hbar^2}{2m} \left( 4 \frac{\sqrt{mk}}{2\hbar} x^2 - 2 \frac{\sqrt{mk}}{2\hbar} \right) + \frac{1}{2} kx^2 \\ &= \frac{-\hbar^2}{2m} \frac{4mkx^2}{4\hbar^2} + \frac{\hbar^2}{2m} \frac{2\sqrt{mk}}{2\hbar} + \frac{1}{2} kx^2 \\ &= -\frac{1}{2} kx^2 + \frac{\hbar\sqrt{mk}}{2m} + \frac{1}{2} kx^2 \end{aligned}$$

$$= \frac{\hbar\sqrt{mk}}{2m}$$

Normalization of  $\psi(x)$  to determine the constant C yields the equation:

$$C^2 \int_{-\infty}^{\infty} \exp\left(-\sqrt{mk} \frac{x^2}{\hbar}\right) dx = 1.$$

Using integral equation (1) gives:

$$C^2 \int_{-\infty}^{\infty} \sqrt{\frac{mk}{\hbar}} \frac{1}{2} dx = 1$$

$$C^2 \frac{\hbar}{\sqrt{mk}} \frac{1}{2} = 1$$

$$C^2 = \frac{\sqrt{mk}}{\hbar} \frac{1}{2}$$

$$C = \frac{\sqrt{mk}}{\hbar} \frac{1}{4}$$

Therefore,  $\psi(x) = \frac{\sqrt{mk}}{\hbar} \frac{1}{4} \exp\left(-\sqrt{mk} \frac{x^2}{2\hbar}\right)$ .

d. Proceeding analogous to part c, for a wavefunction in momentum space  $\psi(p) = C \exp(-p^2)$ , evaluating the derivatives of this expression gives:

$$\frac{d}{dp} \psi(p) = \frac{d}{dp} C \exp(-p^2) = -2pC \exp(-p^2)$$

$$\begin{aligned} \frac{d^2}{dp^2} \psi(p) &= \frac{d^2}{dp^2} C \exp(-p^2) = \frac{d}{dp} (-2pC \exp(-p^2)) \\ &= (-2pC) (-2p \exp(-p^2)) + (-2C) (\exp(-p^2)) \\ &= (4p^2 - 2) C \exp(-p^2). \end{aligned}$$

**H**  $\psi(p) = E_p \psi(p)$  then becomes:

$$\mathbf{H} \psi(p) = \frac{1}{2m} p^2 - \frac{\hbar^2 k}{2} (4p^2 - 2) \psi(p)$$

Once again the energy (eigenvalue) expression corresponding to  $E_p$  must be independent of  $p$  and the two terms containing  $p^2$  terms must cancel with the appropriate choice of  $C$ . We also desire our choice of  $C$  to give us the same energy we found in part c (in coordinate space).

$$E_p = \frac{1}{2m} p^2 - \frac{\hbar^2 k}{2} (4p^2 - 2)$$

Therefore we can find  $C$  either of two ways:

$$(1) \quad \frac{1}{2m} p^2 = \frac{k\hbar^2}{2} \psi^2, \text{ or}$$

$$(2) \quad \frac{k\hbar^2}{2} \psi^2 = \frac{\hbar\sqrt{mk}}{2m}.$$

Both equations yield  $\psi = 2\hbar\sqrt{mk}^{-1/2}$ .

Normalization of  $\psi(p)$  to determine the constant C yields the equation:

$$C^2 \int_{-\infty}^{\infty} \exp(-2p^2) dp = 1.$$

Using integral equation (1) gives:

$$C^2 \int_{-\infty}^{\infty} \frac{1}{2} \sqrt{2} (2)^{-1/2} dp = 1$$

$$C^2 \sqrt{2} \int_{-\infty}^{\infty} \frac{1}{2} \sqrt{2} (2)^{-1/2} dp = 1$$

$$C^2 \sqrt{2} \int_{-\infty}^{\infty} \frac{1}{2} \sqrt{2} dp = 1$$

$$C^2 \int_{-\infty}^{\infty} dp = 1$$

$$C^2 = \frac{1}{\sqrt{2\pi}}$$

$$C = \frac{1}{\sqrt{2\pi}}$$

Therefore,  $\psi(p) = \frac{1}{\sqrt{2\pi}} \exp(-p^2/(2\hbar\sqrt{mk}))$ .

Showing that  $\psi(p)$  is the proper fourier transform of  $\psi(x)$  suggests that the fourier integral theorem should hold for the two wavefunctions  $\psi(x)$  and  $\psi(p)$  we have obtained, e.g.

$$\psi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \psi(x) e^{ipx/\hbar} dx, \text{ for}$$

$$\psi(x) = \frac{\sqrt{mk}}{\hbar} \frac{1}{4} \exp(-\sqrt{mk} \frac{x^2}{2\hbar}), \text{ and}$$

$$\psi(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \psi(x) e^{ipx/\hbar} dx.$$

So, verify that:

$$\frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} \psi(x) e^{ipx/\hbar} dx$$

$$= \frac{1}{\sqrt{2\hbar}} \frac{\sqrt{mk}}{\hbar} \frac{1}{4} \exp\left(-\sqrt{mk} \frac{x^2}{2\hbar}\right) e^{ipx/\hbar} dx .$$

Working with the right-hand side of the equation:

$$= \frac{1}{\sqrt{2\hbar}} \frac{\sqrt{mk}}{\hbar} \frac{1}{4} \exp\left(-\sqrt{mk} \frac{x^2}{2\hbar}\right) \cos \frac{px}{\hbar} + i \sin \frac{px}{\hbar} dx ,$$

the Sin term is odd and the integral will therefore vanish. The remaining integral can be evaluated using the given expression:

$$e^{-x^2} \cos bx dx = \sqrt{\frac{\pi}{b^2}} e^{-b^2/4}$$

$$= \frac{1}{\sqrt{2\hbar}} \frac{\sqrt{mk}}{\hbar} \frac{1}{4} \exp\left(-\frac{\sqrt{mk}}{2\hbar} x^2\right) \cos \frac{p}{\hbar} x dx$$

$$= \frac{1}{\sqrt{2\hbar}} \frac{\sqrt{mk}}{\hbar} \frac{1}{4} \frac{2\hbar}{\sqrt{mk}} \frac{1}{2} \exp\left[-\frac{p^2}{\hbar^2} \frac{2\hbar}{4\sqrt{mk}}\right]$$

$$= \frac{\sqrt{mk}}{\hbar} \frac{1}{4} \frac{1}{\sqrt{mk}} \frac{1}{2} \exp\left[-\frac{p^2}{2\hbar\sqrt{mk}}\right]$$

$$= \frac{\sqrt{mk}}{mk \hbar} \frac{1}{4} \exp\left[-\frac{p^2}{2\hbar\sqrt{mk}}\right]$$

$$= \frac{1}{\hbar \sqrt{mk}} \frac{1}{4} \exp\left[-\frac{p^2}{2\hbar\sqrt{mk}}\right] = (p)Q.E.D.$$

6. a. The lowest energy level for a particle in a 3-dimensional box is when  $n_1 = 1$ ,  $n_2 = 1$ , and  $n_3 = 1$ . The total energy (with  $L_1 = L_2 = L_3$ ) will be:

$$E_{\text{total}} = \frac{h^2}{8mL^2}(n_1^2 + n_2^2 + n_3^2) = \frac{3h^2}{8mL^2}$$

Note that  $n = 0$  is not possible. The next lowest energy level is when one of the three quantum numbers equals 2 and the other two equal 1:

$$\begin{aligned} n_1 = 1, n_2 = 1, n_3 = 2 \\ n_1 = 1, n_2 = 2, n_3 = 1 \\ n_1 = 2, n_2 = 1, n_3 = 1. \end{aligned}$$

Each of these three states have the same energy:

$$E_{\text{total}} = \frac{h^2}{8mL^2}(n_1^2 + n_2^2 + n_3^2) = \frac{6h^2}{8mL^2}$$

Note that these three states are only degenerate if  $L_1 = L_2 = L_3$ .

b. distortion

$$L_1 = L_2 = L_3$$

$$L_3 \quad L_1 = L_2$$

For  $L_1 = L_2 = L_3$ ,  $V = L_1L_2L_3 = L_1^3$ ,

$$\begin{aligned} E_{\text{total}}(L_1) &= \frac{2h^2}{8m} \frac{1^2}{L_1^2} + \frac{1^2}{L_2^2} + \frac{1^2}{L_3^2} + \frac{1h^2}{8m} \frac{1^2}{L_1^2} + \frac{1^2}{L_2^2} + \frac{2^2}{L_3^2} \\ &= \frac{2h^2}{8m} \frac{3}{L_1^2} + \frac{1h^2}{8m} \frac{6}{L_1^2} = \frac{h^2}{8m} \frac{12}{L_1^2} \end{aligned}$$

For  $L_3 \quad L_1 = L_2$ ,  $V = L_1L_2L_3 = L_1^2L_3$ ,  $L_3 = V/L_1^2$

$$\begin{aligned} E_{\text{total}}(L_1) &= \frac{2h^2}{8m} \frac{1^2}{L_1^2} + \frac{1^2}{L_2^2} + \frac{1^2}{L_3^2} + \frac{1h^2}{8m} \frac{1^2}{L_1^2} + \frac{1^2}{L_2^2} + \frac{2^2}{L_3^2} \\ &= \frac{2h^2}{8m} \frac{2}{L_1^2} + \frac{1}{L_3^2} + \frac{1h^2}{8m} \frac{2}{L_1^2} + \frac{4}{L_3^2} \\ &= \frac{2h^2}{8m} \frac{2}{L_1^2} + \frac{1}{L_3^2} + \frac{1}{L_1^2} + \frac{2}{L_3^2} \\ &= \frac{2h^2}{8m} \frac{3}{L_1^2} + \frac{3}{L_3^2} = \frac{h^2}{8m} \frac{6}{L_1^2} + \frac{6}{L_3^2} \end{aligned}$$

In comparing the total energy at constant volume of the undistorted box ( $L_1 = L_2 = L_3$ )

versus the distorted box ( $L_3 \quad L_1 = L_2$ ) it can be seen that:

$$\frac{h^2}{8m} \frac{6}{L_1^2} + \frac{6}{L_3^2} \quad \frac{h^2}{8m} \frac{12}{L_1^2} \quad \text{as long as } L_3 \quad L_1.$$

c. In order to minimize the total energy expression, take the derivative of the energy

with respect to  $L_1$  and set it equal to zero.  $\frac{E_{\text{total}}}{L_1} = 0$

$$\frac{1}{L_1} \frac{\hbar^2}{8m} \frac{6}{L_1^2} + \frac{6}{L_3^2} = 0$$

But since  $V = L_1 L_2 L_3 = L_1^2 L_3$ , then  $L_3 = V/L_1^2$ . This substitution gives:

$$\frac{1}{L_1} \frac{\hbar^2}{8m} \frac{6}{L_1^2} + \frac{6L_1^4}{V^2} = 0$$

$$\frac{\hbar^2}{8m} \frac{(-2)6}{L_1^3} + \frac{(4)6L_1^3}{V^2} = 0$$

$$-\frac{12}{L_1^3} + \frac{24L_1^3}{V^2} = 0$$

$$\frac{24L_1^3}{V^2} = \frac{12}{L_1^3}$$

$$24L_1^6 = 12V^2$$

$$L_1^6 = \frac{1}{2} V^2 = \frac{1}{2} (L_1^2 L_3)^2 = \frac{1}{2} L_1^4 L_3^2$$

$$L_1^2 = \frac{1}{2} L_3^2$$

$$L_3 = \sqrt{2} L_1$$

d. Calculate energy upon distortion:

cube:  $V = L_1^3, L_1 = L_2 = L_3 = (V)^{\frac{1}{3}}$

distorted:  $V = L_1^2 L_3 = L_1^2 \sqrt{2} L_1 = \sqrt{2} L_1^3$

$$L_3 = \sqrt{2} \frac{V}{\sqrt{2}}^{\frac{1}{3}} \quad L_1 = L_2 = \frac{V}{\sqrt{2}}^{\frac{1}{3}}$$

$$E = E_{\text{total}}(L_1 = L_2 = L_3) - E_{\text{total}}(L_3 \quad L_1 = L_2)$$

$$= \frac{\hbar^2}{8m} \frac{12}{L_1^2} - \frac{\hbar^2}{8m} \frac{6}{L_1^2} + \frac{6}{L_3^2}$$

$$= \frac{\hbar^2}{8m} \frac{12}{V^{2/3}} - \frac{6(2)^{1/3}}{V^{2/3}} + \frac{6(2)^{1/3}}{2V^{2/3}}$$

$$= \frac{\hbar^2}{8m} \frac{12 - 9(2)^{1/3}}{V^{2/3}}$$

Since  $V = 8\text{\AA}^3, V^{2/3} = 4\text{\AA}^2 = 4 \times 10^{-16} \text{ cm}^2$ , and  $\frac{\hbar^2}{8m} = 6.01 \times 10^{-27} \text{ erg cm}^2$ :

$$E = 6.01 \times 10^{-27} \text{ erg cm}^2 \frac{12 - 9(2)^{1/3}}{4 \times 10^{-16} \text{ cm}^2}$$

$$E = 6.01 \times 10^{-27} \text{ erg cm}^2 \frac{0.66}{4 \times 10^{-16} \text{ cm}^2}$$

$$E = 0.99 \times 10^{-11} \text{ erg}$$

$$E = 0.99 \times 10^{-11} \text{ erg} \frac{1 \text{ eV}}{1.6 \times 10^{-12} \text{ erg}}$$

$$E = 6.19 \text{ eV}$$

$$7. \quad a. \quad \int_{-\infty}^{\infty} \psi^*(x) \psi(x) dx = 1.$$

$$\int_{-\infty}^{\infty} A^2 e^{-2a|x|} dx = 1.$$

$$\int_{-\infty}^0 A^2 e^{2ax} dx + \int_0^{\infty} A^2 e^{-2ax} dx = 1$$

Making use of integral equation (4) this becomes:

$$A^2 \frac{1}{2a} + \frac{1}{2a} = \frac{2A^2}{2a} = 1$$

$$A^2 = a$$

$A = \pm\sqrt{a}$ , therefore A is not unique.

$$\psi(x) = Ae^{-a|x|} = \pm\sqrt{a} e^{-a|x|}$$

Since a has units of  $\text{\AA}^{-1}$ ,  $\psi(x)$  must have units of  $\text{\AA}^{-\frac{1}{2}}$ .

$$b. \quad |x| = \begin{cases} x & \text{if } x \geq 0 \\ -x & \text{if } x < 0 \end{cases}$$

$$\psi(x) = \sqrt{a} \begin{cases} e^{-ax} & \text{if } x \geq 0 \\ e^{ax} & \text{if } x < 0 \end{cases}$$

Sketching this wavefunction with respect to x (keeping constant a fixed; a = 1) gives:

$$c. \quad \frac{d\psi(x)}{dx} = \sqrt{a} \begin{cases} -ae^{-ax} & \text{if } x \geq 0 \\ ae^{ax} & \text{if } x < 0 \end{cases}$$

$$\frac{d\psi(x)}{dx} \Big|_{0^+} = -a\sqrt{a}$$

$$\frac{d\psi(x)}{dx} \Big|_{0^-} = a\sqrt{a}$$

The magnitude of discontinuity is  $a\sqrt{a} + a\sqrt{a} = 2a\sqrt{a}$  as x goes through x = 0. This also indicates that the potential V undergoes a discontinuity of magnitude  $2a\sqrt{a}$  at x = 0.

$$d. \quad \langle |x| \rangle = \int_{-\infty}^{\infty} \psi^*(x) |x| \psi(x) dx$$



$$\begin{aligned}
 & + \\
 \psi(p) &= \frac{1}{\sqrt{2\hbar}} \int_{-\infty}^{\infty} \psi(x) e^{-ipx/\hbar} dx . \\
 & + \\
 \psi(p) &= \frac{1}{\sqrt{2\hbar}} \int_0^{\infty} \sqrt{a} e^{-a|x|} e^{-ipx/\hbar} dx . \\
 & + \\
 \psi(p) &= \frac{1}{\sqrt{2\hbar}} \int_0^{\infty} \sqrt{a} e^{ax} e^{-ipx/\hbar} dx + \frac{1}{\sqrt{2\hbar}} \int_0^{\infty} \sqrt{a} e^{-ax} e^{-ipx/\hbar} dx . \\
 & = \sqrt{\frac{a}{2\hbar}} \frac{1}{a-ip/\hbar} + \frac{1}{a+ip/\hbar} \\
 & = \sqrt{\frac{a}{2\hbar}} \frac{2a}{a^2+p^2/\hbar^2}
 \end{aligned}$$

$$\begin{aligned}
 \text{g. } \frac{\psi(p=2a\hbar)^2}{\psi(p=-a\hbar)^2} &= \frac{1/(a^2+(2a\hbar)^2/\hbar^2)}{1/(a^2+(-a\hbar)^2/\hbar^2)} \\
 &= \frac{1/(a^2+4a^2)}{1/(a^2+a^2)} \\
 &= \frac{1/(5a^2)}{1/(2a^2)} \\
 &= \frac{2}{5} = 0.4 = 40\%
 \end{aligned}$$

8. a.  $\mathbf{H} = \frac{-\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right)$  (cartesian coordinates)

Finding  $\frac{\partial}{\partial x}$  and  $\frac{\partial}{\partial y}$  from the chain rule gives:

$$\frac{\partial}{\partial x} = \frac{\partial r}{\partial x} \frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x} \frac{\partial}{\partial \theta}, \quad \frac{\partial}{\partial y} = \frac{\partial r}{\partial y} \frac{\partial}{\partial r} + \frac{\partial \theta}{\partial y} \frac{\partial}{\partial \theta},$$

Evaluation of the "coefficients" gives the following:

$$\frac{\partial r}{\partial x} = \cos \theta, \quad \frac{\partial \theta}{\partial x} = -\frac{\sin \theta}{r},$$

$$\frac{\partial r}{\partial y} = \sin \theta, \quad \text{and} \quad \frac{\partial \theta}{\partial y} = \frac{\cos \theta}{r},$$

Upon substitution of these "coefficients":

$$\frac{\partial}{\partial x} = \cos \frac{\phi}{r} - \frac{\sin \phi}{r} \frac{\partial \phi}{\partial x} = -\frac{\sin \phi}{r} \frac{\partial \phi}{\partial x}; \text{ at fixed } r.$$

$$\frac{\partial}{\partial y} = \sin \frac{\phi}{r} + \frac{\cos \phi}{r} \frac{\partial \phi}{\partial y} = \frac{\cos \phi}{r} \frac{\partial \phi}{\partial y}; \text{ at fixed } r.$$

$$\begin{aligned} \frac{\partial^2}{\partial x^2} &= -\frac{\sin^2 \phi}{r^2} \frac{\partial^2 \phi}{\partial x^2} - \frac{\sin \phi \cos \phi}{r^2} \frac{\partial^2 \phi}{\partial x \partial y} \\ &= \frac{\sin^2 \phi}{r^2} \frac{\partial^2 \phi}{\partial x^2} + \frac{\sin \phi \cos \phi}{r^2} \frac{\partial^2 \phi}{\partial x \partial y}; \text{ at fixed } r. \end{aligned}$$

$$\begin{aligned} \frac{\partial^2}{\partial y^2} &= \frac{\cos^2 \phi}{r^2} \frac{\partial^2 \phi}{\partial y^2} - \frac{\cos \phi \sin \phi}{r^2} \frac{\partial^2 \phi}{\partial x \partial y} \\ &= \frac{\cos^2 \phi}{r^2} \frac{\partial^2 \phi}{\partial y^2} - \frac{\cos \phi \sin \phi}{r^2} \frac{\partial^2 \phi}{\partial x \partial y}; \text{ at fixed } r. \end{aligned}$$

$$\begin{aligned} \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} &= \frac{\sin^2 \phi}{r^2} \frac{\partial^2 \phi}{\partial x^2} + \frac{\sin \phi \cos \phi}{r^2} \frac{\partial^2 \phi}{\partial x \partial y} + \frac{\cos^2 \phi}{r^2} \frac{\partial^2 \phi}{\partial y^2} - \frac{\cos \phi \sin \phi}{r^2} \frac{\partial^2 \phi}{\partial x \partial y} \\ &= \frac{1}{r^2} \frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2}; \text{ at fixed } r. \end{aligned}$$

$$\begin{aligned} \text{So, } \mathbf{H} &= \frac{-\hbar^2}{2mr^2} \frac{\partial^2 \phi}{\partial \phi^2} \text{ (cylindrical coordinates, fixed } r) \\ &= \frac{-\hbar^2}{2I} \frac{\partial^2 \phi}{\partial \phi^2} \end{aligned}$$

The Schrödinger equation for a particle on a ring then becomes:

$$\begin{aligned} \mathbf{H} \psi &= E \psi \\ \frac{-\hbar^2}{2I} \frac{\partial^2 \psi}{\partial \phi^2} &= E \psi \\ \frac{\partial^2 \psi}{\partial \phi^2} &= \frac{-2IE}{\hbar^2} \psi \end{aligned}$$

The general solution to this equation is the now familiar expression:

$$\psi(\phi) = C_1 e^{-im\phi} + C_2 e^{im\phi}, \text{ where } m = \frac{2IE}{\hbar^2} \frac{1}{2}$$

Application of the cyclic boundary condition,  $\psi(\phi) = \psi(\phi + 2\pi)$ , results in the quantization

of the energy expression:  $E = \frac{m^2 \hbar^2}{2I}$  where  $m = 0, \pm 1, \pm 2, \pm 3, \dots$ . It can be seen that the

$\pm m$  values correspond to angular momentum of the same magnitude but opposite

directions. Normalization of the wavefunction (over the region 0 to  $2\pi$ ) corresponding to +

or -  $m$  will result in a value of  $\frac{1}{\sqrt{2\pi}}$  for the normalization constant.

$$(\ ) = \frac{1}{2} \frac{1}{2} e^{im}$$

$$\frac{(\pm 4)^2 \hbar^2}{2I}$$

$$\frac{(\pm 3)^2 \hbar^2}{2I}$$

$$\frac{(\pm 2)^2 \hbar^2}{2I}$$

$$\frac{(\pm 1)^2 \hbar^2}{2I}$$

$$\frac{(0)^2 \hbar^2}{2I}$$

$$b. \frac{\hbar^2}{2m} = 6.06 \times 10^{-28} \text{ erg cm}^2$$

$$\frac{\hbar^2}{2mr^2} = \frac{6.06 \times 10^{-28} \text{ erg cm}^2}{(1.4 \times 10^{-8} \text{ cm})^2}$$

$$= 3.09 \times 10^{-12} \text{ erg}$$

$$E = (2^2 - 1^2) 3.09 \times 10^{-12} \text{ erg} = 9.27 \times 10^{-12} \text{ erg}$$

but  $E = h \nu = hc/\lambda$  So  $\lambda = hc/E$

$$= \frac{(6.63 \times 10^{-27} \text{ erg sec})(3.00 \times 10^{10} \text{ cm sec}^{-1})}{9.27 \times 10^{-12} \text{ erg}}$$

$$= 2.14 \times 10^{-5} \text{ cm} = 2.14 \times 10^3 \text{ \AA}$$

Sources of error in this calculation include:

- i. The attractive force of the carbon nuclei is not included in the Hamiltonian.
- ii. The repulsive force of the other  $\pi$ -electrons is not included in the Hamiltonian.
- iii. Benzene is not a ring.
- iv. Electrons move in three dimensions not one.
- v. Etc.

$$9. (\ , 0) = \sqrt{\frac{4}{3}} \text{ Cos}^2 \ .$$

This wavefunction needs to be expanded in terms of the eigenfunctions of the angular

momentum operator,  $-i\hbar \frac{\partial}{\partial \phi}$  . This is most easily accomplished by an exponential

expansion of the Cos function.

$$\begin{aligned} (\ , 0) &= \sqrt{\frac{4}{3}} \frac{e^i + e^{-i}}{2} \frac{e^i + e^{-i}}{2} \\ &= \frac{1}{4} \sqrt{\frac{4}{3}} (e^{2i} + e^{-2i} + 2e^{(0)i} ) \end{aligned}$$

The wavefunction is now written in terms of the eigenfunctions of the angular momentum operator,  $-i\hbar$ , but they need to include their normalization constant,  $\frac{1}{\sqrt{2}}$ .

$$\begin{aligned} \psi(\theta, \phi) &= \frac{1}{4} \sqrt{\frac{4}{3}} \sqrt{2} \left[ \frac{1}{\sqrt{2}} e^{2i\phi} + \frac{1}{\sqrt{2}} e^{-2i\phi} + 2 \frac{1}{\sqrt{2}} e^{0i\phi} \right] \\ &= \sqrt{\frac{1}{6}} \left[ \frac{1}{\sqrt{2}} e^{2i\phi} + \frac{1}{\sqrt{2}} e^{-2i\phi} + 2 \frac{1}{\sqrt{2}} e^{0i\phi} \right] \end{aligned}$$

Once the wavefunction is written in this form (in terms of the normalized eigenfunctions of the angular momentum operator having  $m\hbar$  as eigenvalues) the probabilities for observing angular momentums of  $0\hbar$ ,  $2\hbar$ , and  $-2\hbar$  can be easily identified as the square of the coefficients of the corresponding eigenfunctions.

$$\begin{aligned} P_{2\hbar} &= \left( \sqrt{\frac{1}{6}} \right)^2 = \frac{1}{6} \\ P_{-2\hbar} &= \left( \sqrt{\frac{1}{6}} \right)^2 = \frac{1}{6} \\ P_{0\hbar} &= \left( 2\sqrt{\frac{1}{6}} \right)^2 = \frac{4}{6} \end{aligned}$$

10. a.  $\frac{-\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \psi(x,y,z) + \frac{1}{2} k(x^2 + y^2 + z^2) \psi(x,y,z) = E \psi(x,y,z)$ .

b. Let  $\psi(x,y,z) = X(x)Y(y)Z(z)$

$$\frac{-\hbar^2}{2m} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) X(x)Y(y)Z(z) + \frac{1}{2} k(x^2 + y^2 + z^2) X(x)Y(y)Z(z) = E X(x)Y(y)Z(z)$$

$$\begin{aligned} &\frac{-\hbar^2}{2m} Y(y)Z(z) \frac{\partial^2 X(x)}{\partial x^2} + \frac{-\hbar^2}{2m} X(x)Z(z) \frac{\partial^2 Y(y)}{\partial y^2} + \frac{-\hbar^2}{2m} X(x)Y(y) \frac{\partial^2 Z(z)}{\partial z^2} + \\ &\frac{1}{2} kx^2 X(x)Y(y)Z(z) + \frac{1}{2} ky^2 X(x)Y(y)Z(z) + \frac{1}{2} kz^2 X(x)Y(y)Z(z) = E X(x)Y(y)Z(z) \end{aligned}$$

Dividing by  $X(x)Y(y)Z(z)$  you obtain:

$$\frac{-\hbar^2}{2m} \frac{1}{X(x)} \frac{\partial^2 X(x)}{\partial x^2} + \frac{1}{2} kx^2 + \frac{-\hbar^2}{2m} \frac{1}{Y(y)} \frac{\partial^2 Y(y)}{\partial y^2} + \frac{1}{2} ky^2 + \frac{-\hbar^2}{2m} \frac{1}{Z(z)} \frac{\partial^2 Z(z)}{\partial z^2} + \frac{1}{2} kz^2 = E$$

Now you have each variable isolated:

$$F(x) + G(y) + H(z) = \text{constant}$$

So,

$$\frac{-\hbar^2}{2m} \frac{1}{X(x)} \frac{\partial^2 X(x)}{\partial x^2} + \frac{1}{2} kx^2 = E_x \quad \frac{-\hbar^2}{2m} \frac{\partial^2 X(x)}{\partial x^2} + \frac{1}{2} kx^2 X(x) = E_x X(x),$$

$$\frac{-\hbar^2}{2m} \frac{1}{Y(y)} \frac{d^2 Y(y)}{y^2} + \frac{1}{2} ky^2 = E_y \quad \frac{-\hbar^2}{2m} \frac{d^2 Y(y)}{y^2} + \frac{1}{2} ky^2 Y(y) = E_y Y(y),$$

$$\frac{-\hbar^2}{2m} \frac{1}{Z(z)} \frac{d^2 Z(z)}{z^2} + \frac{1}{2} kz^2 = E_z \quad \frac{-\hbar^2}{2m} \frac{d^2 Z(z)}{z^2} + \frac{1}{2} kz^2 Z(z) = E_z Z(z),$$

and  $E = E_x + E_y + E_z$ .

c. All three of these equations are one-dimensional harmonic oscillator equations and thus each have one-dimensional harmonic oscillator solutions which taken from the text are:

$$X_n(x) = \frac{1}{n!2^n} \left(\frac{1}{2} - \frac{1}{4} \frac{-x^2}{2}\right) H_n\left(\frac{1}{2} x\right),$$

$$Y_n(y) = \frac{1}{n!2^n} \left(\frac{1}{2} - \frac{1}{4} \frac{-y^2}{2}\right) H_n\left(\frac{1}{2} y\right), \text{ and}$$

$$Z_n(z) = \frac{1}{n!2^n} \left(\frac{1}{2} - \frac{1}{4} \frac{-z^2}{2}\right) H_n\left(\frac{1}{2} z\right),$$

where  $= \frac{k\mu}{\hbar^2} \frac{1}{2}$ .

d.  $E_{n_x, n_y, n_z} = E_{n_x} + E_{n_y} + E_{n_z}$

$$= \frac{\hbar^2 k}{\mu} \frac{1}{2} n_x + \frac{1}{2} + n_y + \frac{1}{2} + n_z + \frac{1}{2}$$

e. Suppose  $E = 5.5 \frac{\hbar^2 k}{\mu} \frac{1}{2}$

$$= \frac{\hbar^2 k}{\mu} \frac{1}{2} n_x + n_y + n_z + \frac{3}{2}$$

$$5.5 = n_x + n_y + n_z + \frac{3}{2}$$

So,  $n_x + n_y + n_z = 4$ . This gives rise to a degeneracy of 15. They are:

States 1-3			States 4-6			States 7-9		
$n_x$	$n_y$	$n_z$	$n_x$	$n_y$	$n_z$	$n_x$	$n_y$	$n_z$
4	0	0	3	1	0	0	3	1
0	4	0	3	0	1	1	0	3
0	0	4	1	3	0	0	1	3
States 10-12			States 13-15					
$n_x$	$n_y$	$n_z$	$n_x$	$n_y$	$n_z$			
2	2	0	2	1	1			
2	0	2	1	2	1			
0	2	2	1	1	2			

f. Suppose  $V = \frac{1}{2}kr^2$  (independent of  $\theta$  and  $\phi$ )

The solutions  $G(\theta, \phi)$  are the spherical harmonics  $Y_{l,m}(\theta, \phi)$ .

$$g. -\frac{\hbar^2}{2\mu r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial G}{\partial r} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial}{\partial \theta} \left( \sin^2 \theta \frac{\partial G}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 G}{\partial \phi^2} + \frac{k}{2}(r - r_e)^2 = E G$$

If  $G(\theta, \phi)$  is replaced by  $F(r)G(\theta, \phi)$ :

$$-\frac{\hbar^2}{2\mu r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial F(r)G(\theta, \phi)}{\partial r} \right) + \frac{F(r)}{r^2 \sin^2 \theta} \frac{\partial}{\partial \theta} \left( \sin^2 \theta \frac{\partial G(\theta, \phi)}{\partial \theta} \right) + \frac{F(r)}{r^2 \sin^2 \theta} \frac{\partial^2 G(\theta, \phi)}{\partial \phi^2} + \frac{k}{2}(r - r_e)^2 F(r)G(\theta, \phi) = E F(r)G(\theta, \phi)$$

and the angle dependence is recognized as the  $L^2$  angular momentum operator. Division by  $G(\theta, \phi)$  further reduces the equation to:

$$-\frac{\hbar^2}{2\mu r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial F(r)}{\partial r} \right) + \frac{J(J+1)\hbar^2}{2\mu r^2} F(r) + \frac{k}{2}(r - r_e)^2 F(r) = E F(r)$$

11. a.  $\frac{1}{2}mv^2 = 100 \text{ eV} \frac{1.602 \times 10^{-12} \text{ erg}}{1 \text{ eV}}$

$$v^2 = \frac{(2)1.602 \times 10^{-10} \text{ erg}}{9.109 \times 10^{-28} \text{ g}}$$

$$v = 0.593 \times 10^9 \text{ cm/sec}$$

The length of the  $N_2$  molecule is  $2\text{\AA} = 2 \times 10^{-8} \text{ cm}$ .

$$v = \frac{d}{t}$$

$$t = \frac{d}{v} = \frac{2 \times 10^{-8} \text{ cm}}{0.593 \times 10^9 \text{ cm/sec}} = 3.37 \times 10^{-17} \text{ sec}$$

b. The normalized ground state harmonic oscillator can be written (from both in the text and in exercise 11) as:

$$\psi_0 = \frac{1}{\pi^{1/4}} e^{-x^2/2}, \text{ where } x = \frac{k\mu}{\hbar^2} \frac{1}{2} \text{ and } x = r - r_e$$

Calculating constants;

$$N_2 = \frac{(2.294 \times 10^6 \text{ g sec}^{-2})(1.1624 \times 10^{-23} \text{ g})^{1/2}}{(1.0546 \times 10^{-27} \text{ erg sec})^2}$$

$$= 0.48966 \times 10^{19} \text{ cm}^{-2} = 489.66 \text{\AA}^{-2}$$

For  $N_2$ :  $\psi_0(r) = 3.53333\text{\AA}^{-1/2} e^{-(244.83\text{\AA}^{-2})(r-1.09769\text{\AA})^2}$

$$N_2^+ = \frac{(2.009 \times 10^6 \text{ g sec}^{-2})(1.1624 \times 10^{-23} \text{ g})^{\frac{1}{2}}}{(1.0546 \times 10^{-27} \text{ erg sec})^2}$$

$$= 0.45823 \times 10^{19} \text{ cm}^{-2} = 458.23 \text{ \AA}^{-2}$$

For  $N_2^+$ :  $\rho(r) = 3.47522 \text{ \AA}^{-\frac{1}{2}} e^{-(229.113 \text{ \AA}^{-2})(r-1.11642 \text{ \AA})^2}$

c.  $P(v=0) = \langle v=0(N_2^+) \quad v=0(N_2) \rangle^2$

Let  $P(v=0) = I^2$  where  $I = \text{integral}$ :

+

$$I = (3.47522 \text{ \AA}^{-\frac{1}{2}} e^{-(229.113 \text{ \AA}^{-2})(r-1.11642 \text{ \AA})^2}) \cdot$$

-

$$(3.53333 \text{ \AA}^{-\frac{1}{2}} e^{-(244.830 \text{ \AA}^{-2})(r-1.09769 \text{ \AA})^2}) dr$$

Let  $C_1 = 3.47522 \text{ \AA}^{-\frac{1}{2}}$ ,  $C_2 = 3.53333 \text{ \AA}^{-\frac{1}{2}}$ ,  
 $A_1 = 229.113 \text{ \AA}^{-2}$ ,  $A_2 = 244.830 \text{ \AA}^{-2}$ ,  
 $r_1 = 1.11642 \text{ \AA}$ ,  $r_2 = 1.09769 \text{ \AA}$ ,

+

$$I = C_1 C_2 e^{-A_1(r-r_1)^2} e^{-A_2(r-r_2)^2} dr.$$

-

Focusing on the exponential:

$$-A_1(r-r_1)^2 - A_2(r-r_2)^2 = -A_1(r^2 - 2r_1r + r_1^2) - A_2(r^2 - 2r_2r + r_2^2)$$

$$= -(A_1 + A_2)r^2 + (2A_1r_1 + 2A_2r_2)r - A_1r_1^2 - A_2r_2^2$$

Let  $A = A_1 + A_2$ ,  
 $B = 2A_1r_1 + 2A_2r_2$ ,  
 $C = C_1C_2$ , and  
 $D = A_1r_1^2 + A_2r_2^2$ .

+

$$I = C e^{-Ar^2 + Br - D} dr$$

-

+

$$= C e^{-A(r-r_0)^2 + D'} dr$$

-

where  $-A(r-r_0)^2 + D' = -Ar^2 + Br - D$

such that,  $-A(r^2 - 2rr_0 + r_0^2) + D' = -Ar^2 + Br - D$   
 $2Ar_0 = B$   
 $-Ar_0^2 + D' = -D$

and,

$$r_0 = \frac{B}{2A}$$

$$D' = Ar_0^2 - D = A \frac{B^2}{4A^2} - D = \frac{B^2}{4A} - D .$$

+

$$I = C \int e^{-A(r-r_0)^2 + D'} dr$$

-

+

$$= Ce^{D'} \int e^{-Ay^2} dy$$

-

$$= Ce^{D'} \sqrt{\frac{\pi}{A}}$$

Now back substituting all of these constants:

$$I = C_1 C_2 \sqrt{\frac{\pi}{A_1 + A_2}} \exp \frac{(2A_1 r_1 + 2A_2 r_2)^2}{4(A_1 + A_2)} - A_1 r_1^2 - A_2 r_2^2$$

$$I = (3.47522)(3.53333) \sqrt{\frac{\pi}{(229.113) + (244.830)}} \\ \cdot \exp \frac{(2(229.113)(1.11642) + 2(244.830)(1.09769))^2}{4((229.113) + (244.830))} \\ \cdot \exp( - (229.113)(1.11642)^2 - (244.830)(1.09769)^2) \\ I = 0.959 \\ P(v=0) = I^2 = 0.92$$

12. a.  $E = \frac{\hbar^2 k^2}{2\mu} + \frac{1}{2}$

$$E = E_{+1} - E$$

$$= \frac{\hbar^2 k^2}{2\mu} + 1 + \frac{1}{2} - \left( - \frac{1}{2} \right) = \frac{\hbar^2 k^2}{\mu}$$

$$= \frac{(1.0546 \times 10^{-27} \text{ erg sec})^2 (1.87 \times 10^6 \text{ g sec}^{-2})^{\frac{1}{2}}}{6.857 \text{ g} / 6.02 \times 10^{23}}$$

$$= 4.27 \times 10^{-13} \text{ erg}$$

$$E = \frac{hc}{\lambda}$$

$$= \frac{hc}{E} = \frac{(6.626 \times 10^{-27} \text{ erg sec})(3.00 \times 10^{10} \text{ cm sec}^{-1})}{4.27 \times 10^{-13} \text{ erg}}$$

$$= 4.66 \times 10^{-4} \text{ cm}$$

$$\frac{1}{\lambda} = 2150 \text{ cm}^{-1}$$

$$\text{b. } \psi_0 = -\frac{1}{\sqrt{2}} e^{-x^2/2}$$

$$\begin{aligned} \langle x \rangle &= \langle \psi_0 | x | \psi_0 \rangle \\ &= \int_{-\infty}^{\infty} \psi_0^* x \psi_0 dx \\ &= \int_{-\infty}^{\infty} -\frac{1}{2} x e^{-x^2} dx \\ &= \frac{1}{2} \int_{-\infty}^{\infty} e^{-x^2} d(-x^2) \\ &= \frac{-1}{2} \int_{-\infty}^{\infty} e^{-x^2} dx = 0 \end{aligned}$$

$$\begin{aligned} \langle x^2 \rangle &= \langle \psi_0 | x^2 | \psi_0 \rangle \\ &= \int_{-\infty}^{\infty} \psi_0^* x^2 \psi_0 dx \\ &= \int_{-\infty}^{\infty} -\frac{1}{2} x^2 e^{-x^2} dx \\ &= 2 \int_{-\infty}^{\infty} x^2 e^{-x^2} dx \end{aligned}$$

Using integral equation (4) this becomes:

$$\begin{aligned} &= 2 \int_{-\infty}^{\infty} x^2 e^{-x^2} dx \\ &= \frac{1}{2} \end{aligned}$$

$$\Delta x = (\langle x^2 \rangle - \langle x \rangle^2)^{1/2} = \frac{1}{2}$$

$$\begin{aligned}
&= \frac{\hbar}{2\sqrt{k\mu}}^{\frac{1}{2}} \\
&= \frac{(1.0546 \times 10^{-27} \text{ erg sec})^2}{4(1.87 \times 10^6 \text{ g sec}^{-2})(6.857 \text{ g} / 6.02 \times 10^{23})}^{\frac{1}{4}} \\
&= 3.38 \times 10^{-10} \text{ cm} = 0.0338 \text{ \AA} \\
\text{c. } x &= \frac{\hbar}{2\sqrt{k\mu}}^{\frac{1}{2}}
\end{aligned}$$

The smaller  $k$  and  $\mu$  become, the larger the uncertainty in the internuclear distance becomes. Helium has a small  $\mu$  and small force between atoms. This results in a very large  $x$ . This implies that it is extremely difficult for He atoms to "vibrate" with small displacement as a solid even as absolute zero is approached.

$$\begin{aligned}
13. \quad \text{a. } W &= Z_e^2 - 2ZZ_e + \frac{5}{8} Z_e \frac{e^2}{a_0} \\
\frac{dW}{dZ_e} &= 2Z_e - 2Z + \frac{5}{8} \frac{e^2}{a_0} = 0
\end{aligned}$$

$$2Z_e - 2Z + \frac{5}{8} = 0$$

$$2Z_e = 2Z - \frac{5}{8}$$

$$Z_e = Z - \frac{5}{16} = Z - 0.3125 \quad (\text{Note this is the shielding factor of one } 1s$$

electron to the other).

$$W = Z_e Z_e - 2Z + \frac{5}{8} \frac{e^2}{a_0}$$

$$W = \left(Z - \frac{5}{16}\right) \left(Z - \frac{5}{16}\right) - 2Z + \frac{5}{8} \frac{e^2}{a_0}$$

$$W = Z^2 - \frac{5}{8} Z - Z + \frac{5}{16} \frac{e^2}{a_0}$$

$$W = -Z - \frac{5}{16} Z + \frac{5}{16} \frac{e^2}{a_0} = -Z - \frac{5}{16} Z + \frac{5}{16} \frac{e^2}{a_0}$$

$$= -(Z - 0.3125)^2 (27.21) \text{ eV}$$

b. Using the above result for  $W$  and the percent error as calculated below we obtain the following:

$$\% \text{ error} = \frac{(\text{Experimental} - \text{Theoretical})}{\text{Experimental}} * 100$$

Z	Atom	Experimental	Calculated	% Error
Z = 1	H-	-14.35 eV	-12.86 eV	10.38%
Z = 2	He	-78.98 eV	-77.46 eV	1.92%
Z = 3	Li+	-198.02 eV	-196.46 eV	0.79%
Z = 4	Be <sup>2+</sup>	-371.5 eV	-369.86 eV	0.44%

Z = 5	B <sup>+3</sup>	-599.3 eV	-597.66 eV	0.27%
Z = 6	C <sup>+4</sup>	-881.6 eV	-879.86 eV	0.19%
Z = 7	N <sup>+5</sup>	-1218.3 eV	-1216.48 eV	0.15%
Z = 8	O <sup>+6</sup>	-1609.5 eV	-1607.46 eV	0.13%

The ignored electron correlation effects are essentially constant over the range of Z, but this correlation effect is a larger percentage error at small Z. At large Z the dominant interaction is electron attraction to the nucleus completely overwhelming the ignored electron correlation and hence reducing the overall percent error.

c. Since -12.86 eV (H<sup>-</sup>) is greater than -13.6 eV (H + e) this simple variational calculation erroneously predicts H<sup>-</sup> to be unstable.

14. a.  $W = \int \psi^* H \psi dx$

$$W = \int \frac{2b}{2} e^{-bx^2} \left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + a|x| \right] e^{-bx^2} dx$$

$$\begin{aligned} \frac{d^2}{dx^2} e^{-bx^2} &= \frac{d}{dx} (-2bx) e^{-bx^2} \\ &= (-2bx) (-2bx) e^{-bx^2} + e^{-bx^2} (-2b) \\ &= 4b^2x^2 e^{-bx^2} - 2b e^{-bx^2} \end{aligned}$$

Making this substitution results in the following three integrals:

$$W = \frac{2b}{2} \int e^{-bx^2} \left[ -\frac{\hbar^2}{2m} 4b^2x^2 e^{-bx^2} + \right. dx +$$

$$\left. \frac{2b}{2} \int e^{-bx^2} (-2b) e^{-bx^2} dx + \right.$$

$$\left. \frac{2b}{2} \int e^{-bx^2} a|x| e^{-bx^2} dx \right]$$

$$= \frac{2b}{2} \frac{1}{2} \frac{-2b^2 \hbar^2}{m} \int x^2 e^{-2bx^2} dx + \frac{2b}{2} \frac{1}{2} \frac{b \hbar^2}{m} \int e^{-2bx^2} dx +$$

$$\frac{2b}{2} \frac{1}{2} a \int |x| e^{-2bx^2} dx$$

Using integral equations (1), (2), and (3) this becomes:

$$= \frac{2b}{2} \frac{1}{2} \frac{-2b^2 \hbar^2}{m} \frac{1}{2 \cdot 2 \cdot 2b} \sqrt{\frac{1}{2b}} + \frac{2b}{2} \frac{1}{2} \frac{b \hbar^2}{m} \frac{1}{2} \sqrt{\frac{1}{2b}} +$$

$$\frac{2b}{2} \frac{1}{2} a \frac{0!}{2b}$$

$$= -\frac{b \hbar^2}{m} \frac{1}{2} + \frac{b \hbar^2}{m} + \frac{2b}{2} \frac{1}{2} \frac{a}{2b}$$

$$W = \frac{b \hbar^2}{2m} + a \frac{1}{2b} \frac{1}{2}$$

b. Optimize b by evaluating  $\frac{dW}{db} = 0$

$$\frac{dW}{db} = \frac{d}{db} \frac{b \hbar^2}{2m} + a \frac{1}{2b} \frac{1}{2}$$

$$= \frac{\hbar^2}{2m} - \frac{a}{2} \frac{1}{2} b^{-\frac{3}{2}}$$

$$\text{So, } \frac{a}{2} \frac{1}{2} b^{-\frac{3}{2}} = \frac{\hbar^2}{2m} \text{ or, } b^{-\frac{3}{2}} = \frac{\hbar^2}{2m} \frac{2}{a} \frac{1}{2} = \frac{\hbar^2}{ma} \sqrt{2},$$

and,  $b = \frac{ma}{\sqrt{2} \hbar^2} \frac{2}{3}$ . Substituting this value of b into the expression for W gives:

$$W = \frac{\hbar^2}{2m} \frac{ma}{\sqrt{2} \hbar^2} \frac{2}{3} + a \frac{1}{2} \frac{1}{2} \frac{ma}{\sqrt{2} \hbar^2} \frac{1}{3}$$

$$= \frac{\hbar^2}{2m} \frac{ma}{\sqrt{2} \hbar^2} \frac{2}{3} + a \frac{1}{2} \frac{1}{2} \frac{ma}{\sqrt{2} \hbar^2} \frac{1}{3}$$

$$= 2 \frac{4}{3} \frac{1}{3 \hbar^3} \frac{2}{a^3} \frac{2}{m} \frac{1}{3} + 2 \frac{1}{3} \frac{1}{3 \hbar^3} \frac{2}{a^3} \frac{2}{m} \frac{1}{3}$$

$$= 2^{\frac{4}{3}} \frac{1}{3} + 2^{\frac{1}{3}} \frac{1}{3} \frac{2}{\hbar^3} \frac{2}{a^3} \frac{1}{m^3} = \frac{3}{2} (2)^{\frac{1}{3}} \frac{2}{\hbar^3} \frac{2}{a^3} \frac{1}{m^3}$$

$$= 0.812889106 \hbar^3 a^3 m^{\frac{1}{3}} \quad \text{in error} = 0.5284\% \quad \text{!!!!}$$

15. a.  $\mathbf{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} kx^2$

$$= \sqrt{\frac{15}{16}} a^{\frac{5}{2}} (a^2 - x^2) \quad \text{for } -a < x < a$$

$$= 0 \quad \text{for } |x| > a$$

+  
\* $\mathbf{H} dx$   
-  
+a

$$= \int_{-a}^{+a} \sqrt{\frac{15}{16}} a^{\frac{5}{2}} (a^2 - x^2) \left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} kx^2 \right] \sqrt{\frac{15}{16}} a^{\frac{5}{2}} (a^2 - x^2) dx$$

$$= \frac{15}{16} a^{-5} \int_{-a}^{+a} (a^2 - x^2) \left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{1}{2} kx^2 \right] (a^2 - x^2) dx$$

$$= \frac{15}{16} a^{-5} \int_{-a}^{+a} (a^2 - x^2) \left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} (a^2 - x^2) \right] dx$$

$$+ \frac{15}{16} a^{-5} \int_{-a}^{+a} (a^2 - x^2) \frac{1}{2} kx^2 (a^2 - x^2) dx$$

$$= \frac{15}{16} a^{-5} \int_{-a}^{+a} (a^2 - x^2) \left[ -\frac{\hbar^2}{2m} (-2) \right] dx$$

$$+ \frac{15}{32} a^{-5} \int_{-a}^{+a} (kx^2)(a^4 - 2a^2x^2 + x^4) dx$$

$$= \frac{15\hbar^2}{16m} a^{-5} \int_{-a}^{+a} (a^2 - x^2) dx + \frac{15}{32} a^{-5} \int_{-a}^{+a} a^4 kx^2 - 2a^2 kx^4 + kx^6 dx$$

$$\begin{aligned}
&= \frac{15\hbar^2}{16m} a^{-5} a^2 x^3 - \frac{1}{3} x^3 a^{-a} \\
&\quad + \frac{15}{32} a^{-5} \frac{a^4 k}{3} x^3 - \frac{2a^2 k}{5} x^5 - \frac{k}{7} x^7 \\
&= \frac{15\hbar^2}{16m} a^{-5} 2a^3 - \frac{2}{3} a^3 + \frac{15}{32} a^{-5} \frac{2a^7 k}{3} - \frac{4a^7 k}{5} + \frac{2k}{7} a^7 \\
&= \frac{15}{16} a^{-5} \frac{4\hbar^2}{3m} a^3 + \frac{a^7 k}{3} - \frac{2a^7 k}{5} + \frac{k}{7} a^7 \\
&= \frac{15}{16} a^{-5} \frac{4\hbar^2}{3m} a^3 + \frac{k}{3} - \frac{2k}{5} + \frac{k}{7} a^7 \\
&= \frac{15}{16} a^{-5} \frac{4\hbar^2}{3m} a^3 + \frac{35k}{105} - \frac{42k}{105} + \frac{15k}{105} a^7 \\
&= \frac{15}{16} a^{-5} \frac{4\hbar^2}{3m} a^3 + \frac{8k}{105} a^7 = \frac{5\hbar^2}{4ma^2} + \frac{ka^2}{14}
\end{aligned}$$

b. Substituting  $a = b \frac{\hbar^2}{km}^{\frac{1}{4}}$  into the above expression for E we obtain:

$$\begin{aligned}
E &= \frac{5\hbar^2}{4b^2 m} \frac{km}{\hbar^2}^{\frac{1}{2}} + \frac{kb^2}{14} \frac{\hbar^2}{km}^{\frac{1}{2}} \\
&= \hbar^{\frac{1}{2}} k^{\frac{1}{2}} m^{-\frac{1}{2}} \frac{5}{4} b^{-2} + \frac{1}{14} b^2
\end{aligned}$$

Plotting this expression for the energy with respect to b having values of 0.2, 0.4, 0.6, 0.8, 1.0, 1.5, 2.0, 2.5, 3.0, 4.0, and 5.0 gives:

$$\begin{aligned}
\text{c. } E &= \frac{5\hbar^2}{4ma^2} + \frac{ka^2}{14} \\
\frac{dE}{da} &= -\frac{10\hbar^2}{4ma^3} + \frac{2ka}{14} = -\frac{5\hbar^2}{2ma^3} + \frac{ka}{7} = 0 \\
\frac{5\hbar^2}{2ma^3} &= \frac{ka}{7} \text{ and } 35\hbar^2 = 2mka^4
\end{aligned}$$

$$\text{So, } a^4 = \frac{35\hbar^2}{2mk}, \text{ or } a = \frac{35\hbar^2}{2mk}^{\frac{1}{4}}$$

$$\text{Therefore } E_{\text{best}} = \sqrt{\frac{15}{16}} \frac{35\hbar^2}{2mk}^{\frac{5}{8}} \frac{35\hbar^2}{2mk}^{\frac{1}{2}} - x^2,$$

$$\text{and } E_{\text{best}} = \frac{5\hbar^2}{4m} \frac{2mk}{35\hbar^2}^{\frac{1}{2}} + \frac{k}{14} \frac{35\hbar^2}{2mk}^{\frac{1}{2}} = \hbar^{\frac{1}{2}} k^{\frac{1}{2}} m^{-\frac{1}{2}} \frac{5}{14}^{\frac{1}{2}}.$$

$$\begin{aligned}
 \text{d. } \frac{E_{\text{best}} - E_{\text{true}}}{E_{\text{true}}} &= \frac{\hbar^2 k^2 m^{-1} \frac{5}{14} - 0.5}{\hbar^2 k^2 m^{-1} 0.5} \\
 &= \frac{\frac{5}{14} - 0.5}{0.5} = \frac{0.0976}{0.5} = 0.1952 = 19.52\%
 \end{aligned}$$

$$\begin{aligned}
 16. \quad \text{a.} \quad E^2 &= m^2 c^4 + p^2 c^2 \\
 &= m^2 c^4 \left( 1 + \frac{p^2}{m^2 c^2} \right) \\
 E &= mc^2 \sqrt{1 + \frac{p^2}{m^2 c^2}} \\
 &= mc^2 \left( 1 + \frac{p^2}{2m^2 c^2} - \frac{p^4}{8m^4 c^4} + \dots \right) \\
 &= mc^2 + \frac{p^2}{2m} - \frac{p^4}{8m^3 c^2} + \dots \\
 \text{Let } V &= -\frac{p^4}{8m^3 c^2}
 \end{aligned}$$

$$\begin{aligned}
 \text{b.} \quad E_{1s}^{(1)} &= \langle (r)_{1s}^* V (r)_{1s} \rangle \\
 E_{1s}^{(1)} &= \int_0^{\infty} \int_0^{\pi} \int_0^{2\pi} \frac{Z^3}{a^3} \frac{1}{2} e^{-\frac{Zr}{a}} \left( -\frac{p^4}{8m^3 c^2} \right) \frac{Z^3}{a^3} \frac{1}{2} e^{-\frac{Zr}{a}} d\Omega dr
 \end{aligned}$$

Substituting  $p = -i\hbar \nabla$ ,  $d\Omega = r^2 dr \sin \theta d\theta d\phi$ , and pulling out constants gives:

$$E_{1s}^{(1)} = -\frac{\hbar^4}{8m^3 c^2} \frac{Z^6}{a^6} \frac{1}{4} \int_0^{\infty} e^{-\frac{2Zr}{a}} r^2 dr \int_0^{\pi} \sin \theta d\theta \int_0^{2\pi} d\phi$$

The integrals over the angles are easy,  $\int_0^{2\pi} d\phi = 2\pi$  and  $\int_0^{\pi} \sin \theta d\theta = 2$ .

The work remaining is in evaluating the integral over  $r$ . Substituting

$u = \frac{1}{r^2} - \frac{r^2}{r}$  we obtain:

$$\int_0^{\infty} e^{-\frac{2Zr}{a}} r^2 dr = \int_0^{\infty} \frac{1}{r^2} - \frac{r^2}{r} e^{-\frac{2Zr}{a}} dr = \int_0^{\infty} \frac{1}{r^2} - \frac{Z}{a} e^{-\frac{2Zr}{a}} dr = \frac{-Z}{a} \int_0^{\infty} \frac{1}{r^2} - \frac{Z}{a} e^{-\frac{2Zr}{a}} dr$$

$$\begin{aligned}
&= \frac{-Z}{a} \frac{1}{r^2} - \frac{Z}{a} \frac{1}{r} r^2 e^{-\frac{Zr}{a}} = \frac{-Z}{a} \frac{1}{r^2} r^2 - \frac{Z}{a} e^{-\frac{Zr}{a}} + e^{-\frac{Zr}{a}} 2r \\
&= \frac{-Z}{a} - \frac{Z}{a} + \frac{2}{r} e^{-\frac{Zr}{a}} = \frac{Z}{a} - \frac{2Z}{ar} e^{-\frac{Zr}{a}}.
\end{aligned}$$

The integral over r then becomes:

$$\begin{aligned}
\int_0^{\infty} e^{-\frac{Zr}{a}} \left( \frac{Z}{a} - \frac{2Z}{ar} \right) r^2 dr &= \int_0^{\infty} \left( \frac{Z}{a} r^2 - \frac{2Z}{a} r \right) e^{-\frac{Zr}{a}} dr \\
&= \int_0^{\infty} \left( \frac{Z}{a} r^2 - \frac{4}{r} \frac{Z}{a} r^3 + \frac{4}{r^2} \frac{Z}{a} r^2 \right) e^{-\frac{Zr}{a}} dr \\
&= \int_0^{\infty} \left( \frac{Z}{a} r^2 - 4 \frac{Z}{a} r + 4 \frac{Z}{a} \right) e^{-\frac{Zr}{a}} dr
\end{aligned}$$

Using integral equation (4) these integrals can easily be evaluated:

$$\begin{aligned}
&= 2 \frac{Z}{a} \frac{a^3}{2Z} - 4 \frac{Z}{a} \frac{a^2}{2Z} + 4 \frac{Z}{a} \frac{a}{2Z} \\
&= \frac{Z}{4a} - \frac{Z}{a} + 2 \frac{Z}{a} = \frac{5Z}{4a}
\end{aligned}$$

$$\text{So, } E_{1s}^{(1)} = -\frac{\hbar^4}{8m^3c^2} \frac{Z^3}{a} \frac{1}{4} \frac{5Z}{4a} = -\frac{5\hbar^4 Z^4}{8m^3c^2 a^4}$$

Substituting  $a_0 = \frac{\hbar^2}{me^2}$  gives:

$$E_{1s}^{(1)} = -\frac{5\hbar^4 Z^4 m^4 e^8}{8m^3 c^2 \hbar^8} = -\frac{5Z^4 m e^8}{8c^2 \hbar^4}$$

Notice that  $E_{1s} = -\frac{Z^2 m e^4}{2\hbar^2}$ , so,  $E_{1s}^2 = -\frac{Z^4 m^2 e^8}{4\hbar^4}$  and that  $E_{1s}^{(1)} = \frac{5m}{2} E_{1s}^2$

$$\text{c. } \frac{E_{1s}^{(1)}}{E_{1s}} = \frac{-\frac{5Z^4 m e^8}{8c^2 \hbar^4}}{-\frac{Z^2 m e^4}{2\hbar^2}} = 10\% = 0.1$$

$$\frac{5Z^2 e^4}{4c^2 \hbar^2} = 0.1, \text{ so, } Z^2 = \frac{(0.1)4c^2 \hbar^2}{5e^4}$$

$$Z^2 = \frac{(0.1)(4)(3.00 \times 10^{10})^2 (1.05 \times 10^{-27})^2}{(5)(4.8 \times 10^{-10})^4}$$

$$Z^2 = 1.50 \times 10^3$$

$$Z = 39$$

$$17. \quad a. \quad \mathbf{H}_0 \quad \begin{matrix} (0) \\ l m \end{matrix} = \frac{\mathbf{L}^2}{2m_e r_0^2} \quad \begin{matrix} (0) \\ l m \end{matrix} = \frac{\mathbf{L}^2}{2m_e r_0^2} Y_{l,m}(\theta, \phi)$$

$$= \frac{1}{2m_e r_0^2} \hbar^2 l(l+1) Y_{l,m}(\theta, \phi)$$

$$E_{lm}^{(0)} = \frac{\hbar^2}{2m_e r_0^2} l(l+1)$$

$$b. \quad V = -e z = -e r_0 \cos \theta$$

$$E_{00}^{(1)} = \langle Y_{00} | V | Y_{00} \rangle = \langle Y_{00} | -e r_0 \cos \theta | Y_{00} \rangle$$

$$= -e r_0 \langle Y_{00} | \cos \theta | Y_{00} \rangle$$

Using the given identity this becomes:

$$E_{00}^{(1)} = -e r_0 \langle Y_{00} | Y_{10} \rangle \sqrt{\frac{(0+0+1)(0-0+1)}{(2(0)+1)(2(0)+3)}} +$$

$$-e r_0 \langle Y_{00} | Y_{-10} \rangle \sqrt{\frac{(0+0)(0-0)}{(2(0)+1)(2(0)-1)}}$$

The spherical harmonics are orthonormal, thus  $\langle Y_{00} | Y_{10} \rangle = \langle Y_{00} | Y_{-10} \rangle = 0$ , and  $E_{00}^{(1)} = 0$ .

$$E_{lm}^{(2)} = \frac{\langle Y_{lm} | V | Y_{00} \rangle^2}{E_{00}^{(0)} - E_{lm}^{(0)}}$$

$$\langle Y_{lm} | V | Y_{00} \rangle = -e r_0 \langle Y_{lm} | \cos \theta | Y_{00} \rangle$$

Using the given identity this becomes:

$$\langle Y_{lm} | V | Y_{00} \rangle = -e r_0 \langle Y_{lm} | Y_{10} \rangle \sqrt{\frac{(0+0+1)(0-0+1)}{(2(0)+1)(2(0)+3)}} +$$

$$-e r_0 \langle Y_{lm} | Y_{-10} \rangle \sqrt{\frac{(0+0)(0-0)}{(2(0)+1)(2(0)-1)}}$$

$$\langle Y_{lm} | V | Y_{00} \rangle = -\frac{e r_0}{\sqrt{3}} \langle Y_{lm} | Y_{10} \rangle$$

This indicates that the only term contributing to the sum in the expression for  $E_{00}^{(2)}$  is when  $l m = 1 0$  ( $l=1$ , and  $m=0$ ), otherwise

$\langle Y_{lm} | V | Y_{00} \rangle$  vanishes (from orthonormality). In quantum chemistry when using

orthonormal functions it is typical to write the term  $\langle Y_{lm} | Y_{10} \rangle$  as a delta function, for

example  $\delta_{lm,10}$ , which only has values of 1 or 0;  $\delta_{ij} = 1$  when  $i = j$  and 0 when  $i \neq j$ . This delta function when inserted into the sum then eliminates the sum by "picking out" the non-zero component. For example,

$$\langle Y_{lm}|V|Y_{00}\rangle = -\frac{e r_0}{\sqrt{3}} \quad l=1, m=0, \text{ so}$$

$$E_{00}^{(2)} = \frac{e^2 r_0^2}{3} \frac{l(l+1)}{E_{00}^{(0)} - E_{lm}^{(0)}} = \frac{e^2 r_0^2}{3} \frac{1}{E_{00}^{(0)} - E_{10}^{(0)}}$$

$$E_{00}^{(0)} = \frac{\hbar^2}{2m_e r_0^2} 0(0+1) = 0 \quad \text{and} \quad E_{10}^{(0)} = \frac{\hbar^2}{2m_e r_0^2} 1(1+1) = \frac{\hbar^2}{m_e r_0^2}$$

Inserting these energy expressions above yields:

$$E_{00}^{(2)} = -\frac{e^2 r_0^2}{3} \frac{m_e r_0^2}{\hbar^2} = -\frac{m_e e^2 r_0^4}{3\hbar^2}$$

$$\begin{aligned} \text{c.} \quad E_{00} &= E_{00}^{(0)} + E_{00}^{(1)} + E_{00}^{(2)} + \dots \\ &= 0 + 0 - \frac{m_e e^2 r_0^4}{3\hbar^2} \\ &= -\frac{m_e e^2 r_0^4}{3\hbar^2} \\ &= -\frac{2E}{2} = -\frac{2}{2} \frac{m_e e^2 r_0^4}{3\hbar^2} \\ &= \frac{2m_e e^2 r_0^4}{3\hbar^2} \end{aligned}$$

$$\begin{aligned} \text{d.} \quad &= \frac{2(9.1095 \times 10^{-28} \text{g})(4.80324 \times 10^{-10} \text{g}^2 \text{cm}^2 \text{s}^{-1})^2 r_0^4}{3(1.05459 \times 10^{-27} \text{g cm}^2 \text{s}^{-1})^2} \\ &= r_0^4 12598 \times 10^6 \text{cm}^{-1} = r_0^4 1.2598 \text{\AA}^{-1} \\ H &= 0.0987 \text{\AA}^3 \\ C_s &= 57.57 \text{\AA}^3 \end{aligned}$$

$$18. \quad \text{a.} \quad V = e^{-x} - \frac{L}{2}, \quad \psi_n^{(0)} = \frac{2}{L} \frac{1}{2} \sin \frac{n x}{L}, \quad \text{and}$$

$$E_n^{(0)} = \frac{\hbar^2 k_n^2}{2mL^2}$$

$$\begin{aligned} E_{n=1}^{(1)} &= \langle \psi_{n=1}^{(0)} | V | \psi_{n=1}^{(0)} \rangle = \left\langle \frac{2}{L} \sin \frac{x}{L} \left| e^{-x} - \frac{L}{2} \right| \frac{2}{L} \sin \frac{x}{L} \right\rangle \\ &= \frac{2}{L} \int_0^L \sin^2 \frac{x}{L} e^{-x} - \frac{L}{2} dx \end{aligned}$$

$$= \frac{2e}{L} \int_0^L \sin^2 \frac{x}{L} x dx - \frac{2e}{L} \frac{L}{2} \int_0^L \sin^2 \frac{x}{L} dx$$

The first integral can be evaluated using integral equation (18) with  $a = \frac{1}{L}$  :

$$\int_0^L \sin^2(ax) x dx = \frac{x^2}{4} - \frac{x \sin(2ax)}{4a} - \frac{\cos(2ax)}{8a^2} \Big|_0^L = \frac{L^2}{4}$$

The second integral can be evaluated using integral equation (10) with  $a = \frac{1}{L}$  and  $d = \frac{1}{L}$  dx :

$$\int_0^L \sin^2 \frac{x}{L} dx = \frac{L}{2} \int_0^L \sin^2 d$$

$$\int_0^L \sin^2 d = -\frac{1}{4} \sin(2d) + \frac{d}{2} \Big|_0^L = \frac{L}{2}$$

Making all of these appropriate substitutions we obtain:

$$E_{n=1}^{(1)} = \frac{2e}{L} \frac{L^2}{4} - \frac{L}{2} \frac{L}{2} = 0$$

$$E_{n=1}^{(1)} = \frac{\left\langle \int_{n=2}^{(0)} e^{-x} - \frac{L}{2} \int_{n=1}^{(0)} \right\rangle_{n=2}^{(0)}}{E_{n=1}^{(0)} - E_{n=2}^{(0)}}$$

$$E_{n=1}^{(1)} = \frac{\int_0^L \left( \frac{2}{L} \sin \frac{2x}{L} e^{-x} - \frac{L}{2} \sin \frac{x}{L} \right) dx}{\frac{\hbar^2}{2mL^2} (1^2 - 2^2)} = \frac{2}{L} \frac{1}{2} \sin \frac{2x}{L}$$

The two integrals in the numerator need to be evaluated:

$$\int_0^L x \sin \frac{2x}{L} \sin \frac{x}{L} dx, \text{ and } \int_0^L \sin \frac{2x}{L} \sin \frac{x}{L} dx.$$

Using trigonometric identity (20), the integral  $\int x \cos(ax) dx = \frac{1}{a^2} \cos(ax) + \frac{x}{a} \sin(ax)$ , and

the integral  $\int \cos(ax) dx = \frac{1}{a} \sin(ax)$ , we obtain the following:

$$\int_0^L \sin \frac{2x}{L} \sin \frac{x}{L} dx = \frac{1}{2} \int_0^L \cos \frac{x}{L} dx - \frac{1}{2} \int_0^L \cos \frac{3x}{L} dx$$

$$= \frac{1}{2} \left[ \frac{L}{1} \sin \frac{x}{L} \right]_0^L - \frac{1}{2} \left[ \frac{L}{3} \sin \frac{3x}{L} \right]_0^L = 0$$

$$\int_0^L x \sin \frac{2x}{L} \sin \frac{x}{L} dx = \frac{1}{2} \int_0^L x \cos \frac{x}{L} dx - \frac{1}{2} \int_0^L x \cos \frac{3x}{L} dx$$

$$= \frac{1}{2} \left[ \frac{L^2}{2} \cos \frac{x}{L} + \frac{Lx}{1} \sin \frac{x}{L} \right]_0^L - \frac{1}{2} \left[ \frac{L^2}{9} \cos \frac{3x}{L} + \frac{Lx}{3} \sin \frac{3x}{L} \right]_0^L$$

$$= \frac{L^2}{2} (\cos(\pi) - \cos(0)) + \frac{L^2}{2} \sin(\pi) - 0$$

$$- \frac{L^2}{18} (\cos(3\pi) - \cos(0)) - \frac{L^2}{6} \sin(3\pi) + 0$$

$$= \frac{-2L^2}{2} - \frac{-2L^2}{18} = \frac{L^2}{9} - \frac{L^2}{2} = -\frac{8L^2}{9}$$

Making all of these appropriate substitutions we obtain:

$$a_{n=1}^{(1)} = \frac{\frac{2}{L} \left( e^{-\frac{8L^2}{9}} - \frac{L}{2} \right)}{\frac{-3\hbar^2}{2mL^2}} \frac{2}{L} \frac{1}{2} \sin \frac{2x}{L}$$

$$a_{n=1}^{(1)} = \frac{32mL^3 e^{-\frac{8L^2}{9}}}{27\hbar^2} \frac{2}{L} \frac{1}{2} \sin \frac{2x}{L}$$

Crudely sketching  $a_{n=1}^{(0)} + a_{n=1}^{(1)}$  gives:

Note that the electron density has been pulled to the left side of the box by the external field!

$$b. \mu_{\text{induced}} = -e \int_0^L \left( a_{n=1}^{(0)} + a_{n=1}^{(1)} \right) x - \frac{L}{2} dx, \text{ where, } a_{n=1}^{(0)} = \frac{2}{L} \sin \frac{2x}{L}$$

$$\mu_{\text{induced}} = -e \int_0^L \left( \frac{2}{L} \sin \frac{2x}{L} + \frac{32mL^3 e^{-\frac{8L^2}{9}}}{27\hbar^2} \frac{2}{L} \frac{1}{2} \sin \frac{2x}{L} \right) x - \frac{L}{2} dx$$

$$= -e \int_0^L \left( \frac{2}{L} \sin \frac{2x}{L} + \frac{32mL^3 e^{-\frac{8L^2}{9}}}{27\hbar^2} \frac{2}{L} \frac{1}{2} \sin \frac{2x}{L} \right) x - \frac{L}{2} dx$$

$$-e \int_0^L \psi_1^{(1)*} \left(x - \frac{L}{2}\right) \psi_1^{(0)} dx - e \int_0^L \psi_1^{(1)*} \left(x - \frac{L}{2}\right) \psi_1^{(1)} dx$$

The first integral is zero (see the evaluation of this integral for  $E_1^{(1)}$  above in part a.) The fourth integral is neglected since it is proportional to  $L^{-2}$ . The second and third integrals are the same and are combined:

$$\mu_{\text{induced}} = -2e \int_0^L \psi_1^{(0)*} \left(x - \frac{L}{2}\right) \psi_1^{(1)} dx$$

Substituting  $\psi_1^{(0)} = \frac{2}{L} \frac{1}{2} \sin \frac{x}{L}$  and  $\psi_1^{(1)} = \frac{32mL^3e}{27\hbar^2} \frac{2}{L} \frac{1}{2} \sin \frac{2x}{L}$ , we obtain:

$$\mu_{\text{induced}} = -2e \frac{32mL^3e}{27\hbar^2} \frac{2}{L} \int_0^L \sin \frac{x}{L} \left(x - \frac{L}{2}\right) \sin \frac{2x}{L} dx$$

These integrals are familiar from part a:

$$\mu_{\text{induced}} = -2e \frac{32mL^3e}{27\hbar^2} \frac{2}{L} \frac{8L^2}{9}$$

$$\mu_{\text{induced}} = \frac{mL^4e^2}{\hbar^2} \frac{2^{10}}{3^5}$$

$$c. \quad = \frac{\mu}{\epsilon_0} = \frac{mL^4e^2}{\hbar^2} \frac{2^{10}}{3^5}$$

The larger the box (molecule), the more polarizable the electron density.

## Section 2 Exercises, Problems, and Solutions

### Review Exercises:

1. Draw qualitative shapes of the (1) s, (3) p and (5) d "tangent sphere" atomic orbitals (note that these orbitals represent only the angular portion and do not contain the radial portion of the hydrogen like atomic wavefunctions) Indicate with  $\pm$  the relative signs of the wavefunctions and the position(s) (if any) of any nodes.
2. Define the symmetry adapted "core" and "valence" orbitals of the following systems:
  - i.  $\text{NH}_3$  in the  $C_{3v}$  point group,
  - ii.  $\text{H}_2\text{O}$  in the  $C_{2v}$  point group,
  - iii.  $\text{H}_2\text{O}_2$  (cis) in the  $C_2$  point group,
  - iv. N in  $D_{4h}$ ,  $D_{2h}$ ,  $C_{2v}$ , and  $C_s$  point groups,
  - v.  $\text{N}_2$  in  $D_{4h}$ ,  $D_{2h}$ ,  $C_{2v}$ , and  $C_s$  point groups.
3. Plot the radial portions of the 4s, 4p, 4d, and 4f hydrogen like atomic wavefunctions.
4. Plot the radial portions of the 1s, 2s, 2p, 3s, and 3p hydrogen like atomic wavefunctions for the Si atom using screening concepts for any inner electrons.

### Exercises:

1. In quantum chemistry it is quite common to use combinations of more familiar and easy-to-handle "basis functions" to approximate atomic orbitals. Two common types of basis functions are the Slater type orbitals (STO's) and gaussian type orbitals (GTO's). STO's have the normalized form:

$$\frac{2}{a_0} \left(\frac{r}{a_0}\right)^{n+\frac{1}{2}} \frac{1}{(2n)!} \frac{1}{2} r^{n-1} e^{-\frac{r}{a_0}} Y_{l,m}(\theta, \phi),$$

whereas GTO's have the form:

$$N r^l e^{-\alpha r^2} Y_{l,m}(\theta, \phi).$$

Orthogonalize (using Löwdin (symmetric) orthogonalization) the following 1s (core), 2s (valence), and 3s (Rydberg) STO's for the Li atom given:

$$\text{Li}_{1s} = 2.6906$$

$$\text{Li}_{2s} = 0.6396$$

$$\text{Li}_{3s} = 0.1503.$$

Express the three resultant orthonormal orbitals as linear combinations of these three normalized STO's.

2. Calculate the expectation value of  $r$  for each of the orthogonalized 1s, 2s, and 3s Li orbitals found in Exercise 1.
3. Draw a plot of the radial probability density (e.g.,  $r^2[R_{nl}(r)]^2$  with R referring to the radial portion of the STO) versus  $r$  for each of the orthonormal Li s orbitals found in Exercise 1.

### Problems:

1. Given the following orbital energies (in hartrees) for the N atom and the coupling elements between two like atoms (these coupling elements are the Fock matrix elements from standard *ab-initio* minimum-basis SCF calculations), calculate the molecular orbital

energy levels and 1-electron wavefunctions. Draw the orbital correlation diagram for formation of the  $N_2$  molecule. Indicate the symmetry of each atomic and molecular orbital. Designate each of the molecular orbitals as bonding, non-bonding, or antibonding.

$$\begin{aligned}
 N_{1s} &= -15.31^* \\
 N_{2s} &= -0.86^* \\
 N_{2p} &= -0.48^* \\
 N_{2g} \text{ Fock matrix}^* & \\
 &\begin{matrix} -6.52 \\ -6.22 & -7.06 \\ 3.61 & 4.00 & -3.92 \end{matrix} \\
 N_{2g} \text{ Fock matrix}^* & \\
 &[0.28] \\
 N_{2u} \text{ Fock matrix}^* & \\
 &\begin{matrix} 1.02 \\ -0.60 & -7.59 \\ 0.02 & 7.42 & -8.53 \end{matrix} \\
 N_{2u} \text{ Fock matrix}^* & \\
 &[-0.58]
 \end{aligned}$$

\*The Fock matrices (and orbital energies) were generated using standard STO3G minimum basis set SCF calculations. The Fock matrices are in the orthogonal basis formed from these orbitals.

2. Given the following valence orbital energies for the C atom and  $H_2$  molecule draw the orbital correlation diagram for formation of the  $CH_2$  molecule (via a  $C_{2v}$  insertion of C into  $H_2$  resulting in bent  $CH_2$ ). Designate the symmetry of each atomic and molecular orbital in both their highest point group symmetry and in that of the reaction path ( $C_{2v}$ ).

$$\begin{aligned}
 C_{1s} &= -10.91^* & H_{2g} &= -0.58^* \\
 C_{2s} &= -0.60^* & H_{2u} &= 0.67^* \\
 C_{2p} &= -0.33^*
 \end{aligned}$$

\*The orbital energies were generated using standard STO3G minimum basis set SCF calculations.

3. Using the empirical parameters given below for C and H (taken from Appendix F and "The HMO Model and its Applications" by E. Heilbronner and H. Bock, Wiley-Interscience, NY, 1976), apply the Hückel model to ethylene in order to determine the valence electronic structure of this system. Note that you will be obtaining the 1-electron energies and wavefunctions by solving the secular equation (as you always will when the energy is dependent upon a set of linear parameters like the MO coefficients in the LCAO-MO approach) using the definitions for the matrix elements found in Appendix F.

$$\begin{aligned}
 C_{2p} &= -11.4 \text{ eV} \\
 C_{sp^2} &= -14.7 \text{ eV} \\
 H_s &= -13.6 \text{ eV} \\
 C-C_{2p-2p} &= -1.2 \text{ eV}
 \end{aligned}$$

$$\text{C-C } sp^2-sp^2 = -5.0 \text{ eV}$$

$$\text{C-H } sp^2-s = -4.0 \text{ eV}$$

a. Determine the C=C ( $2p$ ) 1-electron molecular orbital energies and wavefunctions. Calculate the  $\pi \rightarrow \pi^*$  transition energy for ethylene within this model.

b. Determine the C-C ( $sp^2$ ) 1-electron molecular orbital energies and wavefunctions.

c. Determine the C-H ( $sp^2-s$ ) 1-electron molecular orbital energies and wavefunctions (note that appropriate choice of symmetry will reduce this  $8 \times 8$  matrix down to 4  $2 \times 2$  matrices; that is, you are encouraged to symmetry adapt the atomic orbitals before starting the Hückel calculation). Draw a qualitative orbital energy diagram using the HMO energies you have calculated.

4. Using the empirical parameters given below for B and H (taken from Appendix F and "The HMO Model and its Applications" by E. Heilbronner and H. Bock, Wiley-Interscience, NY, 1976), apply the Hückel model to borane ( $BH_3$ ) in order to determine the valence electronic structure of this system.

$$\text{B } 2p = -8.5 \text{ eV}$$

$$\text{B } sp^2 = -10.7 \text{ eV}$$

$$\text{H } s = -13.6 \text{ eV}$$

$$\text{B-H } sp^2-s = -3.5 \text{ eV}$$

Determine the symmetries of the resultant molecular orbitals in the  $D_{3h}$  point group. Draw a qualitative orbital energy diagram using the HMO energies you have calculated.

5. Qualitatively analyze the electronic structure (orbital energies and 1-electron wavefunctions) of  $PF_5$ . Analyze only the 3s and 3p electrons of P and the one 2p bonding electron of each F. Proceed with a  $D_{3h}$  analysis in the following manner:

a. Symmetry adapt the top and bottom F atomic orbitals.

b. Symmetry adapt the three (trigonal) F atomic orbitals.

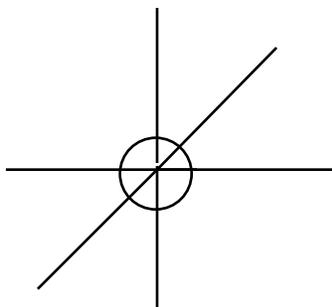
c. Symmetry adapt the P 3s and 3p atomic orbitals.

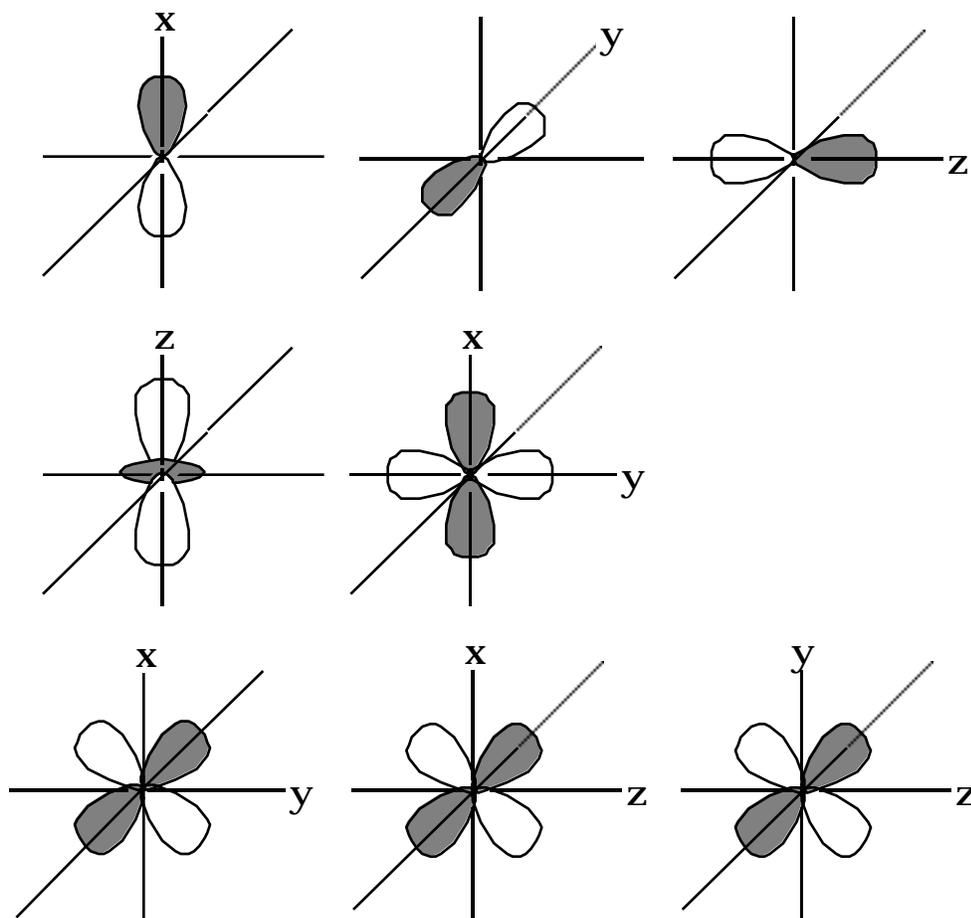
d. Allow these three sets of  $D_{3h}$  orbitals to interact and draw the resultant orbital energy diagram. Symmetry label each of these molecular energy levels. Fill this energy diagram with 10 "valence" electrons.

## Solutions

### Review Exercises

1.





2. i. In ammonia the only "core" orbital is the N 1s and this becomes an  $a_1$  orbital in  $C_{3v}$  symmetry. The N 2s orbitals and 3 H 1s orbitals become 2  $a_1$  and an e set of orbitals. The remaining N 2p orbitals also become 1  $a_1$  and a set of e orbitals. The total valence orbitals in  $C_{3v}$  symmetry are 3 $a_1$  and 2e orbitals.

2. ii. In water the only core orbital is the O 1s and this becomes an  $a_1$  orbital in  $C_{2v}$  symmetry. Placing the molecule in the yz plane allows us to further analyze the remaining valence orbitals as: O  $2p_z = a_1$ , O  $2p_y$  as  $b_2$ , and O  $2p_x$  as  $b_1$ . The H 1s + H 1s combination is an  $a_1$  whereas the H 1s - H 1s combination is a  $b_2$ .

2. iii. Placing the oxygens of  $H_2O_2$  in the yz plane (z bisecting the oxygens) and the (cis) hydrogens distorted slightly in +x and -x directions allows us to analyze the orbitals as follows. The core O 1s + O 1s combination is an  $a$  orbital whereas the O 1s - O 1s combination is a  $b$  orbital. The valence orbitals are: O  $2s + O 2s = a$ , O  $2s - O 2s = b$ , O  $2p_x + O 2p_x = b$ , O  $2p_x - O 2p_x = a$ , O  $2p_y + O 2p_y = a$ , O  $2p_y - O 2p_y = b$ , O  $2p_z + O 2p_z = b$ , O  $2p_z - O 2p_z = a$ , H 1s + H 1s = a, and finally the H 1s - H 1s = b.

2. iv. For the next two problems we will use the convention of choosing the z axis as principal axis for the  $D_h$ ,  $D_{2h}$ , and  $C_{2v}$  point groups and the xy plane as the horizontal reflection plane in  $C_s$  symmetry.

	$D_h$	$D_{2h}$	$C_{2v}$	$C_s$
N 1s	g	$a_g$	$a_1$	$a'$
N 2s	g	$a_g$	$a_1$	$a'$

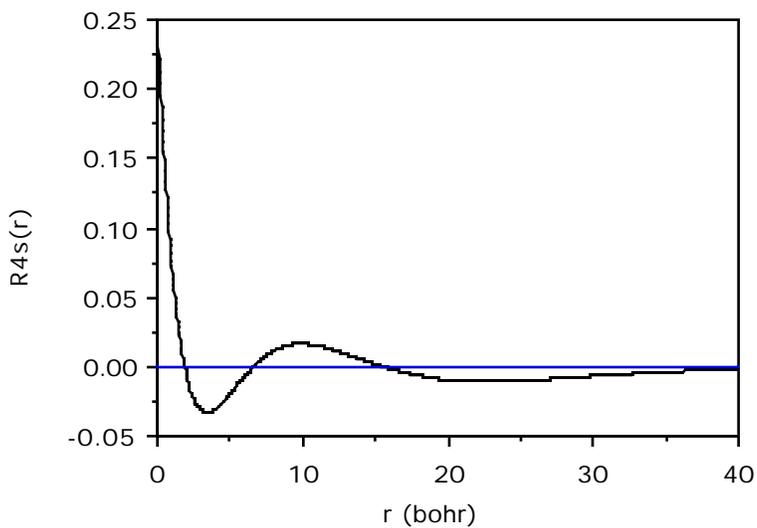
N 2p <sub>x</sub>	xu	b <sub>3u</sub>	b <sub>1</sub>	a'
N 2p <sub>y</sub>	yu	b <sub>2u</sub>	b <sub>2</sub>	a'
N 2p <sub>z</sub>	u	b <sub>1u</sub>	a <sub>1</sub>	a''

2. v. The Nitrogen molecule is in the yz plane for all point groups except the C<sub>s</sub> in which case it is placed in the xy plane.

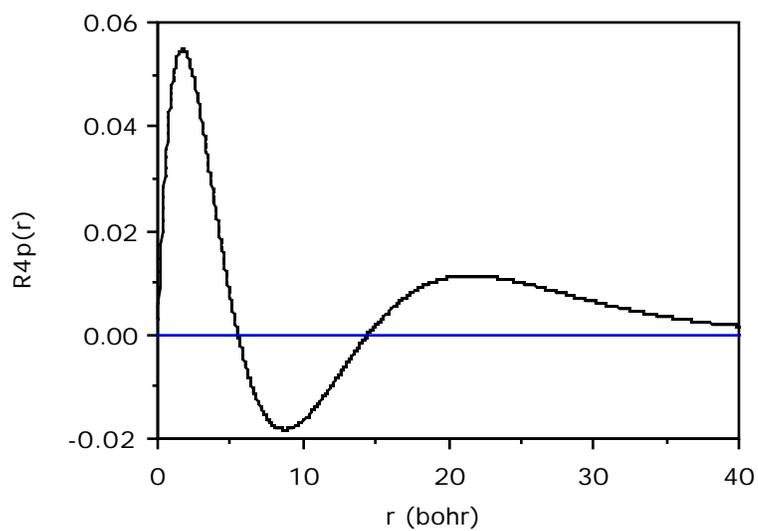
	D <sub>h</sub>	D <sub>2h</sub>	C <sub>2v</sub>	C <sub>s</sub>
N 1s + N 1s	g	a <sub>g</sub>	a <sub>1</sub>	a'
N 1s - N 1s	u	b <sub>1u</sub>	b <sub>2</sub>	a'
N 2s + N 2s	g	a <sub>g</sub>	a <sub>1</sub>	a'
N 2s - N 2s	u	b <sub>1u</sub>	b <sub>2</sub>	a'
N 2p <sub>x</sub> + N 2p <sub>x</sub>	xu	b <sub>3u</sub>	b <sub>1</sub>	a'
N 2p <sub>x</sub> - N 2p <sub>x</sub>	xg	b <sub>2g</sub>	a <sub>2</sub>	a'
N 2p <sub>y</sub> + N 2p <sub>y</sub>	yu	b <sub>2u</sub>	a <sub>1</sub>	a'
N 2p <sub>y</sub> - N 2p <sub>y</sub>	yg	b <sub>3g</sub>	b <sub>2</sub>	a'
N 2p <sub>z</sub> + N 2p <sub>z</sub>	u	b <sub>1u</sub>	b <sub>2</sub>	a''
N 2p <sub>z</sub> - N 2p <sub>z</sub>	g	a <sub>g</sub>	a <sub>1</sub>	a''

3.

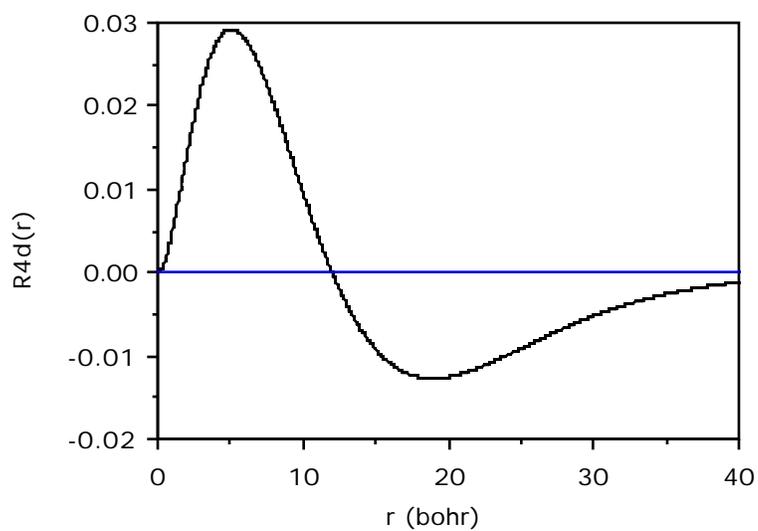
Hydrogen 4s Radial Function



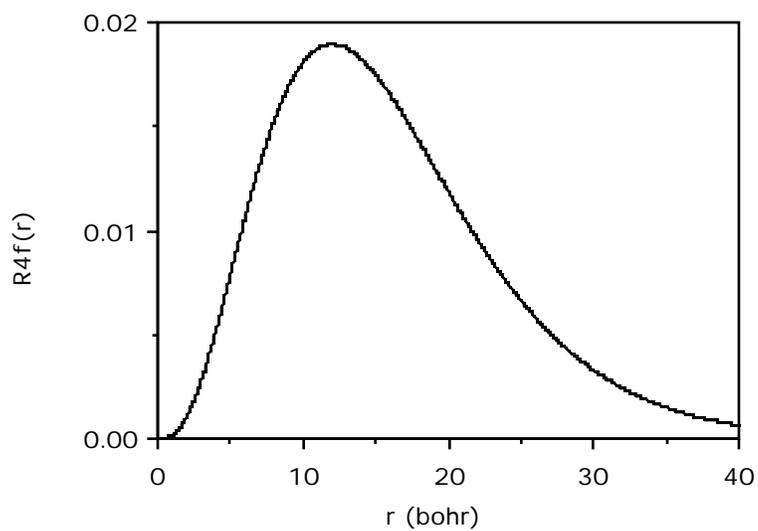
Hydrogen 4p Radial Function



Hydrogen 4d Radial Function

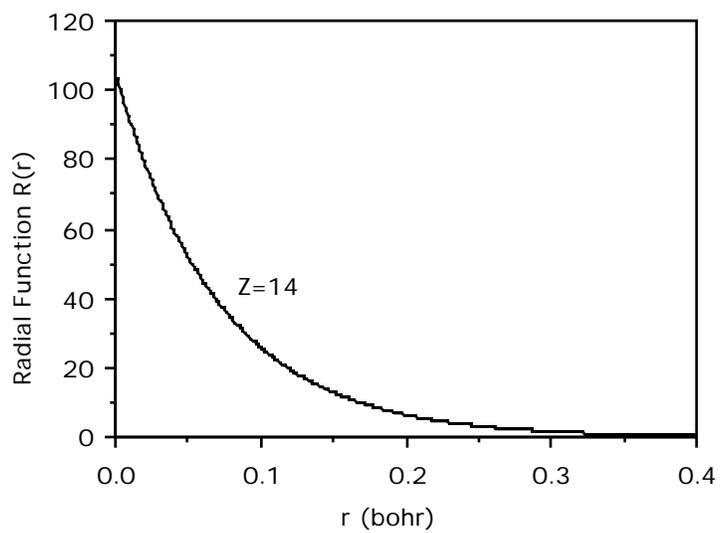


Hydrogen 4f Radial Function

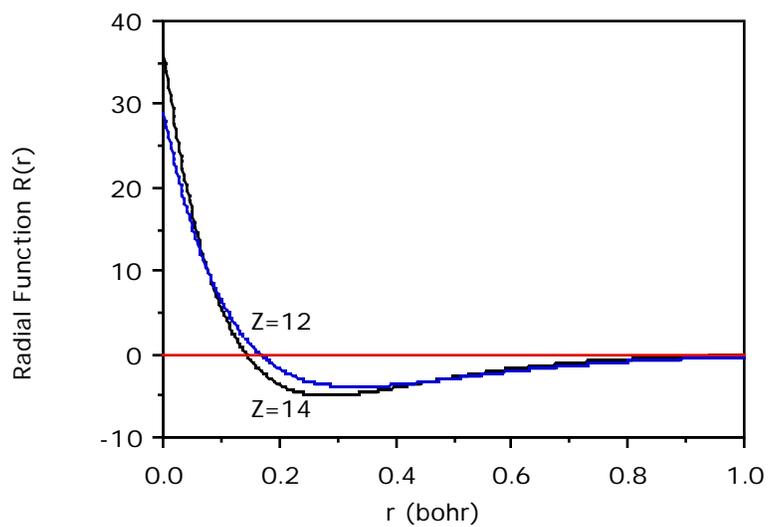


4.

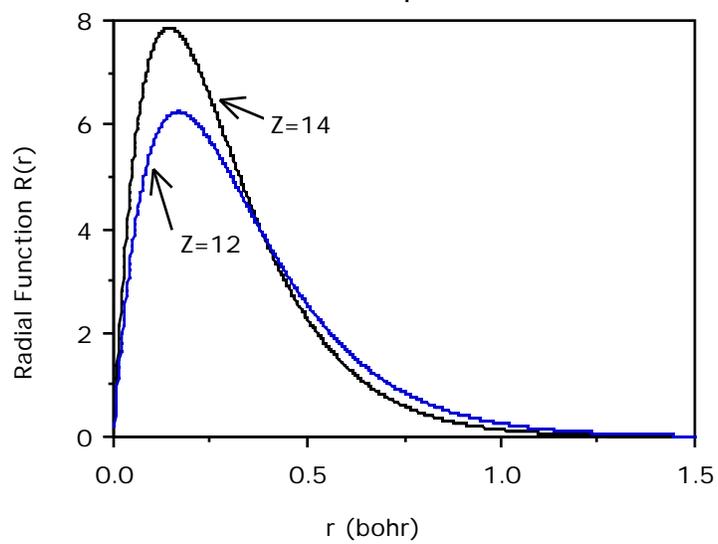
Si 1s



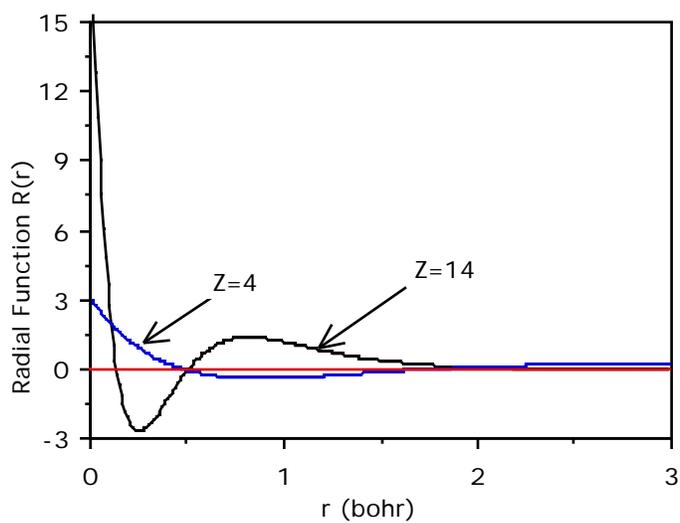
Si 2s



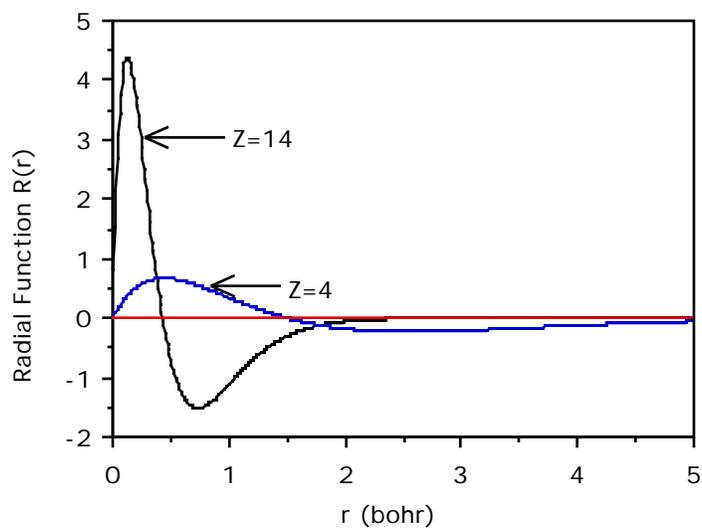
Si 2p



Si 3s



Si 3p



Exercises

1. Two Slater type orbitals,  $i$  and  $j$ , centered on the same point results in the following overlap integrals:

$$S_{ij} = \int_0^{\infty} \int_0^{\pi} \int_0^{2\pi} \frac{2}{a_0} r^{n_i+1/2} \frac{1}{(2n_i)!} \frac{1}{2} r^{(n_i-1)} e^{-\frac{r}{a_0}} Y_{l_i, m_i}(\theta, \phi) \cdot \frac{2}{a_0} r^{n_j+1/2} \frac{1}{(2n_j)!} \frac{1}{2} r^{(n_j-1)} e^{-\frac{r}{a_0}} Y_{l_j, m_j}(\theta, \phi) \cdot r^2 \sin \theta \, dr \, d\theta \, d\phi.$$

For these s orbitals  $l = m = 0$  and  $Y_{0,0}(\theta, \phi) = \frac{1}{\sqrt{4\pi}}$ . Performing the integrations over  $\theta$  and  $\phi$  yields  $4\pi$  which then cancels with these Y terms. The integral then reduces to:

$$S_{ij} = \frac{2}{a_0} r^{n_i+1/2} \frac{1}{(2n_i)!} \frac{1}{2} \frac{2}{a_0} r^{n_j+1/2} \frac{1}{(2n_j)!} \frac{1}{2} \int_0^{\infty} r^{(n_i-1+n_j-1)} e^{-\frac{(n_i+n_j)r}{a_0}} r^2 dr$$

$$= \frac{2}{a_0} r^{n_i+1/2} \frac{1}{(2n_i)!} \frac{1}{2} \frac{2}{a_0} r^{n_j+1/2} \frac{1}{(2n_j)!} \frac{1}{2} \int_0^{\infty} r^{(n_i+n_j)} e^{-\frac{(n_i+n_j)r}{a_0}} dr$$

Using integral equation (4) the integral then reduces to:

$$S_{ij} = \frac{2}{a_0} r^{n_i+1/2} \frac{1}{(2n_i)!} \frac{1}{2} \frac{2}{a_0} r^{n_j+1/2} \frac{1}{(2n_j)!} \frac{1}{2} (n_i+n_j)! \frac{a_0^{n_i+n_j+1}}{(n_i+n_j)!}.$$

We then substitute in the values for each of these constants:

$$\text{for } i=1; n=1, l=m=0, \text{ and } a_0 = 2.6906$$

$$\text{for } i=2; n=2, l=m=0, \text{ and } a_0 = 0.6396$$

$$\text{for } i=3; n=3, l=m=0, \text{ and } a_0 = 0.1503.$$

Evaluating each of these matrix elements we obtain:

$$\begin{aligned} S_{11} &= (12.482992)(0.707107)(12.482992) \\ &= (0.707107)(2.000000)(0.006417) \\ &= 1.000000 \\ S_{21} = S_{12} &= (1.850743)(0.204124)(12.482992) \\ &= (0.707107)(6.000000)(0.008131) \\ &= 0.162673 \\ S_{22} &= (1.850743)(0.204124)(1.850743) \\ &= (0.204124)(24.000000)(0.291950) \\ &= 1.000000 \end{aligned}$$

$$\begin{aligned}
S_{31} = S_{13} &= (0.014892)(0.037268)(12.482992) \\
&= (0.707107)(24.000000)(0.005404) \\
&= 0.000635 \\
S_{32} = S_{23} &= (0.014892)(0.037268)(1.850743) \\
&= (0.204124)(120.000000)(4.116872) \\
&= 0.103582 \\
S_{33} &= (0.014892)(0.037268)(0.014892) \\
&= (0.037268)(720.000000)(4508.968136) \\
&= 1.000000
\end{aligned}$$

$$\begin{aligned}
&1.000000 \\
S &= \begin{matrix} 0.162673 & 1.000000 \\ 0.000635 & 0.103582 & 1.000000 \end{matrix}
\end{aligned}$$

We now solve the matrix eigenvalue problem  $S U = \lambda U$ .

The eigenvalues,  $\lambda$ , of this overlap matrix are:

$$[ 0.807436 \quad 0.999424 \quad 1.193139 ] ,$$

and the corresponding eigenvectors,  $U$ , are:

$$\begin{matrix}
0.596540 & -0.537104 & -0.596372 \\
-0.707634 & -0.001394 & -0.706578 \\
0.378675 & 0.843515 & -0.380905
\end{matrix}$$

The  $\frac{1}{2}$  matrix becomes:

$$\frac{1}{2} = \begin{matrix} 1.112874 & 0.000000 & 0.000000 \\ 0.000000 & 1.000288 & 0.000000 \\ 0.000000 & 0.000000 & 0.915492 \end{matrix}$$

Back transforming into the original eigenbasis gives  $S^{-\frac{1}{2}}$ , e.g.

$$\begin{aligned}
S^{-\frac{1}{2}} &= U^{-\frac{1}{2}} U^T \\
&= \begin{matrix} 1.010194 \\ -0.083258 & 1.014330 \\ 0.006170 & -0.052991 & 1.004129 \end{matrix}
\end{aligned}$$

The old ao matrix can be written as:

$$\begin{aligned}
C &= \begin{matrix} 1.000000 & 0.000000 & 0.000000 \\ 0.000000 & 1.000000 & 0.000000 \\ 0.000000 & 0.000000 & 1.000000 \end{matrix}
\end{aligned}$$

The new ao matrix (which now gives each ao as a linear combination of the original aos) then becomes:

$$C' = S^{-\frac{1}{2}} C = \begin{pmatrix} 1.010194 & -0.083258 & 0.006170 \\ -0.083258 & 1.014330 & -0.052991 \\ 0.006170 & -0.052991 & 1.004129 \end{pmatrix}$$

These new aos have been constructed to meet the orthonormalization requirement  $C'^T S C' = 1$  since:

$$S^{-\frac{1}{2}} C'^T S S^{-\frac{1}{2}} C' = C'^T S^{-\frac{1}{2}} S S^{-\frac{1}{2}} C' = C'^T C' = 1.$$

But, it is always good to check our result and indeed:

$$C'^T S C' = \begin{pmatrix} 1.000000 & 0.000000 & 0.000000 \\ 0.000000 & 1.000000 & 0.000000 \\ 0.000000 & 0.000000 & 1.000000 \end{pmatrix}$$

2. The least time consuming route here is to evaluate each of the needed integrals first. These are evaluated analogous to exercise 1, letting  $\psi_i$  denote each of the individual Slater Type Orbitals.

$$\int_0^{\infty} \psi_i^* r \psi_j r^2 dr = \langle r \rangle_{ij}$$

$$= \frac{2}{a_0} \psi_i^{n_i + \frac{1}{2}} \frac{1}{(2n_i)!} \frac{1}{2} \frac{2}{a_0} \psi_j^{n_j + \frac{1}{2}} \frac{1}{(2n_j)!} \int_0^{\infty} r^{(n_i + n_j + 1)} e^{-\frac{(i+j)r}{a_0}} dr$$

Once again using integral equation (4) the integral reduces to:

$$= \frac{2}{a_0} \psi_i^{n_i + \frac{1}{2}} \frac{1}{(2n_i)!} \frac{1}{2} \frac{2}{a_0} \psi_j^{n_j + \frac{1}{2}} \frac{1}{(2n_j)!} \frac{1}{2^{(n_i + n_j + 1)}} \frac{a_0^{n_i + n_j + 2}}{(i+j)}.$$

Again, upon substituting in the values for each of these constants, evaluation of these expectation values yields:

$$\begin{aligned} \langle r \rangle_{11} &= (12.482992)(0.707107)(12.482992) \\ &= (0.707107)(6.000000)(0.001193) \\ &= 0.557496 \\ \langle r \rangle_{21} = \langle r \rangle_{12} &= (1.850743)(0.204124)(12.482992) \\ &= (0.707107)(24.000000)(0.002441) \\ &= 0.195391 \\ \langle r \rangle_{22} &= (1.850743)(0.204124)(1.850743) \\ &= (0.204124)(120.000000)(0.228228) \\ &= 3.908693 \\ \langle r \rangle_{31} = \langle r \rangle_{13} &= (0.014892)(0.037268)(12.482992) \\ &= (0.707107)(120.000000)(0.001902) \\ &= 0.001118 \\ \langle r \rangle_{32} = \langle r \rangle_{23} &= (0.014892)(0.037268)(1.850743) \\ &= (0.204124)(720.000000)(5.211889) \end{aligned}$$

$$\begin{aligned} &= 0.786798 \\ \langle r \rangle_{33} &= \frac{(0.014892)(0.037268)(0.014892)}{(0.037268)(5040.000000)(14999.893999)} \\ &= \frac{23.286760}{0.557496} \end{aligned}$$

$$\int_0^{\infty} r_i r_j r^2 dr = \langle r \rangle_{ij} = \begin{matrix} 0.195391 & 3.908693 \\ 0.001118 & 0.786798 & 23.286760 \end{matrix}$$

Using these integrals one then proceeds to evaluate the expectation values of each of the orthogonalized aos,  $\psi_n$ , as:

$$\int_0^{\infty} \psi_n r \psi_n r^2 dr = \sum_{i=1}^3 \sum_{j=1}^3 C'_{ni} C'_{nj} \langle r \rangle_{ij}.$$

This results in the following expectation values (in atomic units):

$$\int_0^{\infty} \psi_{1s} r \psi_{1s} r^2 dr = 0.563240 \text{ bohr}$$

$$\int_0^{\infty} \psi_{2s} r \psi_{2s} r^2 dr = 3.973199 \text{ bohr}$$

$$\int_0^{\infty} \psi_{3s} r \psi_{3s} r^2 dr = 23.406622 \text{ bohr}$$

3. The radial density for each orthogonalized orbital,  $\psi_n$ , assuming integrations over  $\theta$  and  $\phi$  have already been performed can be written as:

$$\int_0^{\infty} \psi_n \psi_n r^2 dr = \sum_{i=1}^3 \sum_{j=1}^3 C'_{ni} C'_{nj} \int_0^{\infty} R_i R_j r^2 dr, \text{ where } R_i \text{ and } R_j \text{ are the radial portions}$$

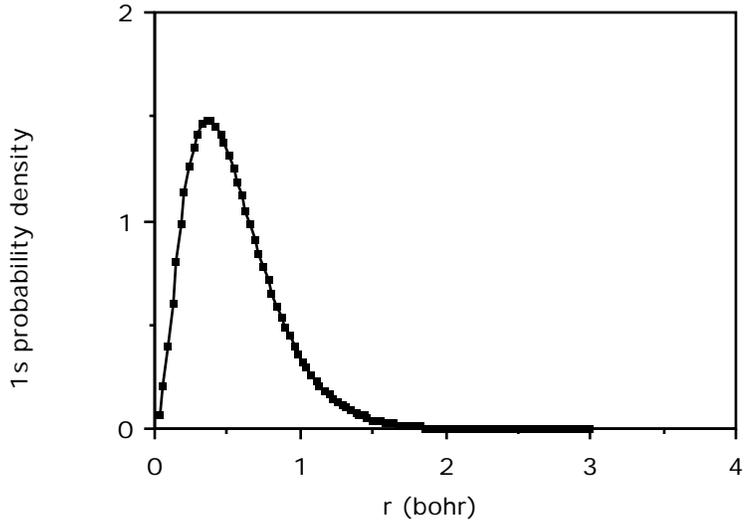
of the individual Slater Type Orbitals, e.g.,

$$R_i R_j r^2 = \frac{2}{a_0} \left( \frac{1}{2} \right)^{n_i + \frac{1}{2}} \frac{1}{(2n_i)!} \frac{1}{2} \frac{2}{a_0} \left( \frac{1}{2} \right)^{n_j + \frac{1}{2}} \frac{1}{(2n_j)!} \frac{1}{2} r^{(n_i + n_j)} e^{-\frac{(i+j)r}{a_0}}$$

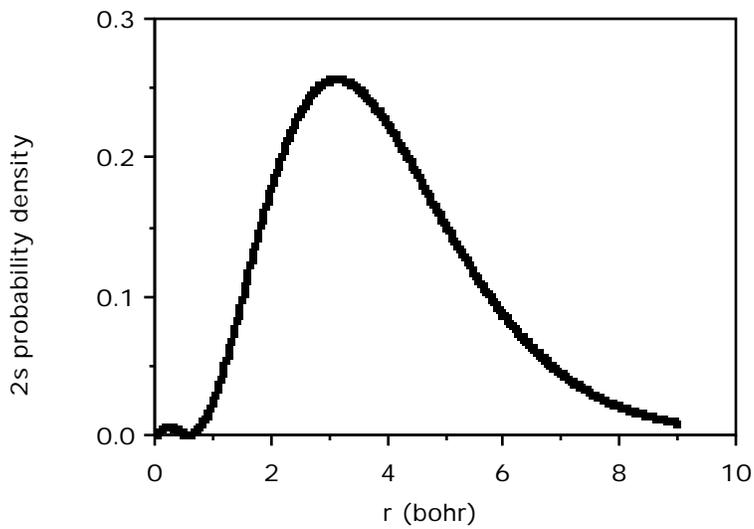
Therefore a plot of the radial probability for a given orthogonalized atomic orbital,  $n$ , will

$$\text{be: } \sum_{i=1}^3 \sum_{j=1}^3 C'_{ni} C'_{nj} R_i R_j r^2 \text{ vs. } r.$$

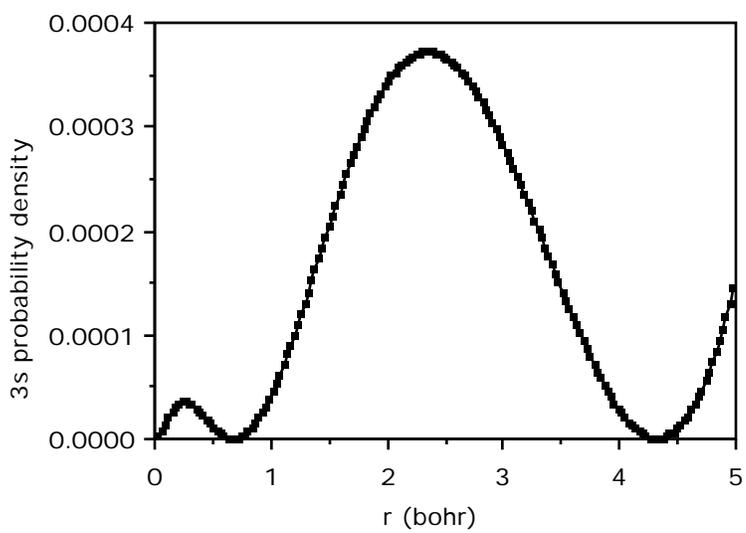
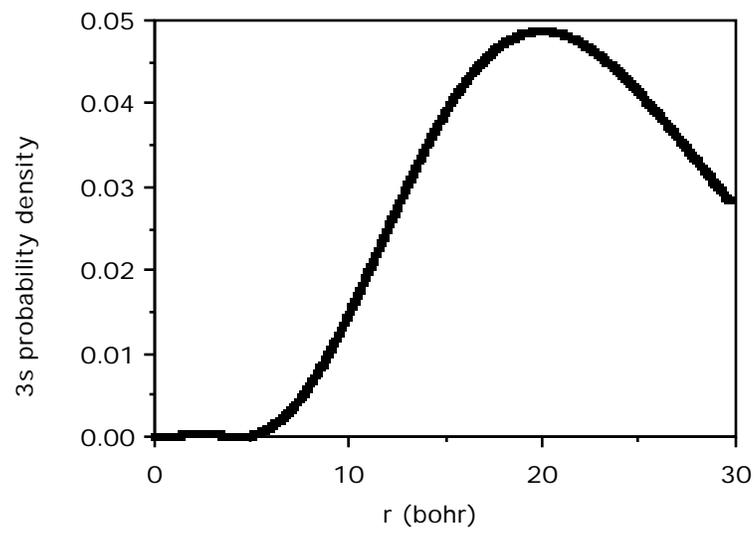
Plot of the orthogonalized 1s orbital probability density vs  $r$ ; note there are no nodes.



Plot of the orthogonalized 2s orbital probability density vs r; note there is one node.

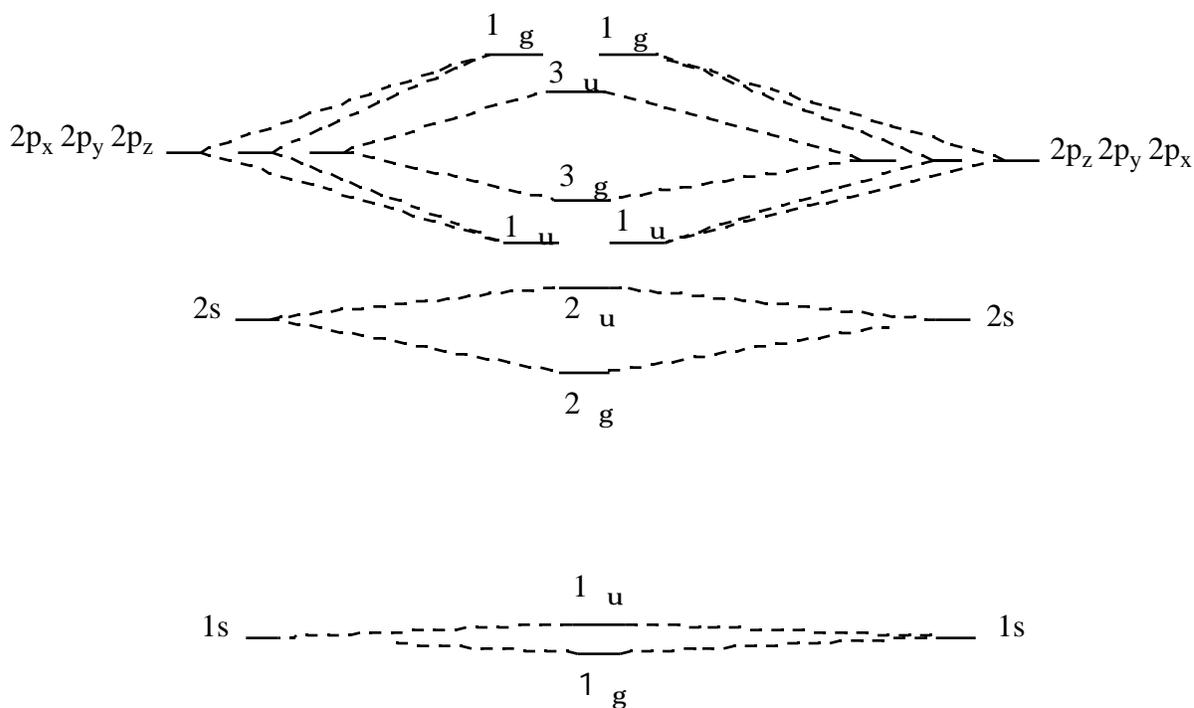


Plot of the orthogonalized 3s orbital probability density vs r; note there are two nodes in the 0-5 bohr region but they are not distinguishable as such. A duplicate plot with this nodal region expanded follows.



Problems

1.



The above diagram indicates how the SALC-AOs are formed from the 1s, 2s, and 2p N atomic orbitals. It can be seen that there are 3  $g$ , 3  $u$ , 1  $u_x$ , 1  $u_y$ , 1  $g_x$ , and 1  $g_y$  SALC-AOs. The Hamiltonian matrices (Fock matrices) are given. Each of these can be diagonalized to give the following MO energies:

3  $g$ ; -15.52, -1.45, and -0.54 (hartrees)

3  $u$ ; -15.52, -0.72, and 1.13

1  $u_x$ ; -0.58

1  $u_y$ ; -0.58

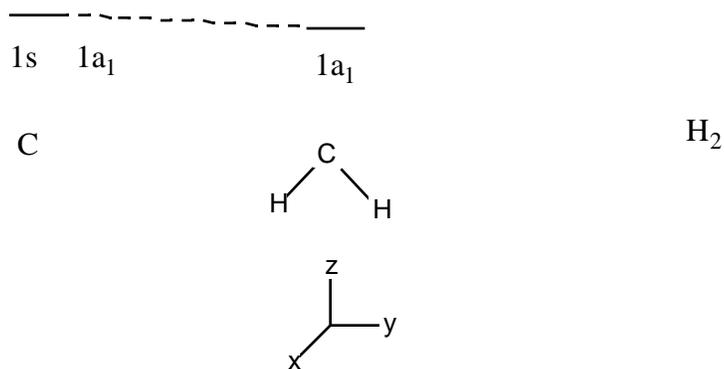
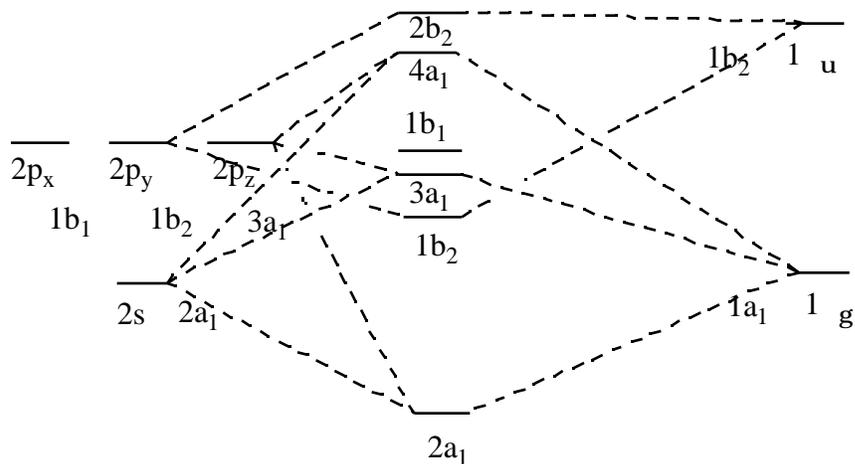
1  $g_x$ ; 0.28

1  $g_y$ ; 0.28

It can be seen that the 3  $g$  orbitals are bonding, the 3  $u$  orbitals are antibonding, the 1  $u_x$  and 1  $u_y$  orbitals are bonding, and the 1  $g_x$  and 1  $g_y$  orbitals are antibonding. The eigenvectors one obtains are in the orthogonal basis and therefore pretty meaningless.

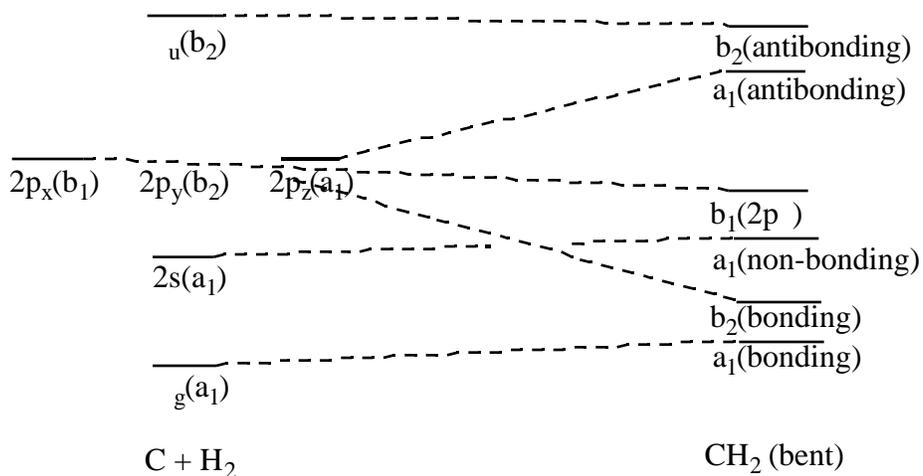
Back transformation into the original basis will generate the expected results for the 1e-MOs (expected combinations of SALC-AOs).

2. Using these approximate energies we can draw the following MO diagram:

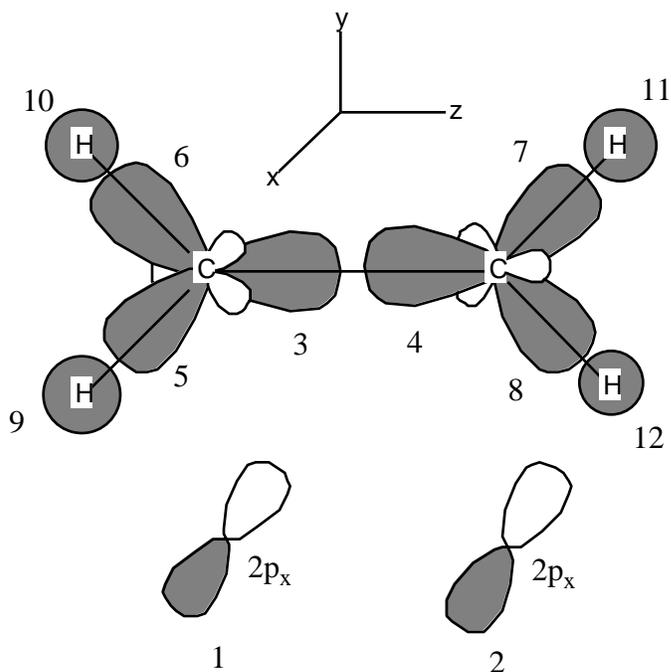


This MO diagram is not an orbital correlation diagram but can be used to help generate one. The energy levels on each side (C and H<sub>2</sub>) can be "superimposed" to generate the left side of the orbital correlation diagram and the center CH<sub>2</sub> levels can be used to form the right side. Ignoring the core levels this generates the following orbital correlation diagram.

Orbital-correlation diagram for the reaction  $C + H_2 \rightarrow CH_2$  (bent)



3.



Using  $D_{2h}$  symmetry and labeling the orbitals ( $f_1$ - $f_{12}$ ) as shown above proceed by using the orbitals to define a reducible representation, which may be subsequently reduced to its irreducible components. Use projectors to find the SALC-AOs for these irreps.

3. a. The  $2P_x$  orbitals on each carbon form the following reducible representation:

$$\begin{array}{cccccccc}
 D_{2h} & E & C_2(z) & C_2(y) & C_2(x) & i & (xy) & (xz) & (yz) \\
 2p_x & 2 & -2 & 0 & 0 & 0 & 0 & 2 & -2
 \end{array}$$

The number of irreducible representations may be found by using the following formula:

$$n_{\text{irrep}} = \frac{1}{g} \sum_{\text{R}} \text{red}(\text{R}) \cdot \text{irrep}(\text{R}),$$

where  $g$  = the order of the point group (8 for  $D_{2h}$ ).

$$\begin{aligned}
 n_{A_g} &= \frac{1}{8} \sum_{\text{R}} 2p_x(\text{R}) \cdot A_g(\text{R}) \\
 &= \frac{1}{8} \{ (2)(1) + (-2)(1) + (0)(1) + (0)(1) + \\
 &\quad (0)(1) + (0)(1) + (2)(1) + (-2)(1) \} = 0
 \end{aligned}$$

Similarly,

$$\begin{aligned}
 n_{B_{1g}} &= 0 \\
 n_{B_{2g}} &= 1 \\
 n_{B_{3g}} &= 0 \\
 n_{A_u} &= 0 \\
 n_{B_{1u}} &= 0 \\
 n_{B_{2u}} &= 0 \\
 n_{B_{3u}} &= 1
 \end{aligned}$$

Projectors using the formula:

$$P_{\text{irrep}} = \sum_{\mathbf{R}} \chi_{\text{irrep}}(\mathbf{R}) \mathbf{R} ,$$

may be used to find the SALC-AOs for these irreducible representations.

$$P_{B_{2g}} = \sum_{\mathbf{R}} \chi_{B_{2g}}(\mathbf{R}) \mathbf{R} ,$$

$$\begin{aligned} P_{B_{2g}} f_1 &= (1)E f_1 + (-1)C_2(z) f_1 + (1)C_2(y) f_1 + (-1)C_2(x) f_1 + \\ &\quad (1)i f_1 + (-1) (xy) f_1 + (1) (xz) f_1 + (-1) (yz) f_1 \\ &= (1) f_1 + (-1) f_1 + (1) f_2 + (-1) f_2 + \\ &\quad (1) f_2 + (-1) f_2 + (1) f_1 + (-1) f_1 \\ &= f_1 + f_1 - f_2 - f_2 - f_2 - f_2 + f_1 + f_1 \\ &= 4f_1 - 4f_2 \end{aligned}$$

Normalization of this SALC-AO (and representing the SALC-AOs with  $b_{2g}$ ) yields:

$$N(f_1 - f_2)N(f_1 - f_2)d = 1$$

$$N^2 (f_1 f_1 d - f_1 f_2 d - f_2 f_1 d + f_2 f_2 d) = 1$$

$$N^2(1 - 1) = 1$$

$$2N^2 = 1$$

$$N = \frac{1}{\sqrt{2}}$$

$$1b_{2g} = \frac{1}{\sqrt{2}}(f_1 - f_2) .$$

The  $B_{3u}$  SALC-AO may be found in a similar fashion:

$$\begin{aligned} P_{B_{3u}} f_1 &= (1) f_1 + (-1) f_1 + (-1) f_2 + (1) f_2 + \\ &\quad (-1) f_2 + (1) f_2 + (1) f_1 + (-1) f_1 \\ &= f_1 + f_1 + f_2 + f_2 + f_2 + f_2 + f_1 + f_1 \\ &= 4f_1 + 4f_2 \end{aligned}$$

Normalization of this SALC-AO yields:

$$1b_{3u} = \frac{1}{\sqrt{2}}(f_1 + f_2) .$$

Since there are only two SALC-AOs and both are of different symmetry types these SALC-AOs are MOs and the 2x2 Hamiltonian matrix reduces to 2 1x1 matrices.

$$\begin{aligned} H_{1b_{2g},1b_{2g}} &= \frac{1}{\sqrt{2}}(f_1 - f_2)H\frac{1}{\sqrt{2}}(f_1 - f_2)d \\ &= \frac{1}{2} (f_1 H f_1 d - 2 f_1 H f_2 d + f_2 H f_2 d) \\ &= \frac{1}{2} (2p - 2 \cdot 2p - 2p + 2p) \\ &= 2p - 2p - 2p \\ &= -11.4 - (-1.2) = -10.2 \end{aligned}$$

$$H_{1b_{3u},1b_{3u}} = \frac{1}{\sqrt{2}}(f_1 + f_2)H\frac{1}{\sqrt{2}}(f_1 + f_2)d$$

$$\begin{aligned}
&= \frac{1}{2} f_1 H f_1 d + 2 f_1 H f_2 d + f_2 H f_2 d \\
&= \frac{1}{2} \begin{matrix} 2p \\ 2p \end{matrix} + 2 \begin{matrix} 2p \\ 2p \end{matrix} - 2p + \begin{matrix} 2p \\ 2p \end{matrix} \\
&= \begin{matrix} 2p \\ 2p \end{matrix} + \begin{matrix} 2p \\ 2p \end{matrix} - 2p \\
&= -11.4 + (-1.2) = -12.6
\end{aligned}$$

This results in a  $\rightarrow$  \* splitting of 2.4 eV.

3. b. The  $sp^2$  orbitals forming the C-C bond generate the following reducible representation:

$$\begin{array}{cccccccc}
D_{2h} & E & C_2(z) & C_2(y) & C_2(x) & i & (xy) & (xz) & (yz) \\
C_{sp^2} & 2 & 2 & 0 & 0 & 0 & 2 & 2 & 
\end{array}$$

This reducible representation reduces to  $1A_g$  and  $1B_{1u}$  irreducible representations.

Projectors are used to find the SALC-AOs for these irreducible representations.

$$\begin{aligned}
P_{A_g} f_3 &= (1)E f_3 + (1)C_2(z) f_3 + (1)C_2(y) f_3 + (1)C_2(x) f_3 + \\
&\quad (1)i f_3 + (1) (xy) f_3 + (1) (xz) f_3 + (1) (yz) f_3 \\
&= (1) f_3 + (1) f_3 + (1) f_4 + (1) f_4 + \\
&\quad (1) f_4 + (1) f_4 + (1) f_3 + (1) f_3 \\
&= 4f_3 + 4f_4
\end{aligned}$$

Normalization of this SALC-AO yields:

$$1_{ag} = \frac{1}{\sqrt{2}}(f_3 + f_4) .$$

The  $B_{1u}$  SALC-AO may be found in a similar fashion:

$$\begin{aligned}
P_{B_{1u}} f_3 &= (1) f_3 + (1) f_3 + (-1) f_4 + (-1) f_4 + \\
&\quad (-1) f_4 + (-1) f_4 + (1) f_3 + (1) f_3 \\
&= 4f_3 - 4f_4
\end{aligned}$$

Normalization of this SALC-AO yields:

$$1_{b_{1u}} = \frac{1}{\sqrt{2}}(f_3 - f_4) .$$

Again since there are only two SALC-AOs and both are of different symmetry types these SALC-AOs are MOs and the  $2 \times 2$  Hamiltonian matrix reduces to 2  $1 \times 1$  matrices.

$$\begin{aligned}
H_{1_{ag}, 1_{ag}} &= \frac{1}{\sqrt{2}}(f_3 + f_4) H \frac{1}{\sqrt{2}}(f_3 + f_4) d \\
&= \frac{1}{2} f_3 H f_3 d + 2 f_3 H f_4 d + f_4 H f_4 d \\
&= \frac{1}{2} \begin{matrix} sp^2 \\ sp^2 \end{matrix} + 2 \begin{matrix} sp^2-sp^2 \\ sp^2 \end{matrix} + \begin{matrix} sp^2 \\ sp^2 \end{matrix} \\
&= \begin{matrix} sp^2 \\ sp^2 \end{matrix} + \begin{matrix} sp^2-sp^2 \\ sp^2 \end{matrix} \\
&= -14.7 + (-5.0) = -19.7
\end{aligned}$$

$$\begin{aligned}
H_{1_{b_{1u}}, 1_{b_{1u}}} &= \frac{1}{\sqrt{2}}(f_3 - f_4) H \frac{1}{\sqrt{2}}(f_3 - f_4) d \\
&= \frac{1}{2} f_3 H f_3 d - 2 f_3 H f_4 d + f_4 H f_4 d
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{2} \text{ sp}^2 - 2 \text{ sp}^2 - \text{sp}^2 + \text{ sp}^2 \\
&= \text{ sp}^2 - \text{ sp}^2 - \text{sp}^2 \\
&= -14.7 - (-5.0) = -9.7
\end{aligned}$$

3. c. The C  $\text{sp}^2$  orbitals and the H s orbitals forming the C-H bonds generate the following reducible representation:

$$\begin{array}{cccccccc}
D_{2h} & E & C_2(z) & C_2(y) & C_2(x) & i & (xy) & (xz) & (yz) \\
\text{sp}^2\text{-s} & 8 & 0 & 0 & 0 & 0 & 0 & 0 & 8
\end{array}$$

This reducible representation reduces to  $2A_g$ ,  $2B_{3g}$ ,  $2B_{1u}$  and  $2B_{2u}$  irreducible representations.

Projectors are used to find the SALC-AOs for these irreducible representations.

$$\begin{aligned}
P_{A_g} f_6 &= (1)E f_6 + (1)C_2(z) f_6 + (1)C_2(y) f_6 + (1)C_2(x) f_6 + \\
&\quad (1)i f_6 + (1) (xy) f_6 + (1) (xz) f_6 + (1) (yz) f_6 \\
&= (1) f_6 + (1) f_5 + (1) f_7 + (1) f_8 + \\
&\quad (1) f_8 + (1) f_7 + (1) f_5 + (1) f_6 \\
&= 2f_5 + 2f_6 + 2f_7 + 2f_8
\end{aligned}$$

Normalization yields:  $2a_g = \frac{1}{2}(f_5 + f_6 + f_7 + f_8)$  .

$$\begin{aligned}
P_{A_g} f_{10} &= (1)E f_{10} + (1)C_2(z) f_{10} + (1)C_2(y) f_{10} + (1)C_2(x) f_{10} + \\
&\quad (1)i f_{10} + (1) (xy) f_{10} + (1) (xz) f_{10} + (1) (yz) f_{10} \\
&= (1) f_{10} + (1) f_9 + (1) f_{11} + (1) f_{12} + \\
&\quad (1) f_{12} + (1) f_{11} + (1) f_9 + (1) f_{10} \\
&= 2f_9 + 2f_{10} + 2f_{11} + 2f_{12}
\end{aligned}$$

Normalization yields:  $3a_g = \frac{1}{2}(f_9 + f_{10} + f_{11} + f_{12})$  .

$$\begin{aligned}
P_{B_{3g}} f_6 &= (1) f_6 + (-1) f_5 + (-1) f_7 + (1) f_8 + \\
&\quad (1) f_8 + (-1) f_7 + (-1) f_5 + (1) f_6 \\
&= -2f_5 + 2f_6 - 2f_7 + 2f_8
\end{aligned}$$

Normalization yields:  $1b_{3g} = \frac{1}{2}(-f_5 + f_6 - f_7 + f_8)$  .

$$\begin{aligned}
P_{B_{3g}} f_{10} &= (1) f_{10} + (-1) f_9 + (-1) f_{11} + (1) f_{12} + \\
&\quad (1) f_{12} + (-1) f_{11} + (-1) f_9 + (1) f_{10} \\
&= -2f_9 + 2f_{10} - 2f_{11} + 2f_{12}
\end{aligned}$$

Normalization yields:  $2b_{3g} = \frac{1}{2}(-f_9 + f_{10} - f_{11} + f_{12})$  .

$$\begin{aligned}
P_{B_{1u}} f_6 &= (1) f_6 + (1) f_5 + (-1) f_7 + (-1) f_8 + \\
&\quad (-1) f_8 + (-1) f_7 + (1) f_5 + (1) f_6 \\
&= 2f_5 + 2f_6 - 2f_7 - 2f_8
\end{aligned}$$

Normalization yields:  $2b_{1u} = \frac{1}{2}(f_5 + f_6 - f_7 - f_8)$  .

$$\begin{aligned}
P_{B_{1u}} f_{10} &= (1) f_{10} + (1) f_9 + (-1) f_{11} + (-1) f_{12} + \\
&\quad (-1) f_{12} + (-1) f_{11} + (1) f_9 + (1) f_{10} \\
&= 2f_9 + 2f_{10} - 2f_{11} - 2f_{12}
\end{aligned}$$

Normalization yields:  $3b_{1u} = \frac{1}{2}(f_9 + f_{10} - f_{11} - f_{12})$  .

$$\begin{aligned}
P_{B_{2u}} f_6 &= (1) f_6 + (-1) f_5 + (1) f_7 + (-1) f_8 + \\
&\quad (-1) f_8 + (1) f_7 + (-1) f_5 + (1) f_6
\end{aligned}$$

$$= -2f_5 + 2f_6 + 2f_7 - 2f_8$$

Normalization yields:  $1b_{2u} = \frac{1}{2}(-f_5 + f_6 + f_7 - f_8)$  .

$$\begin{aligned} P_{B_{2u}} f_{10} &= (1) f_{10} + (-1) f_9 + (1) f_{11} + (-1) f_{12} + \\ &\quad (-1) f_{12} + (1) f_{11} + (-1) f_9 + (1) f_{10} \\ &= -2f_9 + 2f_{10} + 2f_{11} - 2f_{12} \end{aligned}$$

Normalization yields:  $2b_{2u} = \frac{1}{2}(-f_9 + f_{10} + f_{11} - f_{12})$  .

Each of these four 2x2 symmetry blocks generate identical Hamiltonian matrices. This will be demonstrated for the  $B_{3g}$  symmetry, the others proceed analogously:

$$\begin{aligned} H_{1b_{3g},1b_{3g}} &= \frac{1}{2}(-f_5 + f_6 - f_7 + f_8)H\frac{1}{2}(-f_5 + f_6 - f_7 + f_8)d \\ &= \frac{1}{4} \left\{ \begin{aligned} &f_5Hf_{5d} - f_5Hf_{6d} + f_5Hf_{7d} - f_5Hf_{8d} - \\ &f_6Hf_{5d} + f_6Hf_{6d} - f_6Hf_{7d} + f_6Hf_{8d} + \\ &f_7Hf_{5d} - f_7Hf_{6d} + f_7Hf_{7d} - f_7Hf_{8d} - \\ &f_8Hf_{5d} + f_8Hf_{6d} - f_8Hf_{7d} + f_8Hf_{8d} \end{aligned} \right\} \\ &= \frac{1}{4} \left\{ \begin{aligned} &sp^2 - 0 + 0 - 0 - \\ &0 + sp^2 - 0 + 0 + \\ &0 - 0 + sp^2 - 0 - \\ &0 + 0 - 0 + sp^2 \end{aligned} \right\} = sp^2 \end{aligned}$$

$$\begin{aligned} H_{1b_{3g},2b_{3g}} &= \frac{1}{2}(-f_5 + f_6 - f_7 + f_8)H\frac{1}{2}(-f_9 + f_{10} - f_{11} + f_{12})d \\ &= \frac{1}{4} \left\{ \begin{aligned} &f_5Hf_{9d} - f_5Hf_{10d} + f_5Hf_{11d} - f_5Hf_{12d} - \\ &f_6Hf_{9d} + f_6Hf_{10d} - f_6Hf_{11d} + f_6Hf_{12d} + \\ &f_7Hf_{9d} - f_7Hf_{10d} + f_7Hf_{11d} - f_7Hf_{12d} - \\ &f_8Hf_{9d} + f_8Hf_{10d} - f_8Hf_{11d} + f_8Hf_{12d} \end{aligned} \right\} \\ &= \frac{1}{4} \left\{ \begin{aligned} &sp^{2-s} - 0 + 0 - 0 - \\ &0 + sp^{2-s} - 0 + 0 + \\ &0 - 0 + sp^{2-s} - 0 - \\ &0 + 0 - 0 + sp^{2-s} \end{aligned} \right\} = sp^{2-s} \end{aligned}$$

$$H_{2b_{3g},2b_{3g}} = \frac{1}{2}(-f_9 + f_{10} - f_{11} + f_{12})H\frac{1}{2}(-f_9 + f_{10} - f_{11} + f_{12})d$$

$$\begin{aligned}
&= \frac{1}{4} \left\{ \begin{aligned}
&f_9 H_{f9d} - f_9 H_{f10d} + f_9 H_{f11d} - f_9 H_{f12d} - \\
&f_{10} H_{f9d} + f_{10} H_{f10d} - f_{10} H_{f11d} + f_{10} H_{f12d} + \\
&f_{11} H_{f9d} - f_{11} H_{f10d} + f_{11} H_{f11d} - f_{11} H_{f12d} - \\
&f_{12} H_{f9d} + f_{12} H_{f10d} - f_{12} H_{f11d} + f_{12} H_{f12d} \end{aligned} \right\} \\
&= \frac{1}{4} \left\{ \begin{aligned}
&s - 0 + 0 - 0 - \\
&0 + s - 0 + 0 + \\
&0 - 0 + s - 0 - \\
&0 + 0 - 0 + s \end{aligned} \right\} = s
\end{aligned}$$

This matrix eigenvalue problem then becomes:

$$\begin{array}{cc}
sp^2 & - \\
sp^2-s & \\
\hline
sp^2-s & s \\
-14.7 & -4.0 \\
\hline
-4.0 & -13.6
\end{array} = 0$$

Solving this yields eigenvalues of:

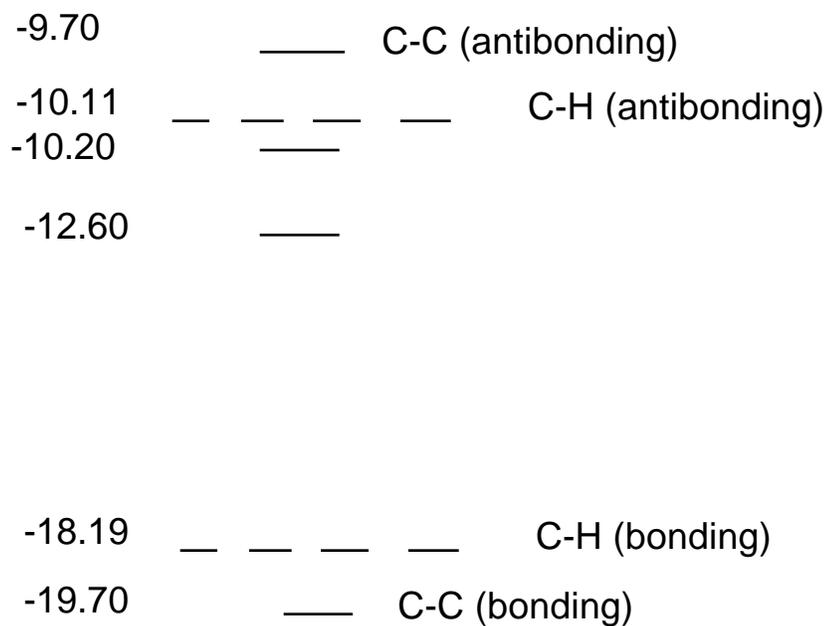
$$| -18.19 \quad -10.11 |$$

and corresponding eigenvectors:

$$-0.7537 \quad -0.6572$$

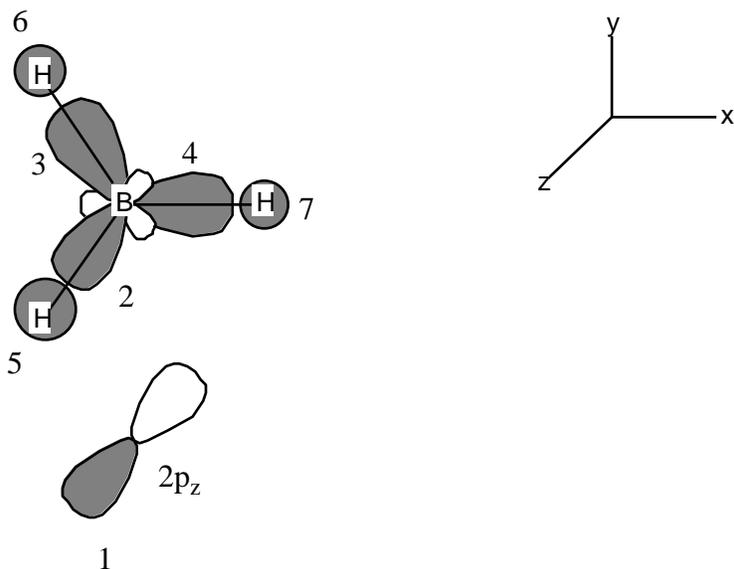
$$-0.6572 \quad 0.7537$$

This results in an orbital energy diagram:



For the ground state of ethylene you would fill the bottom 3 levels (the C-C, C-H, and bonding orbitals), with 12 electrons.

4.



Using the hybrid atomic orbitals as labeled above (functions  $f_1$ - $f_7$ ) and the  $D_{3h}$  point group symmetry it is easiest to construct three sets of reducible representations:

- i. the B  $2p_z$  orbital (labeled function 1)
  - ii. the 3 B  $sp^2$  hybrids (labeled functions 2 - 4)
  - iii. the 3 H  $1s$  orbitals (labeled functions 5 - 7).
- i. The B  $2p_z$  orbital generates the following irreducible representation:

$D_{3h}$	E	$2C_3$	$3C_2$	h	$2S_3$	3	v
$2p_z$	1	1	-1	-1	-1	1	

This irreducible representation is  $A_2''$  and is its own SALC-AO.

ii. The B sp<sup>2</sup> orbitals generate the following reducible representation:

$$\begin{array}{ccccccc} D_{3h} & E & 2C_3 & 3C_2 & \sigma_h & 2S_3 & 3\sigma_v \\ sp^2 & 3 & 0 & 1 & 3 & 0 & 1 \end{array}$$

This reducible representation reduces to 1A<sub>1</sub>' and 1E' irreducible representations.

Projectors are used to find the SALC-AOs for these irreducible representations.

Define: C<sub>3</sub> = 120 degree rotation, C<sub>3</sub>' = 240 degree rotation,

C<sub>2</sub> = rotation around f<sub>4</sub>, C<sub>2</sub>' = rotation around f<sub>2</sub>, and

C<sub>2</sub>'' = rotation around f<sub>3</sub>. S<sub>3</sub> and S<sub>3</sub>' are defined analogous to C<sub>3</sub> and C<sub>3</sub>' with accompanying horizontal reflection.

σ<sub>v</sub> = a reflection plane through f<sub>4</sub>, σ<sub>v</sub>' = a reflection plane

through f<sub>2</sub>, and σ<sub>v</sub>'' = a reflection plane through f<sub>3</sub>

$$\begin{aligned} P_{A_1'} f_2 &= (1)E f_2 + (1)C_3 f_2 + (1)C_3' f_2 + \\ & (1)C_2 f_2 + (1)C_2' f_2 + (1)C_2'' f_2 + \\ & (1)\sigma_h f_2 + (1)S_3 f_2 + (1)S_3' f_2 \\ & (1)\sigma_v f_2 + (1)\sigma_v' f_2 + (1)\sigma_v'' f_2 \\ &= (1)f_2 + (1)f_3 + (1)f_4 + \\ & (1)f_3 + (1)f_2 + (1)f_4 + \\ & (1)f_2 + (1)f_3 + (1)f_4 + \\ & (1)f_3 + (1)f_2 + (1)f_4 \\ &= 4f_2 + 4f_3 + 4f_4 \end{aligned}$$

Normalization yields:  $1a_1' = \frac{1}{\sqrt{3}}(f_2 + f_3 + f_4)$  .

$$\begin{aligned} P_{E'} f_2 &= (2)E f_2 + (-1)C_3 f_2 + (-1)C_3' f_2 + \\ & (0)C_2 f_2 + (0)C_2' f_2 + (0)C_2'' f_2 + \\ & (2)\sigma_h f_2 + (-1)S_3 f_2 + (-1)S_3' f_2 \\ & (0)\sigma_v f_2 + (0)\sigma_v' f_2 + (0)\sigma_v'' f_2 \\ &= (2)f_2 + (-1)f_3 + (-1)f_4 + \\ & (2)f_2 + (-1)f_3 + (-1)f_4 + \\ &= 4f_2 - 2f_3 - 2f_4 \end{aligned}$$

Normalization yields:  $1e' = \frac{1}{\sqrt{6}}(2f_2 - f_3 - f_4)$  .

To find the second e' (orthogonal to the first), projection on f<sub>3</sub> yields (2f<sub>3</sub> - f<sub>2</sub> - f<sub>4</sub>) and projection on f<sub>4</sub> yields (2f<sub>4</sub> - f<sub>2</sub> - f<sub>3</sub>). Neither of these functions are orthogonal to the first, but a combination of the two (2f<sub>3</sub> - f<sub>2</sub> - f<sub>4</sub>) - (2f<sub>4</sub> - f<sub>2</sub> - f<sub>3</sub>) yields a function which is orthogonal to the first.

Normalization yields:  $2e' = \frac{1}{\sqrt{2}}(f_3 - f_4)$  .

iii. The H 1s orbitals generate the following reducible representation:

$$\begin{array}{ccccccc} D_{3h} & E & 2C_3 & 3C_2 & \sigma_h & 2S_3 & 3\sigma_v \\ sp^2 & 3 & 0 & 1 & 3 & 0 & 1 \end{array}$$

This reducible representation reduces to 1A<sub>1</sub>' and 1E'

irreducible representations. exactly like part ii. and in addition the projectors used to find the SALC-AOs for these irreducible representations is exactly analogous to part ii.

$$2a_1' = \frac{1}{\sqrt{3}}(f_5 + f_6 + f_7)$$

$$3e' = \frac{1}{\sqrt{6}}(2f_5 - f_6 - f_7) .$$

$$4e' = \frac{1}{\sqrt{2}}(f_6 - f_7) .$$

So, there are  $1A_2''$ ,  $2A_1'$  and  $2E'$  orbitals. Solving the Hamiltonian matrix for each symmetry block yields:

$A_2''$  Block:

$$\begin{aligned} H_{1a_2',1a_2'} &= f_1 H f_1 d \\ &= 2p = -8.5 \end{aligned}$$

$A_1'$  Block:

$$\begin{aligned} H_{1a_1',1a_1'} &= \frac{1}{\sqrt{3}}(f_2 + f_3 + f_4) H \frac{1}{\sqrt{3}}(f_2 + f_3 + f_4) d \\ &= \frac{1}{3} \left\{ f_2 H f_2 d + f_2 H f_3 d + f_2 H f_4 d + \right. \\ &\quad \left. f_3 H f_2 d + f_3 H f_3 d + f_3 H f_4 d + \right. \\ &\quad \left. f_4 H f_2 d + f_4 H f_3 d + f_4 H f_4 d \right\} \end{aligned}$$

$$= \frac{1}{3} \left\{ \begin{array}{l} sp^2 + 0 + 0 + \\ 0 + sp^2 + 0 + \\ 0 + 0 + sp^2 \end{array} \right\} = sp^2$$

$$\begin{aligned} H_{1a_1',2a_1'} &= \frac{1}{\sqrt{3}}(f_2 + f_3 + f_4) H \frac{1}{\sqrt{3}}(f_5 + f_6 + f_7) d \\ &= \frac{1}{3} \left\{ f_2 H f_5 d + f_2 H f_6 d + f_2 H f_7 d + \right. \\ &\quad \left. f_3 H f_5 d + f_3 H f_6 d + f_3 H f_7 d + \right. \\ &\quad \left. f_4 H f_5 d + f_4 H f_6 d + f_4 H f_7 d \right\} \end{aligned}$$

$$= \frac{1}{3} \left\{ \begin{array}{l} sp^{2-s} + 0 + 0 + \\ 0 + sp^{2-s} + 0 + \\ 0 + 0 + sp^{2-s} \end{array} \right\} = sp^{2-s}$$

$$\begin{aligned} H_{2a_1',2a_1'} &= \frac{1}{\sqrt{3}}(f_5 + f_6 + f_7) H \frac{1}{\sqrt{3}}(f_5 + f_6 + f_7) d \\ &= \frac{1}{3} \left\{ f_5 H f_5 d + f_5 H f_6 d + f_5 H f_7 d + \right. \\ &\quad \left. f_6 H f_5 d + f_6 H f_6 d + f_6 H f_7 d + \right. \\ &\quad \left. f_7 H f_5 d + f_7 H f_6 d + f_7 H f_7 d \right\} \end{aligned}$$

$$\begin{aligned}
 & f_7Hf_{5d} + f_7Hf_{6d} + f_7Hf_{7d} \} \\
 = & \frac{1}{3} \left\{ \begin{array}{l} s + 0 + 0 + \\ 0 + s + 0 + \\ 0 + 0 + s \end{array} \right\} = s
 \end{aligned}$$

This matrix eigenvalue problem then becomes:

$$\begin{array}{cc}
 sp^2 - & sp^2-s \\
 & = 0
 \end{array}$$

$$\begin{array}{cc}
 sp^2-s & s - \\
 -10.7 - & -3.5 \\
 & = 0 \\
 -3.5 & -13.6 -
 \end{array}$$

Solving this yields eigenvalues of:

$$\begin{array}{cc}
 | -15.94 & -8.36 |
 \end{array}$$

and corresponding eigenvectors:

$$\begin{array}{cc}
 -0.5555 & -0.8315 \\
 -0.8315 & 0.5555
 \end{array}$$

E' Block:

This 4x4 symmetry block factors to two 2x2 blocks: where one 2x2 block includes the SALC-AOs

$$e' = \frac{1}{\sqrt{6}}(2f_2 - f_3 - f_4)$$

$$e' = \frac{1}{\sqrt{6}}(2f_5 - f_6 - f_7) ,$$

and the other includes the SALC-AOs

$$e' = \frac{1}{\sqrt{2}}(f_3 - f_4)$$

$$e' = \frac{1}{\sqrt{2}}(f_6 - f_7) .$$

Both of these 2x2 matrices are identical to the  $A_1'$  2x2 array and therefore yield identical energies and MO coefficients.

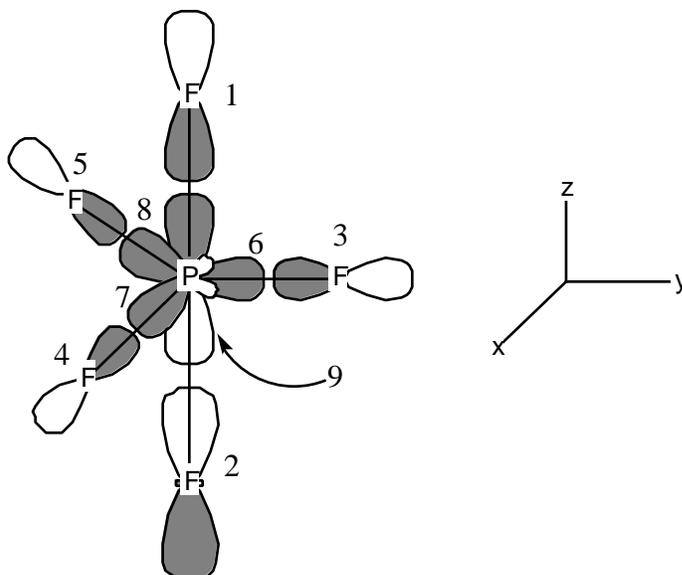
This results in an orbital energy diagram:

$$\begin{array}{cccc}
 -8.36 & \text{---} & \text{---} & \text{---} & a_1', e' \\
 & & -8.5 & \text{---} & a_2''
 \end{array}$$

$$\begin{array}{cccc}
 -15.94 & \text{---} & \text{---} & \text{---} & a_1', e'
 \end{array}$$

For the ground state of  $\text{BH}_3$  you would fill the bottom level (B-H bonding),  $a_1'$  and  $e'$  orbitals, with 6 electrons.

5.



5. a. The two F p orbitals (top and bottom) generate the following reducible representation:

$$\begin{array}{cccccc}
 D_{3h} & E & 2C_3 & 3C_2 & h & 2S_3 & 3 \sigma_v \\
 \rho & 2 & 2 & 0 & 0 & 0 & 2
 \end{array}$$

This reducible representation reduces to  $1A_1'$  and  $1A_2''$  irreducible representations.

Projectors may be used to find the SALC-AOs for these irreducible representations.

$$a_1' = \frac{1}{\sqrt{2}}(f_1 - f_2)$$

$$a_2'' = \frac{1}{\sqrt{2}}(f_1 + f_2)$$

5. b. The three trigonal F p orbitals generate the following reducible representation:

$$\begin{array}{cccccc}
 D_{3h} & E & 2C_3 & 3C_2 & h & 2S_3 & 3 \sigma_v \\
 \rho & 3 & 0 & 1 & 3 & 0 & 1
 \end{array}$$

This reducible representation reduces to  $1A_1'$  and  $1E'$  irreducible representations.

Projectors may be used to find the SALC-AOs for these irreducible representations (but they are exactly analogous to the previous few problems):

$$a_1' = \frac{1}{\sqrt{3}}(f_3 + f_4 + f_5)$$

$$e' = \frac{1}{\sqrt{6}}(2f_3 - f_4 - f_5)$$

$$e' = \frac{1}{\sqrt{2}}(f_4 - f_5) .$$

5. c. The 3 P sp<sup>2</sup> orbitals generate the following reducible representation:

$$\begin{array}{ccccccc} D_{3h} & E & 2C_3 & 3C_2 & \sigma_h & 2S_3 & 3\sigma_v \\ sp^2 & 3 & 0 & 1 & 3 & 0 & 1 \end{array}$$

This reducible representation reduces to 1A<sub>1</sub>' and 1E' irreducible representations. Again, projectors may be used to find the SALC-AOs for these irreducible representations.(but again they are exactly analogous to the previous few problems):

$$a_1' = \frac{1}{\sqrt{3}}(f_6 + f_7 + f_8)$$

$$e' = \frac{1}{\sqrt{6}}(2f_6 - f_7 - f_8)$$

$$e' = \frac{1}{\sqrt{2}}(f_7 - f_8) .$$

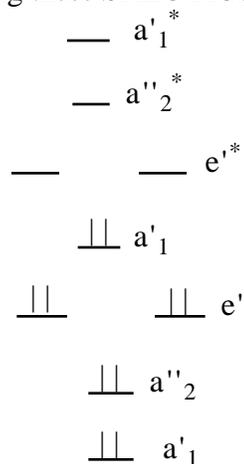
The leftover P p<sub>z</sub> orbital generate the following irreducible representation:

$$\begin{array}{ccccccc} D_{3h} & E & 2C_3 & 3C_2 & \sigma_h & 2S_3 & 3\sigma_v \\ p_z & 1 & 1 & -1 & -1 & -1 & 1 \end{array}$$

This irreducible representation is an A<sub>2</sub>''

$$a_2'' = f_9.$$

Drawing an energy level diagram using these SALC-AOs would result in the following:



## Section 3 Exercises, Problems, and Solutions

### Review Exercises

1. For the given orbital occupations (configurations) of the following systems, determine all possible states (all possible allowed combinations of spin and space states). There is no need to form the determinantal wavefunctions simply label each state with its proper term symbol. One method commonly used is Harry Grays "box method" found in Electrons and Chemical Bonding.

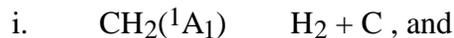
- a.) CH<sub>2</sub>      1a<sub>1</sub><sup>2</sup>2a<sub>1</sub><sup>2</sup>1b<sub>2</sub><sup>2</sup>3a<sub>1</sub><sup>1</sup>1b<sub>1</sub><sup>1</sup>
- b.) B<sub>2</sub> 1 g<sup>2</sup>1 u<sup>2</sup>2 g<sup>2</sup>2 u<sup>2</sup>1 u<sup>1</sup>2 u<sup>1</sup>
- c.) O<sub>2</sub> 1 g<sup>2</sup>1 u<sup>2</sup>2 g<sup>2</sup>2 u<sup>2</sup>1 u<sup>4</sup>3 g<sup>2</sup>1 g<sup>2</sup>
- d.) Ti 1s<sup>2</sup>2s<sup>2</sup>2p<sup>6</sup>3s<sup>2</sup>3p<sup>6</sup>4s<sup>2</sup>3d<sup>1</sup>4d<sup>1</sup>
- e.) Ti 1s<sup>2</sup>2s<sup>2</sup>2p<sup>6</sup>3s<sup>2</sup>3p<sup>6</sup>4s<sup>2</sup>3d<sup>2</sup>

### Exercises

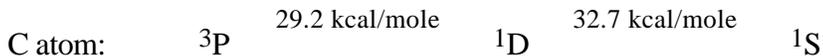
1. Show that the configuration (determinant) corresponding to the Li<sup>+</sup> 1s( )1s( ) state vanishes.
2. Construct the 3 triplet and 1 singlet wavefunctions for the Li<sup>+</sup> 1s<sup>1</sup>2s<sup>1</sup> configuration. Show that each state is a proper eigenfunction of S<sup>2</sup> and S<sub>z</sub> (use raising and lowering operators for S<sup>2</sup>)
3. Construct wavefunctions for each of the following states of CH<sub>2</sub>:
  - a.) <sup>1</sup>B<sub>1</sub> (1a<sub>1</sub><sup>2</sup>2a<sub>1</sub><sup>2</sup>1b<sub>2</sub><sup>2</sup>3a<sub>1</sub><sup>1</sup>1b<sub>1</sub><sup>1</sup>)
  - b.) <sup>3</sup>B<sub>1</sub> (1a<sub>1</sub><sup>2</sup>2a<sub>1</sub><sup>2</sup>1b<sub>2</sub><sup>2</sup>3a<sub>1</sub><sup>1</sup>1b<sub>1</sub><sup>1</sup>)
  - c.) <sup>1</sup>A<sub>1</sub> (1a<sub>1</sub><sup>2</sup>2a<sub>1</sub><sup>2</sup>1b<sub>2</sub><sup>2</sup>3a<sub>1</sub><sup>2</sup>)
4. Construct wavefunctions for each state of the 1<sup>2</sup> 2<sup>2</sup> 3<sup>2</sup> 2<sup>1</sup> 2 configuration of NH.
5. Construct wavefunctions for each state of the 1s<sup>1</sup>2s<sup>1</sup>3s<sup>1</sup> configuration of Li.
6. Determine all term symbols that arise from the 1s<sup>2</sup>2s<sup>2</sup>2p<sup>2</sup>3d<sup>1</sup> configuration of the excited N atom.
7. Calculate the energy (using Slater Condon rules) associated with the 2p valence electrons for the following states of the C atom.
  - i. <sup>3</sup>P(M<sub>L</sub>=1, M<sub>S</sub>=1),
  - ii. <sup>3</sup>P(M<sub>L</sub>=0, M<sub>S</sub>=0),
  - iii. <sup>1</sup>S(M<sub>L</sub>=0, M<sub>S</sub>=0), and
  - iv. <sup>1</sup>D(M<sub>L</sub>=0, M<sub>S</sub>=0).
8. Calculate the energy (using Slater Condon rules) associated with the valence electrons for the following states of the NH molecule.
  - i. <sup>1</sup> (M<sub>L</sub>=2, M<sub>S</sub>=0),
  - ii. <sup>1</sup> (M<sub>L</sub>=0, M<sub>S</sub>=0), and
  - iii. <sup>3</sup> (M<sub>L</sub>=0, M<sub>S</sub>=0).

### Problems

1. Let us investigate the reactions:



under an assumed  $\text{C}_{2v}$  reaction pathway utilizing the following information:



$\text{IP}(\text{H}_2) > \text{IP}(2s \text{ carbon})$ .

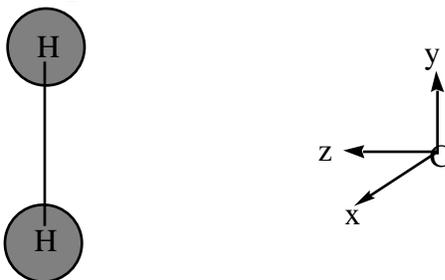
a. Write down (first in terms of  $2p_{1,0,-1}$  orbitals and then in terms of  $2p_{x,y,z}$  orbitals) the:

i. three Slater determinant (SD) wavefunctions belonging to the  $^3\text{P}$  state all of which have  $M_S = 1$ ,

ii. five  $^1\text{D}$  SD wavefunctions, and

iii. one  $^1\text{S}$  SD wavefunction.

b. Using the coordinate system shown below, label the hydrogen orbitals  $g$ ,  $u$  and the carbon  $2s$ ,  $2p_x$ ,  $2p_y$ ,  $2p_z$  orbitals as  $a_1$ ,  $b_1(x)$ ,  $b_2(y)$ , or  $a_2$ . Do the same for the  $s$ ,  $p$ ,  $d$ ,  $f$ ,  $g$ ,  $h$ , and  $i$  orbitals of  $\text{CH}_2$ .



c. Draw an orbital correlation diagram for the  $\text{CH}_2 \rightarrow \text{H}_2 + \text{C}$  reactions. Try to represent the relative energy orderings of the orbitals correctly.

d. Draw (on graph paper) a configuration correlation diagram for  $\text{CH}_2(^3\text{B}_1) \rightarrow \text{H}_2 + \text{C}$  showing all configurations which arise from the  $\text{C}(^3\text{P}) + \text{H}_2$  products. You can assume that doubly excited configurations lie much ( $\sim 100 \text{ kcal/mole}$ ) above their parent configurations.

e. Repeat step d. for  $\text{CH}_2(^1\text{A}_1) \rightarrow \text{H}_2 + \text{C}$  again showing all configurations which arise from the  $\text{C}(^1\text{D}) + \text{H}_2$  products.

f. Do you expect the reaction  $\text{C}(^3\text{P}) + \text{H}_2 \rightarrow \text{CH}_2$  to have a large activation barrier? About how large? What state of  $\text{CH}_2$  is produced in this reaction? Would distortions away from  $\text{C}_{2v}$  symmetry be expected to raise or lower the activation barrier? Show how one could estimate where along the reaction path the barrier top occurs.

g. Would  $\text{C}(^1\text{D}) + \text{H}_2 \rightarrow \text{CH}_2$  be expected to have a larger or smaller barrier than you found for the  $^3\text{P}$  C reaction?

2. The decomposition of the ground-state singlet carbene,



to produce acetylene and  $^1D$  carbon is known to occur with an activation energy equal to the reaction endothermicity. However, when triplet carbene decomposes to acetylene and ground-state (triplet) carbon, the activation energy exceeds this reaction's endothermicity. Construct orbital, configuration, and state correlation diagrams which permit you to explain the above observations. Indicate whether single configuration or configuration interaction wavefunctions would be required to describe the above singlet and triplet decomposition processes.

3. We want to carry out a configuration interaction calculation on  $H_2$  at  $R=1.40$  au. A minimal basis consisting of normalized  $1s$  Slater orbitals with  $\alpha=1.0$  gives rise to the following overlap ( $S$ ), one-electron ( $h$ ), and two-electron atomic integrals:

$$\langle 1s_A | 1s_B \rangle = 0.753 \quad S,$$

$$\langle 1s_A | h | 1s_A \rangle = -1.110, \quad \langle 1s_B | h | 1s_A \rangle = -0.968,$$

$$\langle 1s_A 1s_A | h | 1s_A 1s_A \rangle = 0.625 \quad \langle AA | AA \rangle$$

$$\langle AA | BB \rangle = 0.323, \quad \langle AB | AB \rangle = 0.504, \text{ and}$$

$$\langle AA | AB \rangle = 0.426.$$

a. The normalized and orthogonal molecular orbitals we will use for this minimal basis will be determined purely by symmetry:

$$g = \frac{1}{\sqrt{2+2S}} (1s_A + 1s_B), \text{ and}$$

$$u = \frac{1}{\sqrt{2+2S}} (1s_A - 1s_B).$$

Show that these orbitals are indeed orthogonal.

b. Evaluate (using the one- and two- electron atomic integrals given above) the unique one- and two- electron integrals over this molecular orbital basis (this is called a transformation from the ao to the mo basis). For example, evaluate  $\langle u | h | u \rangle$ ,  $\langle uu | uu \rangle$ ,  $\langle gu | gu \rangle$ , etc.

c. Using the two  $^1g^+$  configurations  $g^2$ , and  $u^2$ , show that the elements of the  $2 \times 2$  configuration interaction Hamiltonian matrix are  $-1.805$ ,  $0.140$ , and  $-0.568$ .

d. Using this configuration interaction matrix, find the configuration interaction (CI) approximation to the ground and excited state energies and wavefunctions.

e. Evaluate and make a rough sketch of the polarized orbitals which result from the above ground state  $g^2$  and  $u^2$  CI wavefunction.

## Solutions

### Review Exercises

1. a. For non-degenerate point groups one can simply multiply the representations (since only one representation will be obtained):

$$a_1 \otimes b_1 = b_1$$

Constructing a "box" in this case is unnecessary since it would only contain a single row. Two unpaired electrons will result in a singlet ( $S=0$ ,  $M_S=0$ ), and three triplets ( $S=1$ ,

$M_S=1$ ;  $S=1$ ,  $M_S=0$ ;  $S=1$ ,  $M_S=-1$ ). The states will be:  $^3B_1(M_S=1)$ ,  $^3B_1(M_S=0)$ ,  $^3B_1(M_S=-1)$ , and  $^1B_1(M_S=0)$ .

1. b. Remember that when coupling non-equivalent linear molecule angular momenta, one simply adds the individual  $L_z$  values and vector couples the electron spin. So, in this case ( $1u^1 2u^1$ ), we have  $M_L$  values of 1+1, 1-1, -1+1, and -1-1 (2, 0, 0, and -2). The term symbol  ${}^1\Delta_g$  is used to denote the spatially doubly degenerate level ( $M_L = \pm 2$ ) and there are two distinct spatially non-degenerate levels denoted by the term symbol  ${}^1\Sigma_g^+$  ( $M_L = 0$ ). Again, two unpaired electrons will result in a singlet ( $S=0, M_S=0$ ), and three triplets ( $S=1, M_S=1; S=1, M_S=0; S=1, M_S=-1$ ). The states generated are then:

- 1 ( $M_L=2$ ); one state ( $M_S=0$ ),
- 1 ( $M_L=-2$ ); one state ( $M_S=0$ ),
- 3 ( $M_L=2$ ); three states ( $M_S=1, 0, \text{ and } -1$ ),
- 3 ( $M_L=-2$ ); three states ( $M_S=1, 0, \text{ and } -1$ ),
- 1 ( $M_L=0$ ); one state ( $M_S=0$ ),
- 1 ( $M_L=0$ ); one state ( $M_S=0$ ),
- 3 ( $M_L=0$ ); three states ( $M_S=1, 0, \text{ and } -1$ ), and
- 3 ( $M_L=0$ ); three states ( $M_S=1, 0, \text{ and } -1$ ).

1. c. Constructing the "box" for two equivalent electrons one obtains:

$M_S$	$M_L$	2	1	0
1				1 -1
0		1 1		1 -1  ,   -1 1

From this "box" one obtains six states:

- 1 ( $M_L=2$ ); one state ( $M_S=0$ ),
- 1 ( $M_L=-2$ ); one state ( $M_S=0$ ),
- 1 ( $M_L=0$ ); one state ( $M_S=0$ ),
- 3 ( $M_L=0$ ); three states ( $M_S=1, 0, \text{ and } -1$ ).

1. d. It is not necessary to construct a "box" when coupling non-equivalent angular momenta since the vector coupling results in a range from the sum of the two individual angular momenta to the absolute value of their difference. In this case,  $3d^1 4d^1$ ,  $L=4, 3, 2, 1, 0$ , and  $S=1, 0$ . The term symbols are:  ${}^3G, {}^1G, {}^3F, {}^1F, {}^3D, {}^1D, {}^3P, {}^1P, {}^3S$ , and  ${}^1S$ . The L and S angular momenta can be vector coupled to produce further splitting into levels:

$$J = L + S \dots |L - S|$$

Denoting J as a term symbol subscript one can identify all the levels and subsequent  $(2J + 1)$  states:

- ${}^3G_5$  (11 states),
- ${}^3G_4$  (9 states),
- ${}^3G_3$  (7 states),

- $^1G_4$  (9 states),
- $^3F_4$  (9 states),
- $^3F_3$  (7 states),
- $^3F_2$  (5 states),
- $^1F_3$  (7 states),
- $^3D_3$  (7 states),
- $^3D_2$  (5 states),
- $^3D_1$  (3 states),
- $^1D_2$  (5 states),
- $^3P_2$  (5 states),
- $^3P_1$  (3 states),
- $^3P_0$  (1 state),
- $^1P_1$  (3 states),
- $^3S_1$  (3 states), and
- $^1S_0$  (1 state).

1. e. Construction of a "box" for the two equivalent d electrons generates (note the "box" has been turned side ways for convenience):

$M_L$	$M_S$	1	0
4			$ d_2 d_2  $
3		$ d_2 d_1  $	$ d_2 d_1  ,$ $ d_2 d_1  $
2		$ d_2 d_0  $	$ d_2 d_0  ,$ $ d_2 d_0  ,$ $ d_1 d_1  $
1		$ d_1 d_0  ,$ $ d_2 d_{-1}  $	$ d_1 d_0  ,$ $ d_1 d_0  ,$ $ d_2 d_{-1}  ,$ $ d_2 d_{-1}  $
0		$ d_2 d_{-2}  ,$ $ d_1 d_{-1}  $	$ d_2 d_{-2}  ,$ $ d_2 d_{-2}  ,$ $ d_1 d_{-1}  ,$ $ d_1 d_{-1}  ,$ $ d_0 d_0  $

The term symbols are:  $^1G$ ,  $^3F$ ,  $^1D$ ,  $^3P$ , and  $^1S$ . The L and S angular momenta can be vector coupled to produce further splitting into levels:

- $^1G_4$  (9 states),
- $^3F_4$  (9 states),
- $^3F_3$  (7 states),

${}^3F_2$  (5 states),  
 ${}^1D_2$  (5 states),  
 ${}^3P_2$  (5 states),  
 ${}^3P_1$  (3 states),  
 ${}^3P_0$  (1 state), and  
 ${}^1S_0$  (1 state).

### Exercises

1. Constructing the Slater determinant corresponding to the "state"  $1s( )1s( )$  with the rows labeling the orbitals and the columns labeling the electron gives:

$$\begin{aligned}
 |1s \ 1s| &= \frac{1}{\sqrt{2!}} \begin{vmatrix} 1s(1) & 1s(2) \\ 1s(1) & 1s(2) \end{vmatrix} \\
 &= \frac{1}{\sqrt{2}} (1s(1)1s(2) - 1s(1)1s(2)) \\
 &= 0
 \end{aligned}$$

2. Starting with the  $M_S=1$   ${}^3S$  state (which in a "box" for this  $M_L=0$ ,  $M_S=1$  case would contain only one product function;  $|1s \ 2s \uparrow \uparrow\rangle$ ) and applying  $S_-$  gives:

$$\begin{aligned}
 S_- {}^3S(S=1, M_S=1) &= \sqrt{1(1+1) - 1(1-1)} \hbar {}^3S(S=1, M_S=0) \\
 &= \hbar\sqrt{2} {}^3S(S=1, M_S=0) \\
 &= (S_-(1) + S_-(2)) |1s \ 2s \uparrow \uparrow\rangle \\
 &= S_-(1)|1s \ 2s \uparrow \uparrow\rangle + S_-(2)|1s \ 2s \uparrow \uparrow\rangle \\
 &= \hbar \sqrt{\frac{1}{2} \frac{1}{2} + 1} |1s \ 2s \uparrow \downarrow\rangle - \hbar \sqrt{\frac{1}{2} \frac{1}{2} - 1} |1s \ 2s \downarrow \uparrow\rangle \\
 &\quad + \hbar \sqrt{\frac{1}{2} \frac{1}{2} + 1} |1s \ 2s \downarrow \downarrow\rangle - \hbar \sqrt{\frac{1}{2} \frac{1}{2} - 1} |1s \ 2s \uparrow \downarrow\rangle \\
 &= \hbar (|1s \ 2s \uparrow \downarrow\rangle + |1s \ 2s \downarrow \uparrow\rangle)
 \end{aligned}$$

So,  $\hbar\sqrt{2} {}^3S(S=1, M_S=0) = \hbar (|1s \ 2s \uparrow \downarrow\rangle + |1s \ 2s \downarrow \uparrow\rangle)$

$${}^3S(S=1, M_S=0) = \frac{1}{\sqrt{2}} (|1s \ 2s \uparrow \downarrow\rangle + |1s \ 2s \downarrow \uparrow\rangle)$$

The three triplet states are then:

$${}^3S(S=1, M_S=1) = |1s \ 2s \uparrow \uparrow\rangle,$$

$${}^3S(S=1, M_S=0) = \frac{1}{\sqrt{2}} (|1s \ 2s \uparrow \downarrow\rangle + |1s \ 2s \downarrow \uparrow\rangle), \text{ and}$$

$${}^3S(S=1, M_S=-1) = |1s \ 2s \downarrow \downarrow\rangle.$$

The singlet state which must be constructed orthogonal to the three singlet states (and in particular to the  $^3S(S=1, M_S=0)$  state) can be seen to be:

$$^1S(S=0, M_S=0) = \frac{1}{\sqrt{2}} (|1s \uparrow 2s \downarrow\rangle - |1s \downarrow 2s \uparrow\rangle).$$

Applying  $S^2$  and  $S_z$  to each of these states gives:

$$\begin{aligned} S_z |1s \uparrow 2s \downarrow\rangle &= (S_z(1) + S_z(2)) |1s \uparrow 2s \downarrow\rangle \\ &= S_z(1)|1s \uparrow 2s \downarrow\rangle + S_z(2)|1s \uparrow 2s \downarrow\rangle \\ &= \hbar \frac{1}{2} |1s \uparrow 2s \downarrow\rangle + \hbar \frac{1}{2} |1s \uparrow 2s \downarrow\rangle \\ &= \hbar |1s \uparrow 2s \downarrow\rangle \end{aligned}$$

$$\begin{aligned} S^2 |1s \uparrow 2s \downarrow\rangle &= (S \cdot S + S_z^2 + \hbar S_z) |1s \uparrow 2s \downarrow\rangle \\ &= S \cdot S |1s \uparrow 2s \downarrow\rangle + S_z^2 |1s \uparrow 2s \downarrow\rangle + \hbar S_z |1s \uparrow 2s \downarrow\rangle \\ &= 0 + \hbar^2 |1s \uparrow 2s \downarrow\rangle + \hbar^2 |1s \uparrow 2s \downarrow\rangle \\ &= 2\hbar^2 |1s \uparrow 2s \downarrow\rangle \end{aligned}$$

$$\begin{aligned} S_z \frac{1}{\sqrt{2}} (|1s \uparrow 2s \downarrow\rangle + |1s \downarrow 2s \uparrow\rangle) &= (S_z(1) + S_z(2)) \frac{1}{\sqrt{2}} (|1s \uparrow 2s \downarrow\rangle + |1s \downarrow 2s \uparrow\rangle) \\ &= \frac{1}{\sqrt{2}} (S_z(1) + S_z(2)) |1s \uparrow 2s \downarrow\rangle \\ &\quad + \frac{1}{\sqrt{2}} (S_z(1) + S_z(2)) |1s \downarrow 2s \uparrow\rangle \\ &= \frac{1}{\sqrt{2}} \hbar \left( \frac{1}{2} - \frac{1}{2} \right) |1s \uparrow 2s \downarrow\rangle \\ &\quad + \frac{1}{\sqrt{2}} \hbar \left( \frac{1}{2} + \frac{1}{2} \right) |1s \downarrow 2s \uparrow\rangle \\ &= 0 \hbar \frac{1}{\sqrt{2}} (|1s \uparrow 2s \downarrow\rangle + |1s \downarrow 2s \uparrow\rangle) \end{aligned}$$

$$\begin{aligned} S^2 \frac{1}{\sqrt{2}} (|1s \uparrow 2s \downarrow\rangle + |1s \downarrow 2s \uparrow\rangle) &= (S \cdot S + S_z^2 + \hbar S_z) \frac{1}{\sqrt{2}} (|1s \uparrow 2s \downarrow\rangle + |1s \downarrow 2s \uparrow\rangle) \\ &= S \cdot S \frac{1}{\sqrt{2}} (|1s \uparrow 2s \downarrow\rangle + |1s \downarrow 2s \uparrow\rangle) \\ &= \frac{1}{\sqrt{2}} (S \cdot (S_+(1) + S_+(2)) |1s \uparrow 2s \downarrow\rangle + S \cdot (S_+(1) + S_+(2)) |1s \downarrow 2s \uparrow\rangle) \\ &= \frac{1}{\sqrt{2}} S \cdot \hbar (|1s \uparrow 2s \downarrow\rangle + |1s \downarrow 2s \uparrow\rangle) \\ &= 2 \hbar \frac{1}{\sqrt{2}} (S \cdot (1) + S \cdot (2)) |1s \uparrow 2s \downarrow\rangle \\ &= 2 \hbar \frac{1}{\sqrt{2}} \hbar (|1s \uparrow 2s \downarrow\rangle + |1s \downarrow 2s \uparrow\rangle) \\ &= 2 \hbar^2 \frac{1}{\sqrt{2}} (|1s \uparrow 2s \downarrow\rangle + |1s \downarrow 2s \uparrow\rangle) \\ S_z |1s \uparrow 2s \downarrow\rangle &= (S_z(1) + S_z(2)) |1s \uparrow 2s \downarrow\rangle \end{aligned}$$

$$\begin{aligned}
&= S_z(1)|1s \ 2s \rangle + S_z(2)|1s \ 2s \rangle \\
&= \hbar \frac{1}{2} |1s \ 2s \rangle + \hbar \frac{1}{2} |1s \ 2s \rangle \\
&= \hbar |1s \ 2s \rangle
\end{aligned}$$

$$\begin{aligned}
S^2 |1s \ 2s \rangle &= (S_+ S_- + S_z^2 - \hbar S_z) |1s \ 2s \rangle \\
&= S_+ S_- |1s \ 2s \rangle + S_z^2 |1s \ 2s \rangle - \hbar S_z |1s \ 2s \rangle \\
&= 0 + \hbar^2 |1s \ 2s \rangle + \hbar^2 |1s \ 2s \rangle \\
&= 2\hbar^2 |1s \ 2s \rangle
\end{aligned}$$

$$\begin{aligned}
S_z \frac{1}{\sqrt{2}} (|1s \ 2s \rangle - |1s \ 2s \rangle) &= (S_z(1) + S_z(2)) \frac{1}{\sqrt{2}} (|1s \ 2s \rangle - |1s \ 2s \rangle) \\
&= \frac{1}{\sqrt{2}} (S_z(1) + S_z(2)) |1s \ 2s \rangle \\
&\quad - \frac{1}{\sqrt{2}} (S_z(1) + S_z(2)) |1s \ 2s \rangle \\
&= \frac{1}{\sqrt{2}} \left( \hbar \frac{1}{2} + \hbar \frac{1}{2} \right) |1s \ 2s \rangle \\
&\quad - \frac{1}{\sqrt{2}} \left( \hbar \frac{1}{2} + \hbar \frac{1}{2} \right) |1s \ 2s \rangle \\
&= 0 \hbar \frac{1}{\sqrt{2}} (|1s \ 2s \rangle - |1s \ 2s \rangle)
\end{aligned}$$

$$\begin{aligned}
S^2 \frac{1}{\sqrt{2}} (|1s \ 2s \rangle - |1s \ 2s \rangle) &= (S_- S_+ + S_z^2 + \hbar S_z) \frac{1}{\sqrt{2}} (|1s \ 2s \rangle - |1s \ 2s \rangle) \\
&= S_- S_+ \frac{1}{\sqrt{2}} (|1s \ 2s \rangle - |1s \ 2s \rangle) \\
&= \frac{1}{\sqrt{2}} (S_- (S_+(1) + S_+(2)) |1s \ 2s \rangle - S_- (S_+(1) + S_+(2)) |1s \ 2s \rangle) \\
&= \frac{1}{\sqrt{2}} S_- \hbar |1s \ 2s \rangle - S_- \hbar |1s \ 2s \rangle \\
&= 0 \hbar \frac{1}{\sqrt{2}} (S_-(1) + S_-(2)) |1s \ 2s \rangle \\
&= 0 \hbar \frac{1}{\sqrt{2}} \hbar |1s \ 2s \rangle - \hbar |1s \ 2s \rangle \\
&= 0 \hbar^2 \frac{1}{\sqrt{2}} (|1s \ 2s \rangle - |1s \ 2s \rangle)
\end{aligned}$$

3. a. Once the spatial symmetry has been determined by multiplication of the irreducible representations, the spin coupling is identical to exercise 2 and gives the result:

$$\frac{1}{\sqrt{2}} (|3a_1 \ 1b_1 \rangle - |3a_1 \ 1b_1 \rangle)$$

3. b. There are three states here (again analogous to exercise 2):

1.)  $|3a_1 \ 1b_1 \rangle$ ,

2.)  $\frac{1}{\sqrt{2}} (|3a_1 \ 1b_1 \rangle + |3a_1 \ 1b_1 \rangle)$ , and

$$3.) |3a_1 \ 1b_1 \rangle$$

$$3. \quad c. |3a_1 \ 3a_1 \rangle$$

4. As shown in review exercise 1c, for two equivalent electrons one obtains six states:

- 1  $(M_L=2)$ ; one state  $(M_S=0)$ ,
- 1  $(M_L=-2)$ ; one state  $(M_S=0)$ ,
- 1  $(M_L=0)$ ; one state  $(M_S=0)$ , and
- 3  $(M_L=0)$ ; three states  $(M_S=1, 0, \text{ and } -1)$ .

By inspecting the "box" in review exercise 1c, it should be fairly straightforward to write down the wavefunctions for each of these:

- 1  $(M_L=2)$ ;  $| \ 1 \ 1 \ \rangle$
- 1  $(M_L=-2)$ ;  $| \ -1 \ -1 \ \rangle$
- 1  $(M_L=0)$ ;  $\frac{1}{\sqrt{2}}(| \ 1 \ -1 \ \rangle - | \ 1 \ -1 \ \rangle)$
- 3  $(M_L=0, M_S=1)$ ;  $| \ 1 \ -1 \ \rangle$
- 3  $(M_L=0, M_S=0)$ ;  $\frac{1}{\sqrt{2}}(| \ 1 \ -1 \ \rangle + | \ 1 \ -1 \ \rangle)$
- 3  $(M_L=0, M_S=-1)$ ;  $| \ 1 \ -1 \ \rangle$

5. We can conveniently couple another s electron to the states generated from the  $1s^1 2s^1$  configuration in exercise 2:

$^3S(L=0, S=1)$  with  $3s^1(L=0, S=\frac{1}{2})$  giving:

$L=0, S=\frac{3}{2}, \frac{1}{2}$ ;  $^4S$  (4 states) and  $^2S$  (2 states).

$^1S(L=0, S=0)$  with  $3s^1(L=0, S=\frac{1}{2})$  giving:

$L=0, S=\frac{1}{2}$ ;  $^2S$  (2 states).

Constructing a "box" for this case would yield:

$M_S$	$M_L$	0
$\frac{3}{2}$		$ 1s \ 2s \ 3s \ \rangle$
$\frac{1}{2}$		$ 1s \ 2s \ 3s \ \rangle,  1s \ 2s \ 3s \ \rangle,  1s \ 2s \ 3s \ \rangle$

One can immediately identify the wavefunctions for two of the quartets (they are single entries):

$$^4S(S=\frac{3}{2}, M_S=\frac{3}{2}): |1s \ 2s \ 3s \ \rangle$$

$${}^4S(S=\frac{3}{2}, M_S=-\frac{3}{2}): |1s \ 2s \ 3s \ |$$

Applying  $S_-$  to  ${}^4S(S=\frac{3}{2}, M_S=\frac{3}{2})$  yields:

$$\begin{aligned} S_- {}^4S(S=\frac{3}{2}, M_S=\frac{3}{2}) &= h \sqrt{\frac{3}{2}(\frac{3}{2} + 1) - \frac{3}{2}(\frac{3}{2} - 1)} \quad {}^4S(S=\frac{3}{2}, M_S=\frac{1}{2}) \\ &= h \sqrt{3} \quad {}^4S(S=\frac{3}{2}, M_S=\frac{1}{2}) \end{aligned}$$

$$S_- |1s \ 2s \ 3s \ | = h (|1s \ 2s \ 3s \ | + |1s \ 2s \ 3s \ | + |1s \ 2s \ 3s \ |)$$

$$\text{So, } {}^4S(S=\frac{3}{2}, M_S=\frac{1}{2}) = \frac{1}{\sqrt{3}} (|1s \ 2s \ 3s \ | + |1s \ 2s \ 3s \ | + |1s \ 2s \ 3s \ |)$$

Applying  $S_+$  to  ${}^4S(S=\frac{3}{2}, M_S=-\frac{3}{2})$  yields:

$$\begin{aligned} S_+ {}^4S(S=\frac{3}{2}, M_S=-\frac{3}{2}) &= h \sqrt{\frac{3}{2}(\frac{3}{2} + 1) - -\frac{3}{2}(-\frac{3}{2} + 1)} \quad {}^4S(S=\frac{3}{2}, M_S=-\frac{1}{2}) \\ &= h \sqrt{3} \quad {}^4S(S=\frac{3}{2}, M_S=-\frac{1}{2}) \end{aligned}$$

$$S_+ |1s \ 2s \ 3s \ | = h (|1s \ 2s \ 3s \ | + |1s \ 2s \ 3s \ | + |1s \ 2s \ 3s \ |)$$

$$\text{So, } {}^4S(S=\frac{3}{2}, M_S=-\frac{1}{2}) = \frac{1}{\sqrt{3}} (|1s \ 2s \ 3s \ | + |1s \ 2s \ 3s \ | + |1s \ 2s \ 3s \ |)$$

It only remains to construct the doublet states which are orthogonal to these quartet states. Recall that the orthogonal combinations for systems having three equal components (for example when symmetry adapting the 3  $sp^2$  hybrids in  $C_{2v}$  or  $D_{3h}$  symmetry) give results of  $+++$ ,  $+2--$ , and  $0+-$ . Notice that the quartets are the  $+++$  combinations and therefore the doublets can be recognized as:

$${}^2S(S=\frac{1}{2}, M_S=\frac{1}{2}) = \frac{1}{\sqrt{6}} (|1s \ 2s \ 3s \ | + |1s \ 2s \ 3s \ | - 2|1s \ 2s \ 3s \ |)$$

$${}^2S(S=\frac{1}{2}, M_S=\frac{1}{2}) = \frac{1}{\sqrt{2}} (|1s \ 2s \ 3s \ | - |1s \ 2s \ 3s \ | + 0|1s \ 2s \ 3s \ |)$$

$${}^2S(S=\frac{1}{2}, M_S=-\frac{1}{2}) = \frac{1}{\sqrt{6}} (|1s \ 2s \ 3s \ | + |1s \ 2s \ 3s \ | - 2|1s \ 2s \ 3s \ |)$$

$${}^2S(S=\frac{1}{2}, M_S=-\frac{1}{2}) = \frac{1}{\sqrt{3}} (|1s \ 2s \ 3s \ | - |1s \ 2s \ 3s \ | + 0|1s \ 2s \ 3s \ |)$$

6. As illustrated in this chapter a  $p^2$  configuration (two equivalent p electrons) gives rise to the term symbols:  ${}^3P$ ,  ${}^1D$ , and  ${}^1S$ . Coupling an additional electron ( $3d^1$ ) to this  $p^2$  configuration will give the desired  $1s^2 2s^2 2p^2 3d^1$  term symbols:

${}^3P(L=1, S=1)$  with  ${}^2D(L=2, S=\frac{1}{2})$  generates;

$L=3, 2, 1$ , and  $S=\frac{3}{2}, \frac{1}{2}$  with term symbols  ${}^4F, {}^2F, {}^4D, {}^2D, {}^4P$ , and  ${}^2P$ ,

${}^1D(L=2, S=0)$  with  ${}^2D(L=2, S=\frac{1}{2})$  generates;

$L=4, 3, 2, 1, 0$ , and  $S=\frac{1}{2}$  with term symbols  ${}^2G, {}^2F, {}^2D, {}^2P$ , and  ${}^2S$ ,

${}^1S(L=0, S=0)$  with  ${}^2D(L=2, S=\frac{1}{2})$  generates;

$L=2$  and  $S=\frac{1}{2}$  with term symbol  ${}^2D$ .

7. The notation used for the Slater Condon rules will be the same as used in the text:  
 (a.) zero (spin orbital) difference;

$$\begin{aligned} \langle F + G \rangle &= \langle i | f | i \rangle + \langle i j | g | i j \rangle - \langle i j | g | j i \rangle \\ &= f_{ii} + (g_{ijij} - g_{ijji}) \end{aligned}$$

(b.) one (spin orbital) difference (  $p$   $p'$  );

$$\begin{aligned} \langle F + G \rangle &= \langle p | f | p \rangle + \langle p j | g | p' j \rangle - \langle p j | g | j p' \rangle \\ &= f_{pp} + (g_{pj p' j} - g_{p j p' j}) \end{aligned}$$

(c.) two (spin orbital) differences (  $p$   $p'$  and  $q$   $q'$  );

$$\begin{aligned} \langle F + G \rangle &= \langle p q | g | p' q' \rangle - \langle p q | g | q' p' \rangle \\ &= g_{pq p' q'} - g_{pq q' p'} \end{aligned}$$

(d.) three or more (spin orbital) differences;

$$\langle F + G \rangle = 0$$

7. i.  ${}^3P(M_L=1, M_S=1) = |p_1 p_0 |$

$$\langle p_1 p_0 | H | p_1 p_0 \rangle =$$

**Error!**. Using the Slater Condon rule (a.) above (SCa):

$$\langle 10 | H | 10 \rangle = f_{11} + f_{00} + g_{1010} - g_{1001}$$

7. ii.  ${}^3P(M_L=0, M_S=0) = \frac{1}{\sqrt{2}} (|p_1 p_{-1} | + |p_1 p_{-1} |)$

$$\begin{aligned} &\langle {}^3P(M_L=0, M_S=0) | H | {}^3P(M_L=0, M_S=0) \rangle \\ &= \frac{1}{2} ( \langle p_1 p_{-1} | H | p_1 p_{-1} \rangle + \langle p_1 p_{-1} | H | p_1 p_{-1} \rangle \\ &\quad + \langle p_1 p_{-1} | H | p_1 p_{-1} \rangle + \langle p_1 p_{-1} | H | p_1 p_{-1} \rangle ) \end{aligned}$$

Evaluating each matrix element gives:

$$\begin{aligned} \langle p_1 p_{-1} | H | p_1 p_{-1} \rangle &= f_{11} + f_{-1-1} + g_{1-11-1} - g_{1-1-11} \quad (\text{SCa}) \\ &= f_{11} + f_{-1-1} + g_{1-11-1} - 0 \end{aligned}$$

$$\begin{aligned} \langle p_1 p_{-1} | H | p_1 p_{-1} \rangle &= g_{1-11-1} - g_{1-1-11} \quad (\text{SCc}) \\ &= 0 - g_{1-1-11} \end{aligned}$$

$$\begin{aligned} \langle p_1 p_{-1} | H | p_1 p_{-1} \rangle &= g_{1-11-1} - g_{1-1-11} \quad (\text{SCc}) \\ &= 0 - g_{1-1-11} \end{aligned}$$

$$\begin{aligned} \langle p_1 p_{-1} | H | p_1 p_{-1} \rangle &= f_{11} + f_{-1-1} + g_{1-11-1} - g_{1-1-11} \quad (\text{SCa}) \\ &= f_{11} + f_{-1-1} + g_{1-11-1} - 0 \end{aligned}$$

Substitution of these expressions give:

$$\langle {}^3P(M_L=0, M_S=0) | H | {}^3P(M_L=0, M_S=0) \rangle$$

$$= \frac{1}{2} (f_{11} + f_{-1-1} + g_{1-11-1} - g_{1-1-11} - g_{1-1-11} \\ + f_{11} + f_{-1-1} + g_{1-11-1}) \\ = f_{11} + f_{-1-1} + g_{1-11-1} - g_{1-1-11}$$

7. iii.  ${}^1S(M_L=0, M_S=0); \frac{1}{\sqrt{3}}(|p_0 p_0 \rangle - |p_1 p_{-1} \rangle - |p_{-1} p_1 \rangle)$

$$\langle {}^1S(M_L=0, M_S=0) | H | {}^1S(M_L=0, M_S=0) \rangle$$

$$= \frac{1}{3} (\langle p_0 p_0 | H | p_0 p_0 \rangle - \langle p_0 p_0 | H | p_1 p_{-1} \rangle \\ - \langle p_0 p_0 | H | p_{-1} p_1 \rangle - \langle p_1 p_{-1} | H | p_0 p_0 \rangle \\ + \langle p_1 p_{-1} | H | p_1 p_{-1} \rangle + \langle p_1 p_{-1} | H | p_{-1} p_1 \rangle \\ - \langle p_{-1} p_1 | H | p_0 p_0 \rangle + \langle p_{-1} p_1 | H | p_1 p_{-1} \rangle \\ + \langle p_{-1} p_1 | H | p_{-1} p_1 \rangle )$$

Evaluating each matrix element gives:

$$\langle p_0 p_0 | H | p_0 p_0 \rangle = f_{00} + f_{00} + g_{0000} - g_{0000} \quad (\text{SCa}) \\ = f_{00} + f_{00} + g_{0000} - 0$$

$$\langle p_0 p_0 | H | p_1 p_{-1} \rangle = \langle p_1 p_{-1} | H | p_0 p_0 \rangle \\ = g_{001-1} - g_{00-11} \quad (\text{SCc}) \\ = g_{001-1} - 0$$

$$\langle p_0 p_0 | H | p_{-1} p_1 \rangle = \langle p_{-1} p_1 | H | p_0 p_0 \rangle \\ = g_{00-11} - g_{001-1} \quad (\text{SCc}) \\ = g_{00-11} - 0$$

$$\langle p_1 p_{-1} | H | p_1 p_{-1} \rangle = f_{11} + f_{-1-1} + g_{1-11-1} - g_{1-1-11} \quad (\text{SCa}) \\ = f_{11} + f_{-1-1} + g_{1-11-1} - 0$$

$$\langle p_1 p_{-1} | H | p_{-1} p_1 \rangle = \langle p_{-1} p_1 | H | p_1 p_{-1} \rangle \\ = g_{1-1-11} - g_{1-11-1} \quad (\text{SCc}) \\ = g_{1-1-11} - 0$$

$$\langle p_{-1} p_1 | H | p_{-1} p_1 \rangle = f_{-1-1} + f_{11} + g_{-11-11} - g_{-11-11} \quad (\text{SCa}) \\ = f_{-1-1} + f_{11} + g_{-11-11} - 0$$

Substitution of these expressions give:

$$\langle {}^1S(M_L=0, M_S=0) | H | {}^1S(M_L=0, M_S=0) \rangle \\ = \frac{1}{3} (f_{00} + f_{00} + g_{0000} - g_{001-1} - g_{00-11} - g_{001-1} + f_{11} + f_{-1-1} \\ + g_{1-11-1} + g_{1-1-11} - g_{00-11} + g_{1-1-11} + f_{-1-1} + f_{11} + g_{-11-11}) \\ = \frac{1}{3} (2f_{00} + 2f_{11} + 2f_{-1-1} + g_{0000} - 4g_{001-1} + 2g_{1-11-1} + 2g_{1-1-11})$$

$$7. \quad \text{iv. } |^1D(M_L=0, M_S=0)\rangle = \frac{1}{\sqrt{6}}(2|p_0 \ p_0 \rangle + |p_1 \ p_{-1} \rangle + |p_{-1} \ p_1 \rangle)$$

Evaluating  $\langle ^1D(M_L=0, M_S=0) | H | ^1D(M_L=0, M_S=0) \rangle$  we note that all the Slater Condon matrix elements generated are the same as those evaluated in part iii. (the signs for the wavefunction components and the multiplicative factor of two for one of the components, however, are different).

$$\begin{aligned} & \langle ^1D(M_L=0, M_S=0) | H | ^1D(M_L=0, M_S=0) \rangle \\ &= \frac{1}{6} (4f_{00} + 4f_{00} + 4g_{0000} + 2g_{001-1} + 2g_{00-11} + 2g_{001-1} + f_{11} \\ & \quad + f_{-1-1} + g_{1-11-1} + g_{1-1-11} + 2g_{00-11} + g_{1-1-11} + f_{-1-1} + f_{11} \\ & \quad + g_{-11-11}) \\ &= \frac{1}{6} (8f_{00} + 2f_{11} + 2f_{-1-1} + 4g_{0000} + 8g_{001-1} + 2g_{1-11-1} + 2g_{1-1-11}) \end{aligned}$$

$$8. \quad \text{i. } |^1(M_L=2, M_S=0)\rangle = | \ 1 \ 1 \ 1 \ 1 \rangle$$

$$\begin{aligned} & \langle ^1(M_L=2, M_S=0) | H | ^1(M_L=2, M_S=0) \rangle \\ &= \langle | \ 1 \ 1 \ 1 \ 1 \rangle | H | | \ 1 \ 1 \ 1 \ 1 \rangle \\ &= f_{1 \ 1 \ 1 \ 1} + f_{1 \ 1 \ 1 \ 1} + g_{1 \ 1 \ 1 \ 1} - g_{1 \ 1 \ 1 \ 1} \quad (\text{SCa}) \\ &= f_{11} + f_{11} + g_{1111} - 0 \\ &= 2f_{11} + g_{1111} \end{aligned}$$

$$8. \quad \text{ii. } |^1(M_L=0, M_S=0)\rangle = \frac{1}{\sqrt{2}}(| \ 1 \ 1 \ -1 \ -1 \rangle - | \ 1 \ -1 \ 1 \ 1 \rangle)$$

$$\begin{aligned} & \langle ^3(M_L=0, M_S=0) | H | ^3(M_L=0, M_S=0) \rangle \\ &= \frac{1}{2} (\langle | \ 1 \ 1 \ -1 \ -1 \rangle | H | | \ 1 \ 1 \ -1 \ -1 \rangle - \langle | \ 1 \ -1 \ 1 \ 1 \rangle | H | | \ 1 \ -1 \ 1 \ 1 \rangle \\ & \quad - \langle | \ 1 \ -1 \ 1 \ 1 \rangle | H | | \ 1 \ 1 \ -1 \ -1 \rangle + \langle | \ 1 \ 1 \ -1 \ -1 \rangle | H | | \ 1 \ 1 \ -1 \ -1 \rangle) \end{aligned}$$

Evaluating each matrix element gives:

$$\begin{aligned} \langle | \ 1 \ 1 \ -1 \ -1 \rangle | H | | \ 1 \ 1 \ -1 \ -1 \rangle &= f_{1 \ 1 \ -1 \ -1} + f_{-1 \ -1 \ 1 \ 1} + g_{1 \ -1 \ 1 \ -1} - g_{1 \ -1 \ -1 \ 1} \quad (\text{SCa}) \\ &= f_{11} + f_{-1-1} + g_{1-11-1} - 0 \end{aligned}$$

$$\begin{aligned} \langle | \ 1 \ -1 \ 1 \ 1 \rangle | H | | \ 1 \ -1 \ 1 \ 1 \rangle &= g_{1 \ -1 \ 1 \ -1} - g_{1 \ -1 \ -1 \ 1} \quad (\text{SCc}) \\ &= 0 - g_{1-1-11} \end{aligned}$$

$$\begin{aligned} \langle | \ 1 \ -1 \ 1 \ 1 \rangle | H | | \ 1 \ 1 \ -1 \ -1 \rangle &= g_{1 \ -1 \ 1 \ -1} - g_{1 \ -1 \ -1 \ 1} \quad (\text{SCc}) \\ &= 0 - g_{1-1-11} \end{aligned}$$

$$\begin{aligned} \langle | \ 1 \ -1 \ 1 \ 1 \rangle | H | | \ 1 \ -1 \ 1 \ 1 \rangle &= f_{1 \ 1 \ 1 \ 1} + f_{-1 \ -1 \ 1 \ 1} + g_{1 \ -1 \ 1 \ -1} - g_{1 \ -1 \ -1 \ 1} \quad (\text{SCa}) \\ &= f_{11} + f_{-1-1} + g_{1-11-1} - 0 \end{aligned}$$

Substitution of these expressions give:

$$\begin{aligned} & \langle ^3(M_L=0, M_S=0) | H | ^3(M_L=0, M_S=0) \rangle \\ &= \frac{1}{2} (f_{11} + f_{-1-1} + g_{1-11-1} + g_{1-1-11} + g_{1-1-11} + f_{11} + f_{-1-1} + g_{1-11-1}) \\ &= f_{11} + f_{-1-1} + g_{1-11-1} + g_{1-1-11} \end{aligned}$$

$$\begin{aligned}
8. \quad \text{iii. } {}^3(M_L=0, M_S=0) &= \frac{1}{\sqrt{2}}(|1 \ 1 \ -1 \rangle + |1 \ -1 \ 1 \rangle) \\
\langle {}^3(M_L=0, M_S=0) | H | {}^3(M_L=0, M_S=0) \rangle & \\
&= \frac{1}{2} (\langle 1 \ 1 \ -1 | H | 1 \ 1 \ -1 \rangle + \langle 1 \ 1 \ -1 | H | 1 \ -1 \ 1 \rangle \\
&\quad + \langle 1 \ -1 \ 1 | H | 1 \ 1 \ -1 \rangle + \langle 1 \ -1 \ 1 | H | 1 \ -1 \ 1 \rangle )
\end{aligned}$$

Evaluating each matrix element gives:

$$\begin{aligned}
\langle 1 \ 1 \ -1 | H | 1 \ 1 \ -1 \rangle &= f_{1 \ 1} + f_{-1 \ -1} + g_{1 \ -1 \ 1 \ -1} - g_{-1 \ -1 \ 1 \ 1} \quad (\text{SCa}) \\
&= f_{11} + f_{-1-1} + g_{1-11-1} - 0
\end{aligned}$$

$$\begin{aligned}
\langle 1 \ 1 \ -1 | H | 1 \ -1 \ 1 \rangle &= g_{1 \ -1 \ 1 \ -1} - g_{-1 \ -1 \ 1 \ 1} \quad (\text{SCc}) \\
&= 0 - g_{1-1-11}
\end{aligned}$$

$$\begin{aligned}
\langle 1 \ -1 \ 1 | H | 1 \ 1 \ -1 \rangle &= g_{1 \ -1 \ 1 \ -1} - g_{-1 \ -1 \ 1 \ 1} \quad (\text{SCc}) \\
&= 0 - g_{1-1-11}
\end{aligned}$$

$$\begin{aligned}
\langle 1 \ -1 \ 1 | H | 1 \ -1 \ 1 \rangle &= f_{1 \ 1} + f_{-1 \ -1} + g_{1 \ -1 \ 1 \ -1} - g_{-1 \ -1 \ 1 \ 1} \quad (\text{SCa}) \\
&= f_{11} + f_{-1-1} + g_{1-11-1} - 0
\end{aligned}$$

Substitution of these expressions give:

$$\begin{aligned}
\langle {}^3(M_L=0, M_S=0) | H | {}^3(M_L=0, M_S=0) \rangle & \\
&= \frac{1}{2} (f_{11} + f_{-1-1} + g_{1-11-1} - g_{1-1-11} - g_{1-1-11} + f_{11} + f_{-1-1} + g_{1-11-1}) \\
&= f_{11} + f_{-1-1} + g_{1-11-1} - g_{1-1-11}
\end{aligned}$$

### Problems

1. a. All the Slater determinants have in common the  $|1s \ 1s \ 2s \ 2s \rangle$  "core" and hence this component will not be written out explicitly for each case.

$$\begin{aligned}
{}^3P(M_L=1, M_S=1) &= |p_1 \ p_0 \ | \\
&= \left| \frac{1}{\sqrt{2}}(p_x + ip_y) \ (p_z) \ | \\
&= \frac{1}{\sqrt{2}}(|p_x \ p_z \ | + i|p_y \ p_z \ |)
\end{aligned}$$

$$\begin{aligned}
{}^3P(M_L=0, M_S=1) &= |p_1 \ p_{-1} \ | \\
&= \left| \frac{1}{\sqrt{2}}(p_x + ip_y) \ \frac{1}{\sqrt{2}}(p_x - ip_y) \ | \\
&= \frac{1}{2}(|p_x \ p_x \ | - i|p_x \ p_y \ | + i|p_y \ p_x \ | + |p_y \ p_y \ |) \\
&= \frac{1}{2}(0 - i|p_x \ p_y \ | - i|p_x \ p_y \ | + 0) \\
&= \frac{1}{2}(-2i|p_x \ p_y \ |) \\
&= -i|p_x \ p_y \ |
\end{aligned}$$

$${}^3P(M_L=-1, M_S=1) = |p_{-1} \ p_0 \ |$$

$$\begin{aligned}
&= \frac{1}{\sqrt{2}}(p_x - ip_y) (p_z) \quad | \\
&= \frac{1}{\sqrt{2}}(|p_x \ p_z \rangle - i|p_y \ p_z \rangle)
\end{aligned}$$

As you can see, the symmetries of each of these states cannot be labeled with a single irreducible representation of the  $C_{2v}$  point group. For example,  $|p_x \ p_z \rangle$  is  $xz$  ( $B_1$ ) and  $|p_y \ p_z \rangle$  is  $yz$  ( $B_2$ ) and hence the  ${}^3P(M_L=1, M_S=1)$  state is a combination of  $B_1$  and  $B_2$  symmetries. But, the three  ${}^3P(M_L, M_S=1)$  functions are degenerate for the C atom and any combination of these three functions would also be degenerate. Therefore we can choose new combinations which can be labeled with "pure"  $C_{2v}$  point group labels.

$$\begin{aligned}
{}^3P(xz, M_S=1) &= |p_x \ p_z \rangle \\
&= \frac{1}{\sqrt{2}}({}^3P(M_L=1, M_S=1) + {}^3P(M_L=-1, M_S=1)) = {}^3B_1
\end{aligned}$$

$$\begin{aligned}
{}^3P(yx, M_S=1) &= |p_y \ p_x \rangle \\
&= \frac{1}{i}({}^3P(M_L=0, M_S=1)) = {}^3A_2
\end{aligned}$$

$$\begin{aligned}
{}^3P(yz, M_S=1) &= |p_y \ p_z \rangle \\
&= \frac{1}{i\sqrt{2}}({}^3P(M_L=1, M_S=1) - {}^3P(M_L=-1, M_S=1)) = {}^3B_2
\end{aligned}$$

Now we can do likewise for the five degenerate  ${}^1D$  states:

$$\begin{aligned}
{}^1D(M_L=2, M_S=0) &= |p_1 \ p_1 \rangle \\
&= \frac{1}{\sqrt{2}}(p_x + ip_y) \quad \frac{1}{\sqrt{2}}(p_x + ip_y) \quad | \\
&= \frac{1}{2}(|p_x \ p_x \rangle + i|p_x \ p_y \rangle + i|p_y \ p_x \rangle - |p_y \ p_y \rangle)
\end{aligned}$$

$$\begin{aligned}
{}^1D(M_L=-2, M_S=0) &= |p_{-1} \ p_{-1} \rangle \\
&= \frac{1}{\sqrt{2}}(p_x - ip_y) \quad \frac{1}{\sqrt{2}}(p_x - ip_y) \quad | \\
&= \frac{1}{2}(|p_x \ p_x \rangle - i|p_x \ p_y \rangle - i|p_y \ p_x \rangle - |p_y \ p_y \rangle)
\end{aligned}$$

$$\begin{aligned}
{}^1D(M_L=1, M_S=0) &= \frac{1}{\sqrt{2}}(|p_0 \ p_1 \rangle - |p_0 \ p_{-1} \rangle) \\
&= \frac{1}{\sqrt{2}} |(p_z) \quad \frac{1}{\sqrt{2}}(p_x + ip_y) \rangle - |(p_z) \quad \frac{1}{\sqrt{2}}(p_x + ip_y) \rangle | \\
&= \frac{1}{2}(|p_z \ p_x \rangle + i|p_z \ p_y \rangle - |p_z \ p_x \rangle - i|p_z \ p_y \rangle)
\end{aligned}$$

$$\begin{aligned}
{}^1D(M_L=-1, M_S=0) &= \frac{1}{\sqrt{2}}(|p_0 \ p_{-1} \rangle - |p_0 \ p_1 \rangle) \\
&= \frac{1}{\sqrt{2}} |(p_z) \quad \frac{1}{\sqrt{2}}(p_x - ip_y) \rangle - |(p_z) \quad \frac{1}{\sqrt{2}}(p_x - ip_y) \rangle | \\
&= \frac{1}{2}(|p_z \ p_x \rangle - i|p_z \ p_y \rangle - |p_z \ p_x \rangle + i|p_z \ p_y \rangle)
\end{aligned}$$

$${}^1D(M_L=0, M_S=0) = \frac{1}{\sqrt{6}}(2|p_0 \ p_0 \rangle + |p_1 \ p_{-1} \rangle + |p_{-1} \ p_1 \rangle)$$

$$\begin{aligned}
&= \frac{1}{\sqrt{6}} \left( 2|p_z \ p_z \rangle + \frac{1}{\sqrt{2}}(p_x + ip_y) \frac{1}{\sqrt{2}}(p_x - ip_y) \right. \\
&\quad \left. + \frac{1}{\sqrt{2}}(p_x - ip_y) \frac{1}{\sqrt{2}}(p_x + ip_y) \right) \\
&= \frac{1}{\sqrt{6}} \left( 2|p_z \ p_z \rangle \right. \\
&\quad \left. + \frac{1}{2}(|p_x \ p_x \rangle - i|p_x \ p_y \rangle + i|p_y \ p_x \rangle + |p_y \ p_y \rangle) \right. \\
&\quad \left. + \frac{1}{2}(|p_x \ p_x \rangle + i|p_x \ p_y \rangle - i|p_y \ p_x \rangle + |p_y \ p_y \rangle) \right) \\
&= \frac{1}{\sqrt{6}} \left( 2|p_z \ p_z \rangle + |p_x \ p_x \rangle + |p_y \ p_y \rangle \right)
\end{aligned}$$

Analogous to the three  $^3P$  states we can also choose combinations of the five degenerate  $^1D$  states which can be labeled with "pure"  $C_{2v}$  point group labels:

$$\begin{aligned}
^1D_{(xx-yy, M_S=0)} &= |p_x \ p_x \rangle - |p_y \ p_y \rangle \\
&= (^1D_{(M_L=2, M_S=0)} + ^1D_{(M_L=-2, M_S=0)}) = ^1A_1
\end{aligned}$$

$$\begin{aligned}
^1D_{(yx, M_S=0)} &= |p_x \ p_y \rangle + |p_y \ p_x \rangle \\
&= \frac{1}{i} (^1D_{(M_L=2, M_S=0)} - ^1D_{(M_L=-2, M_S=0)}) = ^1A_2
\end{aligned}$$

$$\begin{aligned}
^1D_{(zx, M_S=0)} &= |p_z \ p_x \rangle - |p_z \ p_x \rangle \\
&= (^1D_{(M_L=1, M_S=0)} + ^1D_{(M_L=-1, M_S=0)}) = ^1B_1
\end{aligned}$$

$$\begin{aligned}
^1D_{(zy, M_S=0)} &= |p_z \ p_y \rangle - |p_z \ p_y \rangle \\
&= \frac{1}{i} (^1D_{(M_L=1, M_S=0)} - ^1D_{(M_L=-1, M_S=0)}) = ^1B_2
\end{aligned}$$

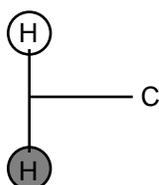
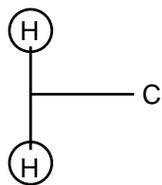
$$\begin{aligned}
^1D_{(2zz+xx+yy, M_S=0)} &= \frac{1}{\sqrt{6}} \left( 2|p_z \ p_z \rangle + |p_x \ p_x \rangle + |p_y \ p_y \rangle \right) \\
&= ^1D_{(M_L=0, M_S=0)} = ^1A_1
\end{aligned}$$

The only state left is the  $^1S$ :

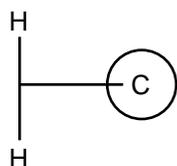
$$\begin{aligned}
^1S_{(M_L=0, M_S=0)} &= \frac{1}{\sqrt{3}} \left( |p_0 \ p_0 \rangle - |p_1 \ p_{-1} \rangle - |p_{-1} \ p_1 \rangle \right) \\
&= \frac{1}{\sqrt{3}} \left( |p_z \ p_z \rangle - \frac{1}{\sqrt{2}}(p_x + ip_y) \frac{1}{\sqrt{2}}(p_x - ip_y) \right. \\
&\quad \left. - \frac{1}{\sqrt{2}}(p_x - ip_y) \frac{1}{\sqrt{2}}(p_x + ip_y) \right) \\
&= \frac{1}{\sqrt{3}} \left( |p_z \ p_z \rangle \right. \\
&\quad \left. - \frac{1}{2}(|p_x \ p_x \rangle - i|p_x \ p_y \rangle + i|p_y \ p_x \rangle + |p_y \ p_y \rangle) \right. \\
&\quad \left. - \frac{1}{2}(|p_x \ p_x \rangle + i|p_x \ p_y \rangle - i|p_y \ p_x \rangle + |p_y \ p_y \rangle) \right) \\
&= \frac{1}{\sqrt{3}} \left( |p_z \ p_z \rangle - |p_x \ p_x \rangle - |p_y \ p_y \rangle \right)
\end{aligned}$$

Each of the components of this state are  $A_1$  and hence this state has  $A_1$  symmetry.

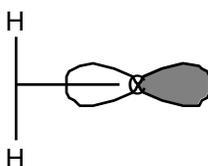
1. b. Forming SALC-AOs from the C and H atomic orbitals would generate the following:



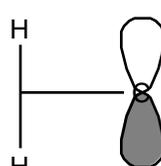
$$H_{1s} + H_{1s} = g = a_1 \quad H_{1s} - H_{1s} = u = b_2$$



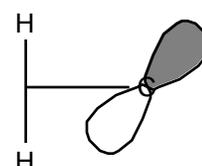
$$C_{2s} = a_1$$



$$C_{2p_z} = a_1$$

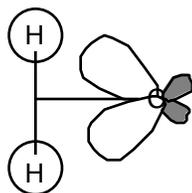


$$C_{2p_y} = b_2$$

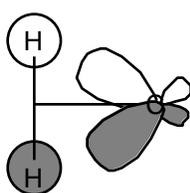


$$C_{2p_x} = b_1$$

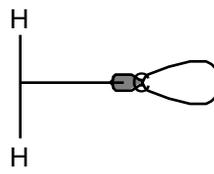
The bonding, nonbonding, and antibonding orbitals of CH<sub>2</sub> can be illustrated in the following manner:



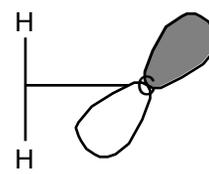
$$= a_1$$



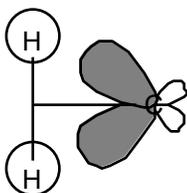
$$= b_2$$



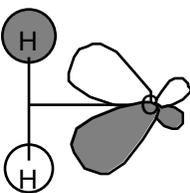
$$n = a_1$$



$$p = b_1$$



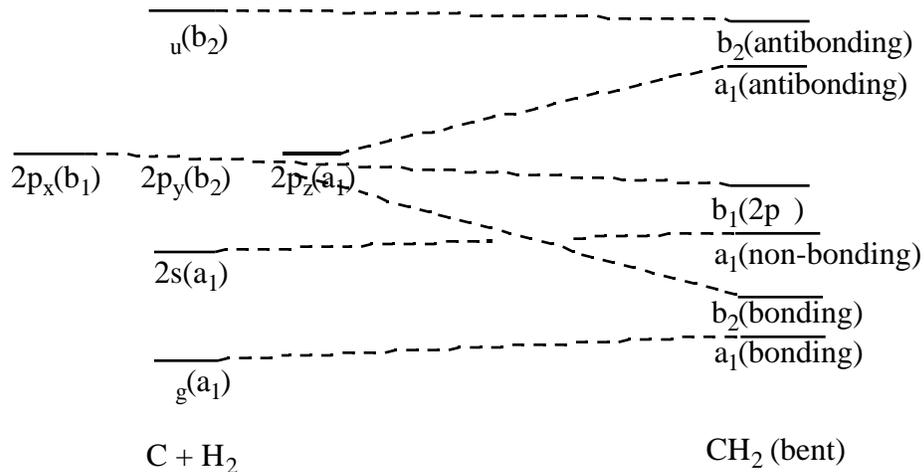
$$* = a_1$$



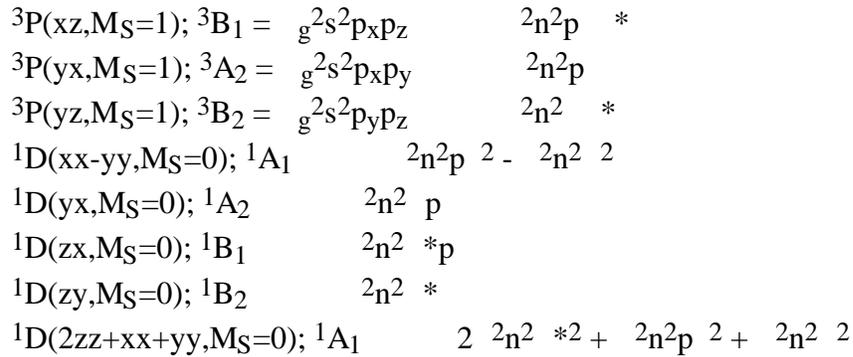
$$* = b_2$$

1. c.

Orbital-correlation diagram for the reaction  $C + H_2 \rightarrow CH_2$  (bent)



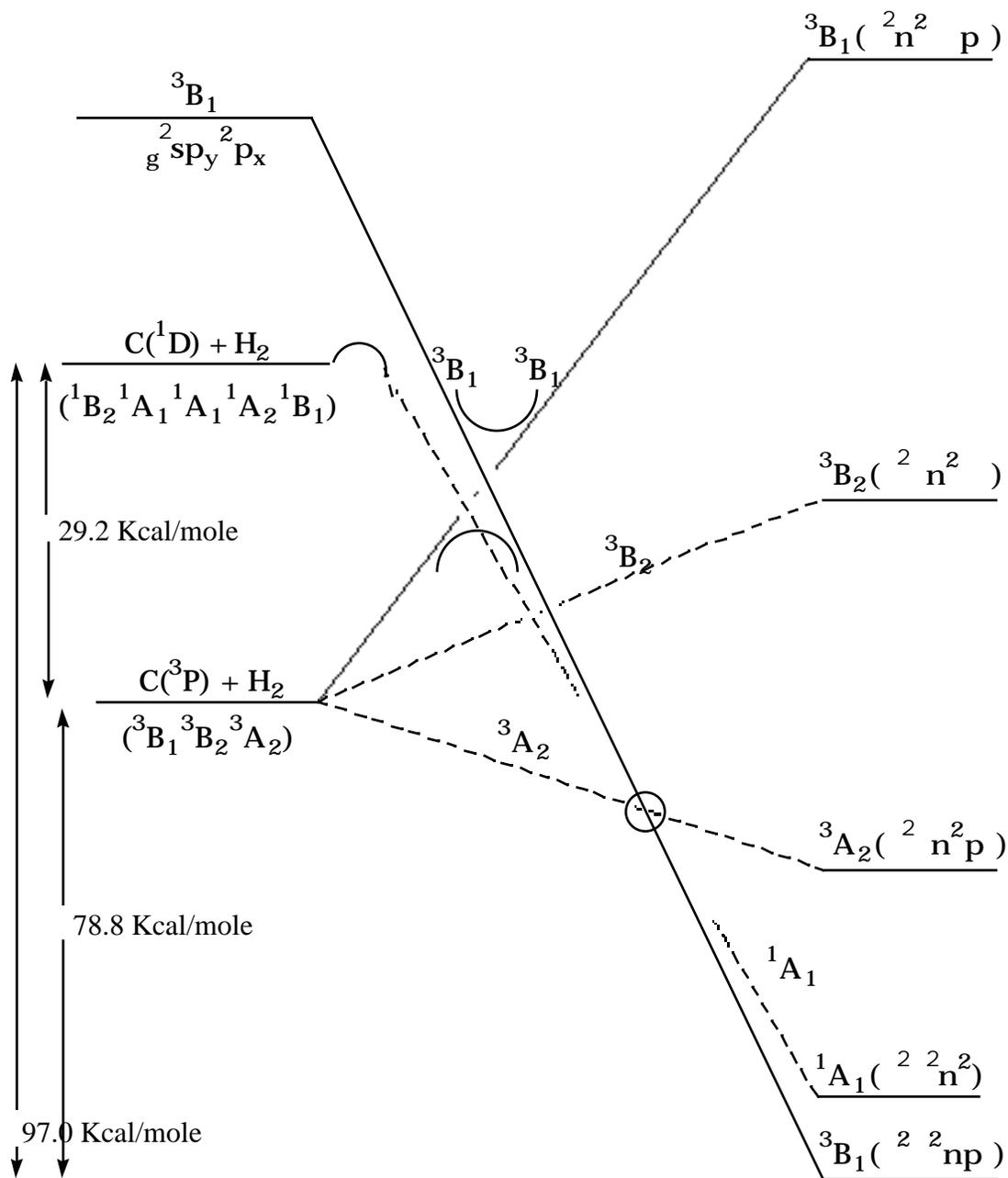
1. d. - e. It is necessary to determine how the wavefunctions found in part a. correlate with states of the  $CH_2$  molecule:



Note, the  $C + H_2$  state to which the lowest  ${}^1A_1$  ( $2n^2 \quad 2$ )  $CH_2$  state decomposes would be  $g^2s^2p_y^2$ . This state ( $g^2s^2p_y^2$ ) cannot be obtained by a simple combination of the  ${}^1D$  states. In order to obtain pure  $g^2s^2p_y^2$  it is necessary to combine  ${}^1S$  with  ${}^1D$ . For example,

$$g^2s^2p_y^2 = \frac{1}{6}(\sqrt{6} \quad {}^1D(0,0) - 2\sqrt{3} \quad {}^1S(0,0)) - \frac{1}{2}({}^1D(2,0) + {}^1D(-2,0)) .$$

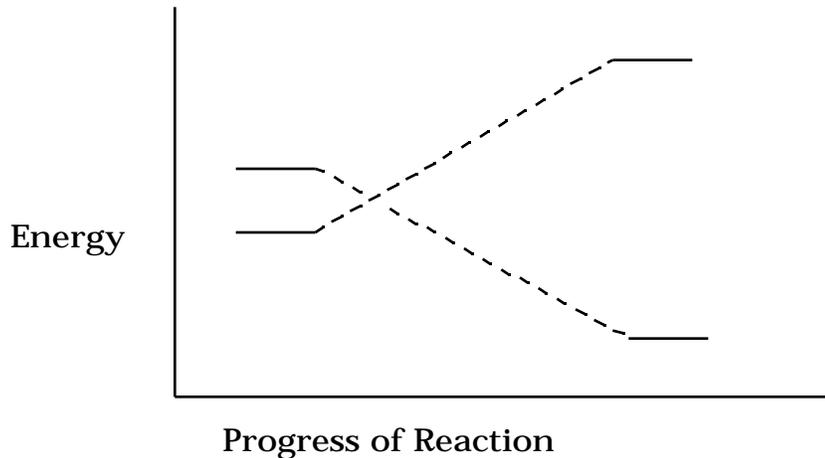
This indicates that a CCD must be drawn with a barrier near the  ${}^1D$  asymptote to represent the fact that  ${}^1A_1$   $CH_2$  correlates with a mixture of  ${}^1D$  and  ${}^1S$  carbon plus hydrogen. The  $C + H_2$  state to which the lowest  ${}^3B_1$  ( $2n^2 \quad 2p$ )  $CH_2$  state decomposes would be  $g^2sp_y^2p_x$ .



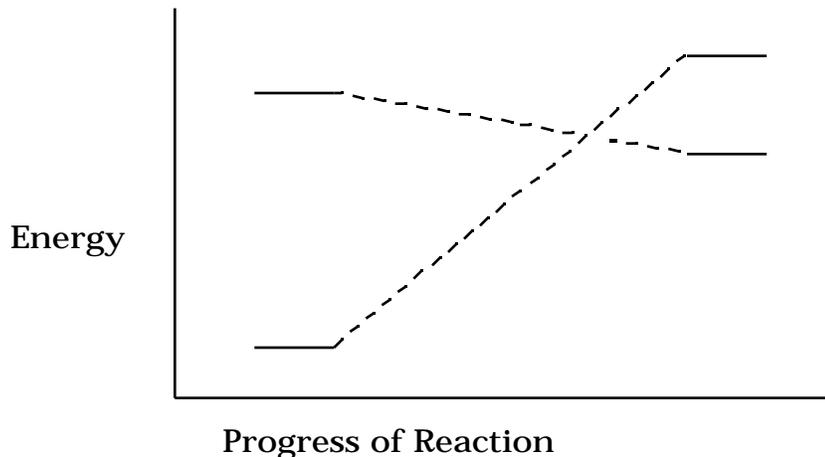
1. f. If you follow the  $^3B_1$  component of the  $C(^3P) + H_2$  (since it leads to the ground-state products) to  $^3B_1 CH_2$  you must go over an approximately 20 Kcal/mole barrier. Of course this path produces  $^3B_1 CH_2$  product. Distortions away from  $C_{2v}$  symmetry, for example to  $C_s$  symmetry, would make the  $a_1$  and  $b_2$  orbitals identical in symmetry ( $a'$ ). The  $b_1$  orbitals would maintain their identity going to  $a''$  symmetry. Thus  $^3B_1$  and  $^3A_2$  (both  $^3A''$  in  $C_s$  symmetry and odd under reflection through the molecular plane) can mix. The system could thus follow the  $^3A_2$  component of the  $C(^3P) + H_2$  surface to the place

(marked with a circle on the CCD) where it crosses the  $^3B_1$  surface upon which it then moves and continues to products. As a result, the barrier would be lowered.

You can estimate when the barrier occurs (late or early) using thermodynamic information for the reaction (i.e. slopes and asymptotic energies). For example, an early barrier would be obtained for a reaction with the characteristics:



and a late barrier would be obtained for a reaction with the characteristics:

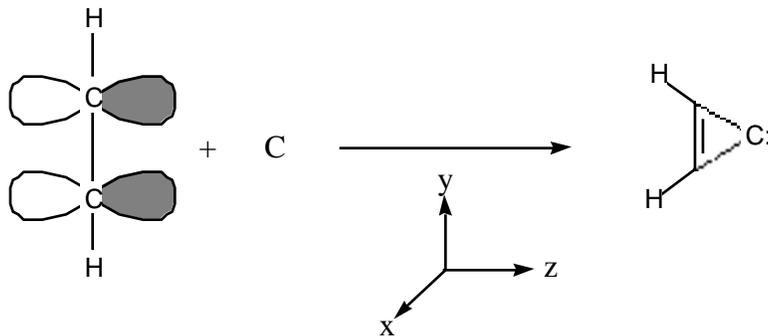


This relation between reaction endothermicity or exothermicity is known as the Hammond postulate. Note that the  $C(^3P_1) + H_2 \rightarrow CH_2$  reaction of interest here (see the CCD) has an early barrier.

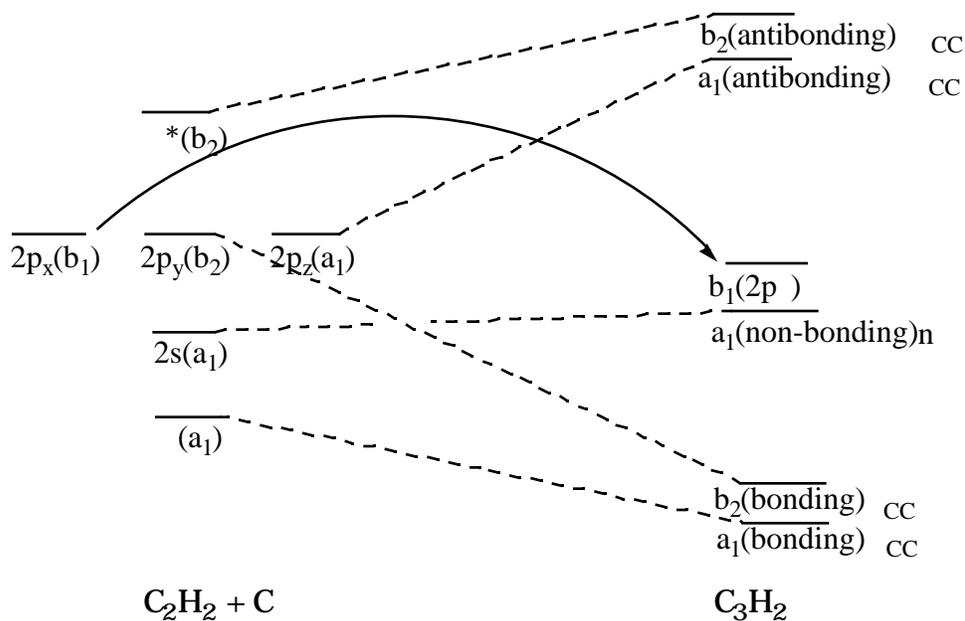
1. g. The reaction  $C(^1D) + H_2 \rightarrow CH_2 (^1A_1)$  should have no symmetry barrier (this can be recognized by following the  $^1A_1 (C(^1D) + H_2)$  reactants down to the  $^1A_1 (CH_2)$  products on the CCD).
2. This problem in many respects is analogous to problem 1.

The  $^3B_1$  surface certainly requires a two configuration CI wavefunction; the  $^2 2np_x$  ( $^2 p_y^2 sp_x$ ) and the  $^2 n^2 p_x * (^2 s^2 p_x p_z)$ . The  $^1A_1$  surface could use the  $^2 2n^2$  ( $^2 s^2 p_y^2$ ) only but once again there is no combination of  $^1D$  determinants which gives purely this configuration ( $^2 s^2 p_y^2$ ). Thus mixing of both  $^1D$  and  $^1S$  determinants are necessary to

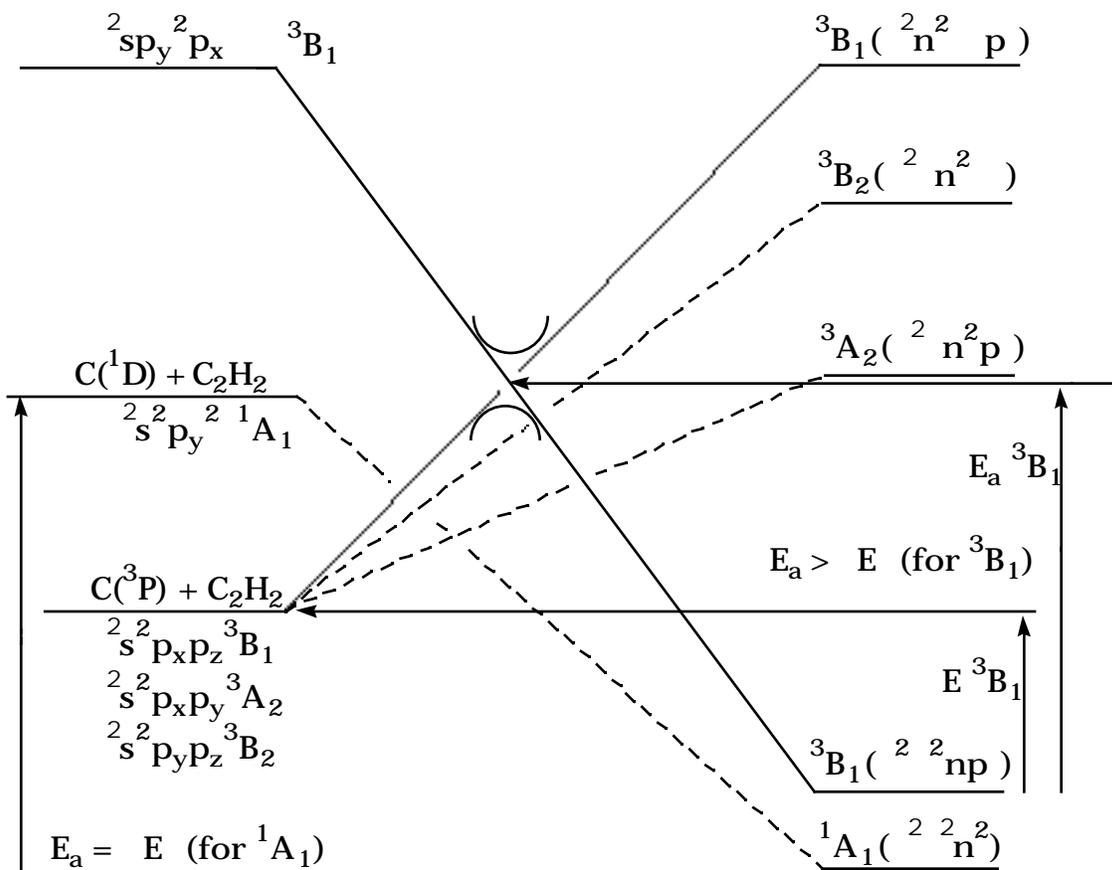
yield the required  $2s^2 2p_y^2$  configuration. Hence even the  $^1A_1$  surface would require a multiconfigurational wavefunction for adequate description.



Orbital-correlation diagram for the reaction  $C_2H_2 + C \rightarrow C_3H_2$



Configuration correlation diagram for the reaction  $C_2H_2 + C \rightarrow C_3H_2$ .



3. a.

$$\begin{aligned}
 \langle g | g \rangle &= \left\langle (2+2S)^{-\frac{1}{2}}(1s_A + 1s_B) \middle| (2+2S)^{-\frac{1}{2}}(1s_A + 1s_B) \right\rangle \\
 &= (2+2S)^{-1} (\langle 1s_A | 1s_A \rangle + \langle 1s_A | 1s_B \rangle + \langle 1s_B | 1s_A \rangle + \langle 1s_B | 1s_B \rangle) \\
 &= (0.285)((1.000) + (0.753) + (0.753) + (1.000)) \\
 &= 0.999 \quad 1
 \end{aligned}$$

$$\begin{aligned}
 \langle g | u \rangle &= \left\langle (2+2S)^{-\frac{1}{2}}(1s_A + 1s_B) \middle| (2-2S)^{-\frac{1}{2}}(1s_A - 1s_B) \right\rangle \\
 &= (2+2S)^{-\frac{1}{2}}(2-2S)^{-\frac{1}{2}} (\langle 1s_A | 1s_A \rangle + \langle 1s_A | 1s_B \rangle \\
 &\quad + \langle 1s_B | 1s_A \rangle + \langle 1s_B | 1s_B \rangle) \\
 &= (1.423)(0.534)((1.000) - (0.753) + (0.753) - (1.000)) \\
 &= 0
 \end{aligned}$$

$$\begin{aligned}
 \langle u | u \rangle &= \left\langle (2-2S)^{-\frac{1}{2}}(1s_A - 1s_B) \middle| (2-2S)^{-\frac{1}{2}}(1s_A - 1s_B) \right\rangle \\
 &= (2-2S)^{-1} (\langle 1s_A | 1s_A \rangle - \langle 1s_A | 1s_B \rangle - \langle 1s_B | 1s_A \rangle + \langle 1s_B | 1s_B \rangle) \\
 &= (2.024)((1.000) - (0.753) - (0.753) + (1.000)) \\
 &= 1.000
 \end{aligned}$$

3. b.

$$\begin{aligned}
 \langle g|h|g\rangle &= \langle (2+2S)^{-\frac{1}{2}}(1s_A + 1s_B) |h|(2+2S)^{-\frac{1}{2}}(1s_A + 1s_B) \rangle \\
 &= (2+2S)^{-1} (\langle 1s_A|h|1s_A\rangle + \langle 1s_A|h|1s_B\rangle \\
 &\quad + \langle 1s_B|h|1s_A\rangle + \langle 1s_B|h|1s_B\rangle) \\
 &= (0.285)((-1.110) + (-0.968) + (-0.968) + (-1.110)) \\
 &= -1.184
 \end{aligned}$$

$$\begin{aligned}
 \langle u|h|u\rangle &= \langle (2-2S)^{-\frac{1}{2}}(1s_A - 1s_B) |h|(2-2S)^{-\frac{1}{2}}(1s_A - 1s_B) \rangle \\
 &= (2-2S)^{-1} (\langle 1s_A|h|1s_A\rangle - \langle 1s_A|h|1s_B\rangle \\
 &\quad - \langle 1s_B|h|1s_A\rangle + \langle 1s_B|h|1s_B\rangle) \\
 &= (2.024)((-1.110) + (0.968) + (0.968) + (-1.110)) \\
 &= -0.575
 \end{aligned}$$

$$\begin{aligned}
 \langle g|g|h|g\rangle &= \langle gg|gg\rangle = (2+2S)^{-1}(2+2S)^{-1} \cdot \\
 &\langle (1s_A + 1s_B)(1s_A + 1s_B) | (1s_A + 1s_B)(1s_A + 1s_B) \rangle \\
 &= (2+2S)^{-2} \cdot (\langle AA|AA\rangle + \langle AA|AB\rangle + \langle AA|BA\rangle + \langle AA|BB\rangle + \\
 &\quad \langle AB|AA\rangle + \langle AB|AB\rangle + \langle AB|BA\rangle + \langle AB|BB\rangle + \\
 &\quad \langle BA|AA\rangle + \langle BA|AB\rangle + \langle BA|BA\rangle + \langle BA|BB\rangle + \\
 &\quad \langle BB|AA\rangle + \langle BB|AB\rangle + \langle BB|BA\rangle + \langle BB|BB\rangle) \\
 &= (0.081) ( (0.625) + (0.426) + (0.426) + (0.323) + \\
 &\quad (0.426) + (0.504) + (0.323) + (0.426) + \\
 &\quad (0.426) + (0.323) + (0.504) + (0.426) + \\
 &\quad (0.323) + (0.426) + (0.426) + (0.625) ) \\
 &= 0.564
 \end{aligned}$$

$$\begin{aligned}
 \langle uu|uu\rangle &= (2-2S)^{-1}(2-2S)^{-1} \cdot \\
 &\langle (1s_A - 1s_B)(1s_A - 1s_B) | (1s_A - 1s_B)(1s_A - 1s_B) \rangle \\
 &= (2-2S)^{-2} \cdot (\langle AA|AA\rangle - \langle AA|AB\rangle - \langle AA|BA\rangle + \langle AA|BB\rangle - \\
 &\quad \langle AB|AA\rangle + \langle AB|AB\rangle + \langle AB|BA\rangle - \langle AB|BB\rangle - \\
 &\quad \langle BA|AA\rangle + \langle BA|AB\rangle + \langle BA|BA\rangle - \langle BA|BB\rangle + \\
 &\quad \langle BB|AA\rangle - \langle BB|AB\rangle - \langle BB|BA\rangle + \langle BB|BB\rangle) \\
 &= (4.100) ( (0.625) - (0.426) - (0.426) + (0.323) - \\
 &\quad (0.426) + (0.504) + (0.323) - (0.426) - \\
 &\quad (0.426) + (0.323) + (0.504) - (0.426) + \\
 &\quad (0.323) - (0.426) - (0.426) + (0.625) ) \\
 &= 0.582
 \end{aligned}$$

$$\begin{aligned}
 \langle gg|uu\rangle &= (2+2S)^{-1}(2-2S)^{-1} \cdot \\
 &\langle (1s_A + 1s_B)(1s_A + 1s_B) | (1s_A - 1s_B)(1s_A - 1s_B) \rangle
 \end{aligned}$$

$$\begin{aligned}
&= (2+2S)^{-1}(2-2S)^{-1} \cdot \\
&\quad (\langle AA|AA \rangle - \langle AA|AB \rangle - \langle AA|BA \rangle + \langle AA|BB \rangle + \\
&\quad \langle AB|AA \rangle - \langle AB|AB \rangle - \langle AB|BA \rangle + \langle AB|BB \rangle + \\
&\quad \langle BA|AA \rangle - \langle BA|AB \rangle - \langle BA|BA \rangle + \langle BA|BB \rangle + \\
&\quad \langle BB|AA \rangle - \langle BB|AB \rangle - \langle BB|BA \rangle + \langle BB|BB \rangle ) \\
&= (0.285)(2.024) ((0.625) - (0.426) - (0.426) + (0.323) + \\
&\quad (0.426) - (0.504) - (0.323) + (0.426) + \\
&\quad (0.426) - (0.323) - (0.504) + (0.426) + \\
&\quad (0.323) - (0.426) - (0.426) + (0.625)) \\
&= 0.140
\end{aligned}$$

$$\begin{aligned}
\langle gu|gu \rangle &= (2+2S)^{-1}(2-2S)^{-1} \cdot \\
&\quad \langle (1s_A + 1s_B)(1s_A - 1s_B) | (1s_A + 1s_B)(1s_A - 1s_B) \rangle \\
&= (2+2S)^{-1}(2-2S)^{-1} \cdot \\
&\quad (\langle AA|AA \rangle - \langle AA|AB \rangle + \langle AA|BA \rangle - \langle AA|BB \rangle - \\
&\quad \langle AB|AA \rangle + \langle AB|AB \rangle - \langle AB|BA \rangle + \langle AB|BB \rangle + \\
&\quad \langle BA|AA \rangle - \langle BA|AB \rangle + \langle BA|BA \rangle - \langle BA|BB \rangle - \\
&\quad \langle BB|AA \rangle + \langle BB|AB \rangle - \langle BB|BA \rangle + \langle BB|BB \rangle ) \\
&= (0.285)(2.024) ((0.625) - (0.426) + (0.426) - (0.323) - \\
&\quad (0.426) + (0.504) - (0.323) + (0.426) + \\
&\quad (0.426) - (0.323) + (0.504) - (0.426) - \\
&\quad (0.323) + (0.426) - (0.426) + (0.625)) \\
&= 0.557
\end{aligned}$$

Note, that  $\langle gg|gu \rangle = \langle uu|ug \rangle = 0$  from symmetry considerations, but this can be easily verified. For example,

$$\begin{aligned}
\langle gg|gu \rangle &= (2+2S)^{-\frac{1}{2}}(2-2S)^{-\frac{3}{2}} \cdot \\
&\quad \langle (1s_A + 1s_B)(1s_A + 1s_B) | (1s_A + 1s_B)(1s_A - 1s_B) \rangle \\
&= (2+2S)^{-\frac{1}{2}}(2-2S)^{-\frac{3}{2}} \cdot \\
&\quad (\langle AA|AA \rangle - \langle AA|AB \rangle + \langle AA|BA \rangle - \langle AA|BB \rangle + \\
&\quad \langle AB|AA \rangle - \langle AB|AB \rangle + \langle AB|BA \rangle - \langle AB|BB \rangle + \\
&\quad \langle BA|AA \rangle - \langle BA|AB \rangle + \langle BA|BA \rangle - \langle BA|BB \rangle + \\
&\quad \langle BB|AA \rangle - \langle BB|AB \rangle + \langle BB|BA \rangle - \langle BB|BB \rangle ) \\
&= (0.534)(2.880) ((0.625) - (0.426) + (0.426) - (0.323) + \\
&\quad (0.426) - (0.504) + (0.323) - (0.426) + \\
&\quad (0.426) - (0.323) + (0.504) - (0.426) + \\
&\quad (0.323) - (0.426) + (0.426) - (0.625)) \\
&= 0.000
\end{aligned}$$

3. c. We can now set up the configuration interaction Hamiltonian matrix. The elements are evaluated by using the Slater-Condon rules as shown in the text.

$$H_{11} = \langle g \quad g | H | g \quad g \rangle$$

$$= 2f_g + g_g$$

$$= 2(-1.184) + 0.564 = -1.804$$

$$H_{21} = H_{12} = \langle g | H | u \rangle$$

$$= \langle g | g | u \rangle$$

$$= 0.140$$

$$H_{22} = \langle u | H | u \rangle$$

$$= 2f_u + g_u$$

$$= 2(-0.575) + 0.582 = -0.568$$

3. d. Solving this eigenvalue problem:

$$\begin{vmatrix} -1.804 - \lambda & 0.140 \\ 0.140 & -0.568 - \lambda \end{vmatrix} = 0$$

$$(-1.804 - \lambda)(-0.568 - \lambda) - (0.140)^2 = 0$$

$$1.025 + 1.804\lambda + 0.568\lambda + \lambda^2 - 0.0196 = 0$$

$$\lambda^2 + 2.372\lambda + 1.005 = 0$$

$$= \frac{-2.372 \pm \sqrt{(2.372)^2 - 4(1)(1.005)}}{(2)(1)}$$

$$= -1.186 \pm 0.634$$

$$= -1.820, \text{ and } -0.552.$$

Solving for the coefficients:

$$\begin{vmatrix} -1.804 - \lambda & 0.140 \\ 0.140 & -0.568 - \lambda \end{vmatrix} \begin{matrix} C_1 \\ C_2 \end{matrix} = 0$$

For the first eigenvalue this becomes:

$$\begin{vmatrix} -1.804 + 1.820 & 0.140 \\ 0.140 & -0.568 + 1.820 \end{vmatrix} \begin{matrix} C_1 \\ C_2 \end{matrix} = 0$$

$$\begin{vmatrix} 0.016 & 0.140 \\ 0.140 & 1.252 \end{vmatrix} \begin{matrix} C_1 \\ C_2 \end{matrix} = 0$$

$$(0.140)(C_1) + (1.252)(C_2) = 0$$

$$C_1 = -8.943 C_2$$

$$C_1^2 + C_2^2 = 1 \text{ (from normalization)}$$

$$(-8.943 C_2)^2 + C_2^2 = 1$$

$$80.975 C_2^2 = 1$$

$$C_2 = 0.111, C_1 = -0.994$$

For the second eigenvalue this becomes:

$$\begin{vmatrix} -1.804 + 0.552 & 0.140 \\ 0.140 & -0.568 + 0.552 \end{vmatrix} \begin{matrix} C_1 \\ C_2 \end{matrix} = 0$$

$$\begin{vmatrix} -1.252 & 0.140 \\ 0.140 & -0.016 \end{vmatrix} \begin{matrix} C_1 \\ C_2 \end{matrix} = 0$$

$$(-1.252)(C_1) + (0.140)(C_2) = 0$$

$$C_1 = 0.112 C_2$$

$$C_1^2 + C_2^2 = 1 \text{ (from normalization)}$$

$$(0.112 C_2)^2 + C_2^2 = 1$$

$$1.0125 C_2^2 = 1$$

$$C_2 = 0.994, C_1 = 0.111$$

3. e. The polarized orbitals,  $R_{\pm}$ , are given by:

$$R_{\pm} = g \pm \sqrt{\frac{C_2}{C_1}} u$$

$$R_{\pm} = g \pm \sqrt{\frac{0.111}{0.994}} u$$

$$R_{\pm} = g \pm 0.334 u$$

$$R_+ = g + 0.334 u \text{ (left polarized)}$$

$$R_- = g - 0.334 u \text{ (right polarized)}$$



$R_+$  Left Polarized



$R_-$  Right Polarized

## Section 4 Exercises, Problems, and Solutions

### Exercises:

- Consider the molecules  $\text{CCl}_4$ ,  $\text{CHCl}_3$ , and  $\text{CH}_2\text{Cl}_2$ .
  - What kind of rotor are they (symmetric top, etc; do not bother with oblate, or near-prolate, etc.)
  - Will they show pure rotational spectra?
  - Assume that ammonia shows a pure rotational spectrum. If the rotational constants are  $9.44$  and  $6.20 \text{ cm}^{-1}$ , use the energy expression:

$$E = (A - B) K^2 + B J(J + 1),$$

to calculate the energies (in  $\text{cm}^{-1}$ ) of the first three lines (i.e., those with lowest  $K$ ,  $J$  quantum number for the absorbing level) in the absorption spectrum (ignoring higher order terms in the energy expression).

- The molecule  $^{11}\text{B } ^{16}\text{O}$  has a vibrational frequency  $\nu_e = 1885 \text{ cm}^{-1}$ , a rotational constant  $B_e = 1.78 \text{ cm}^{-1}$ , and a bond energy from the bottom of the potential well of  $D_e^0 = 8.28 \text{ eV}$ . Use integral atomic masses in the following:

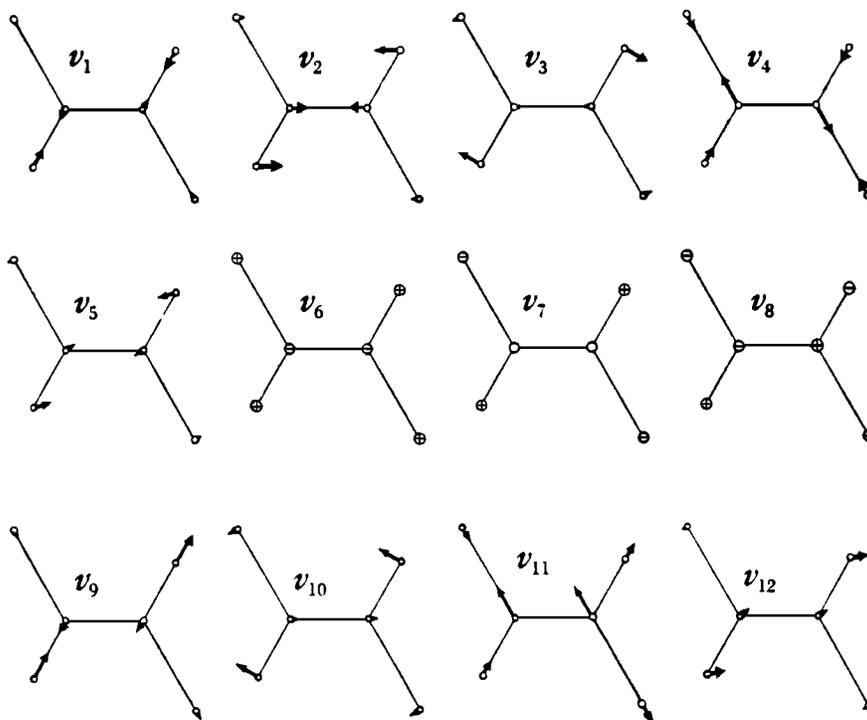
- In the approximation that the molecule can be represented as a Morse oscillator, calculate the bond length,  $R_e$  in angstroms, the centrifugal distortion constant,  $D_e$  in  $\text{cm}^{-1}$ , the anharmonicity constant,  $x_e$  in  $\text{cm}^{-1}$ , the zero-point corrected bond energy,  $D_0^0$  in eV, the vibration rotation interaction constant,  $B_1$  in  $\text{cm}^{-1}$ , and the vibrational state specific rotation constants,  $B_0$  and  $B_1$  in  $\text{cm}^{-1}$ . Use the vibration-rotation energy expression for a Morse oscillator:

$$E = \hbar \nu_e (v + 1/2) - \hbar \nu_e x_e (v + 1/2)^2 + B_v J(J + 1) - D_e J^2(J + 1)^2, \text{ where}$$

$$B_v = B_e - \nu_e (v + 1/2), \quad x_e = \frac{-6B_e^2}{\hbar \nu_e} + \frac{6\sqrt{B_e^3 \hbar \nu_e x_e}}{\hbar \nu_e}, \text{ and } D_e = \frac{4B_e^3}{\hbar \nu_e^2}.$$

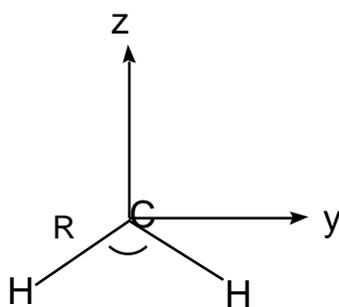
- Will this molecule show a pure rotation spectrum? A vibration-rotation spectrum? Assume that it does, what are the energies (in  $\text{cm}^{-1}$ ) of the first three lines in the P branch ( $v = +1$ ,  $J = -1$ ) of the fundamental absorption?

- Consider  $\text{trans-C}_2\text{H}_2\text{Cl}_2$ . The vibrational normal modes of this molecule are shown below. What is the symmetry of the molecule? Label each of the modes with the appropriate irreducible representation.



Problems:

1. Suppose you are given two molecules (one is  $\text{CH}_2$  and the other is  $\text{CH}_2^-$  but you don't know which is which). Both molecules have  $C_{2v}$  symmetry. The CH bond length of molecule I is  $1.121 \text{ \AA}$  and for molecule II it is  $1.076 \text{ \AA}$ . The bond angle of molecule I is  $104^\circ$  and for molecule II it is  $136^\circ$ .



a. Using a coordinate system centered on the C nucleus as shown above (the molecule is in the YZ plane), compute the moment of inertia tensors of both species (I and II). The definitions of the components of the tensor are, for example:

$$I_{xx} = \sum_j m_j(y_j^2 + z_j^2) - M(Y^2 + Z^2)$$

$$I_{xy} = - \sum_j m_j x_j y_j - MXY$$

Here,  $m_j$  is the mass of the nucleus  $j$ ,  $M$  is the mass of the entire molecule, and  $X, Y, Z$  are the coordinates of the center of mass of the molecule. Use Å for distances and amu's for masses.

b. Find the principal moments of inertia  $I_a < I_b < I_c$  for both compounds (in amu Å<sup>2</sup> units) and convert these values into rotational constants  $A, B,$  and  $C$  in cm<sup>-1</sup> using, for example,

$$A = h(8 \pi^2 c I_a)^{-1}.$$

c. Both compounds are "nearly prolate tops" whose energy levels can be well approximated using the prolate top formula:

$$E = (A - B) K^2 + B J(J + 1),$$

if one uses for the  $B$  constant the average of the  $B$  and  $C$  values determined earlier. Thus, take  $B$  and  $C$  values (for each compound) and average them to produce an effective  $B$  constant to use in the above energy formula. Write down (in cm<sup>-1</sup> units) the energy formula for both species. What values are  $J$  and  $K$  allowed to assume? What is the degeneracy of the level labeled by a given  $J$  and  $K$ ?

d. Draw a picture of both compounds and show the directions of the three principle axes (a,b,c). On these pictures show the kind of rotational motion associated with the quantum number  $K$ .

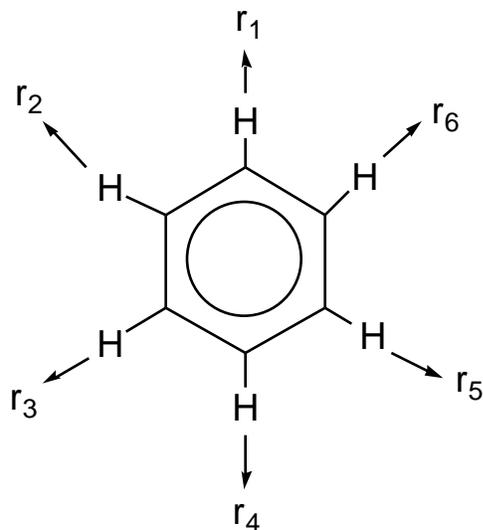
e. Given that the electronic transition moment vector  $\mu$  connecting species I and II is directed along the  $Y$  axis, what are the selection rules  $J$  and  $K$ ?

f. Suppose you are given the photoelectron spectrum of  $\text{CH}_2^-$ . In this spectrum  $J_j = J_i + 1$  transitions are called R-branch absorptions and those obeying  $J_j = J_i - 1$  are called P-branch transitions. The spacing between lines can increase or decrease as functions of  $J_i$  depending on the changes in the moment of inertia for the transition. If spacings grow closer and closer, we say that the spectrum exhibits a so-called band head formation. In the photoelectron spectrum that you are given, a rotational analysis of the vibrational lines in this spectrum is carried out and it is found that the R-branches show band head formation but the P-branches do not. Based on this information, determine which compound I or II is the  $\text{CH}_2^-$  anion. Explain your reasoning.

g. At what  $J$  value (of the absorbing species) does the band head occur and at what rotational energy difference?

2. Let us consider the vibrational motions of benzene. To consider all of the vibrational modes of benzene we should attach a set of displacement vectors in the  $x, y,$  and  $z$  directions to each atom in the molecule (giving 36 vectors in all), and evaluate how these transform under the symmetry operations of  $D_{6h}$ . For this problem, however, let's only inquire about the C-H stretching vibrations.

a. Represent the C-H stretching motion on each C-H bond by an outward-directed vector on each H atom, designated  $r_i$ :



These vectors form the basis for a reducible representation. Evaluate the characters for this reducible representation under the symmetry operations of the  $D_{6h}$  group.

b. Decompose the reducible representation you obtained in part a. into its irreducible components. These are the symmetries of the various C-H stretching vibrational modes in benzene.

c. The vibrational state with zero quanta in each of the vibrational modes (the ground vibrational state) of any molecule always belongs to the totally symmetric representation. For benzene the ground vibrational state is therefore of  $A_{1g}$  symmetry. An excited state which has one quantum of vibrational excitation in a mode which is of a given symmetry species has the same symmetry species as the mode which is excited (because the vibrational wave functions are given as Hermite polynomials in the stretching coordinate). Thus, for example, excitation (by one quantum) of a vibrational mode of  $A_{2u}$  symmetry gives a wavefunction of  $A_{2u}$  symmetry. To resolve the question of what vibrational modes may be excited by the absorption of infrared radiation we must examine the x, y, and z components of the transition dipole integral for initial and final state wave functions  $\psi_i$  and  $\psi_f$ , respectively:

$$|\langle \psi_f | x | \psi_i \rangle|, |\langle \psi_f | y | \psi_i \rangle|, \text{ and } |\langle \psi_f | z | \psi_i \rangle|.$$

Using the information provided above, which of the C-H vibrational modes of benzene will be infrared-active, and how will the transitions be polarized? How many C-H vibrations will you observe in the infrared spectrum of benzene?

d. A vibrational mode will be active in Raman spectroscopy only if one of the following integrals is nonzero:

$$|\langle \psi_f | xy | \psi_i \rangle|, |\langle \psi_f | xz | \psi_i \rangle|, |\langle \psi_f | yz | \psi_i \rangle|, \\ |\langle \psi_f | x^2 | \psi_i \rangle|, |\langle \psi_f | y^2 | \psi_i \rangle|, \text{ and } |\langle \psi_f | z^2 | \psi_i \rangle|.$$

Using the fact that the quadratic operators transform according to the irreducible representations:

$$(x^2 + y^2, z^2) \quad A_{1g}$$

$$(xz, yz) \quad E_{1g}$$

$$(x^2 - y^2, xy) \quad E_{2g}$$

Determine which of the C-H vibrational modes will be Raman-active.

e. Are there any of the C-H stretching vibrational motions of benzene which cannot be observed in either infrared or Raman spectroscopy? Give the irreducible representation label for these unobservable modes.

3. In treating the vibrational and rotational motion of a diatomic molecule having reduced mass  $\mu$ , equilibrium bond length  $r_e$  and harmonic force constant  $k$ , we are faced with the following radial Schrödinger equation:

$$\frac{-\hbar^2}{2\mu r^2} \frac{d}{dr} \left( r^2 \frac{dR}{dr} \right) + \frac{J(J+1)\hbar^2}{2\mu r^2} R + \frac{1}{2} k(r - r_e)^2 R = E R$$

a. Show that the substitution  $R = r^{-1}F$  leads to:

$$\frac{-\hbar^2}{2\mu} F'' + \frac{J(J+1)\hbar^2}{2\mu r^2} F + \frac{1}{2} k(r - r_e)^2 F = E F$$

b. Taking  $r = r_e + \rho$  and expanding  $(1 + x)^{-2} = 1 - 2x + 3x^2 + \dots$ ,

show that the so-called vibration-rotation coupling term  $\frac{J(J+1)\hbar^2}{2\mu r^2}$  can be approximated

(for small  $\rho$ ) by  $\frac{J(J+1)\hbar^2}{2\mu r_e^2} \left( 1 - \frac{2\rho}{r_e} + \frac{3\rho^2}{r_e^2} \right)$ . Keep terms only through order  $\rho^2$ .

c. Show that, through terms of order  $\rho^2$ , the above equation for  $F$  can be rearranged to yield a new equation of the form:

$$\frac{-\hbar^2}{2\mu} F'' + \frac{1}{2} k(r - r_e)^2 F = E F - \frac{J(J+1)\hbar^2}{2\mu r_e^2} F + \dots$$

Give explicit expressions for how the modified force constant  $k$ , bond length  $r_e$ , and energy shift depend on  $J$ ,  $k$ ,  $r_e$ , and  $\mu$ .

d. Given the above modified vibrational problem, we can now conclude that the modified energy levels are:

$$E = \hbar \sqrt{\frac{k}{\mu}} \left( v + \frac{1}{2} \right) + \frac{J(J+1)\hbar^2}{2\mu r_e^2} - \dots$$

Explain how the conclusion is "obvious", how for  $J = 0$ ,  $k = k$ , and  $r_e = r_e$ , we obtain the usual harmonic oscillator energy levels. Describe how the energy levels would be expected to vary as  $J$  increases from zero and explain how these changes arise from changes in  $k$  and  $r_e$ . Explain in terms of physical forces involved in the rotating-vibrating molecule why  $r_e$  and  $k$  are changed by rotation.

## Solutions

### Exercises:

1. a.  $\text{CCl}_4$  is tetrahedral and therefore is a spherical top.  $\text{CHCl}_3$  has  $C_{3v}$  symmetry and therefore is a symmetric top.  $\text{CH}_2\text{Cl}_2$  has  $C_{2v}$  symmetry and therefore is an asymmetric top.

b.  $\text{CCl}_4$  has such high symmetry that it will not exhibit pure rotational spectra.  $\text{CHCl}_3$  and  $\text{CH}_2\text{Cl}_2$  will both exhibit pure rotation spectra.

c.  $\text{NH}_3$  is a symmetric top (oblate). Use the given energy expression,

$$E = (A - B) K^2 + B J(J + 1),$$

$A = 6.20 \text{ cm}^{-1}$ ,  $B = 9.44 \text{ cm}^{-1}$ , selection rules  $J = \pm 1$ , and the fact that  $\mu_0$  lies along the figure axis such that  $K = 0$ , to give:

$$E = 2B (J + 1) = 2B, 4B, \text{ and } 6B \text{ (} J = 0, 1, \text{ and } 2\text{)}.$$

So, lines are at  $18.88 \text{ cm}^{-1}$ ,  $37.76 \text{ cm}^{-1}$ , and  $56.64 \text{ cm}^{-1}$ .

2. To convert between  $\text{cm}^{-1}$  and energy, multiply by  $hc = (6.62618 \times 10^{-34} \text{ J sec})(2.997925 \times 10^{10} \text{ cm sec}^{-1}) = 1.9865 \times 10^{23} \text{ J cm}$ .

Let all quantities in  $\text{cm}^{-1}$  be designated with a bar,

e.g.  $B_e = 1.78 \text{ cm}^{-1}$ .

$$\text{a. } hcB_e = \frac{\hbar^2}{2\mu R_e^2}$$

$$R_e = \frac{\hbar}{\sqrt{2\mu hcB_e}},$$

$$\mu = \frac{m_B m_O}{m_B + m_O} = \frac{(11)(16)}{(11 + 16)} \times 1.66056 \times 10^{-27} \text{ kg} \\ = 1.0824 \times 10^{-26} \text{ kg}.$$

$$hcB_e = hc(1.78 \text{ cm}^{-1}) = 3.5359 \times 10^{-23} \text{ J}$$

$$R_e = \frac{1.05459 \times 10^{-34} \text{ J sec}}{\sqrt{(2)(1.0824 \times 10^{-26} \text{ kg})(3.5359 \times 10^{-23} \text{ J})}}$$

$$R_e = 1.205 \times 10^{-10} \text{ m} = 1.205 \text{ \AA}$$

$$D_e = \frac{4B_e^3}{\hbar^2 e^2}, D_e = \frac{4B_e^3}{e^2} = \frac{(4)(1.78 \text{ cm}^{-1})^3}{(1885 \text{ cm}^{-1})^2} = 6.35 \times 10^{-6} \text{ cm}^{-1}$$

$$e^{x_e} = \frac{\hbar^2 e^2}{4D_e^0}, e^{x_e} = \frac{e^2}{4D_e^0} = \frac{(1885 \text{ cm}^{-1})^2}{(4)(66782.2 \text{ cm}^{-1})} = 13.30 \text{ cm}^{-1}.$$

$$D_0^0 = D_e^0 - \frac{\hbar^2 e}{2} + \frac{\hbar^2 e^{x_e}}{4}, D_0^0 = D_e^0 - \frac{e}{2} + \frac{e^{x_e}}{4} \\ = 66782.2 - \frac{1885}{2} + \frac{13.3}{4} \\ = 65843.0 \text{ cm}^{-1} = 8.16 \text{ eV}.$$

$$e = \frac{-6B_e^2}{\hbar^2 e} + \frac{6\sqrt{B_e^3 \hbar^2 e^{x_e}}}{\hbar^2 e}$$

$$e = \frac{-6B_e^2}{e} + \frac{6\sqrt{B_e^3 e^{x_e}}}{e}$$

$$e = \frac{(-6)(1.78)^2}{(1885)} + \frac{6\sqrt{(1.78)^3(13.3)}}{(1885)} = 0.0175 \text{ cm}^{-1}$$

$$B_0 = B_e - e(1/2), B_0 = B_e - e(1/2) = 1.78 - 0.0175/2 = 1.77 \text{ cm}^{-1}$$

$$B_1 = B_e - e(3/2), B_1 = B_e - e(3/2) = 1.78 - 0.0175(1.5) = 1.75 \text{ cm}^{-1}$$

b. The molecule has a dipole moment and so it should have a pure rotational spectrum. In addition, the dipole moment should change with R and so it should have a vibration rotation spectrum.

The first three lines correspond to  $J = 1 \rightarrow 0, J = 2 \rightarrow 1, J = 3 \rightarrow 2$

$$E = \hbar \omega_e(v + 1/2) - \hbar \omega_e x_e(v + 1/2)^2 + B_v J(J + 1) - D_e J^2(J + 1)^2$$

$$E = \hbar \omega_e - 2\hbar \omega_e x_e - B_0 J(J + 1) + B_1 J(J - 1) - 4D_e J^3$$

$$E = \omega_e - 2 \omega_e x_e - B_0 J(J + 1) + B_1 J(J - 1) - 4D_e J^3$$

$$E = 1885 - 2(13.3) - 1.77J(J + 1) + 1.75J(J - 1) - 4(6.35 \times 10^{-6})J^3$$

$$= 1858.4 - 1.77J(J + 1) + 1.75J(J - 1) - 2.54 \times 10^{-5}J^3$$

$$E(J = 1) = 1854.9 \text{ cm}^{-1}$$

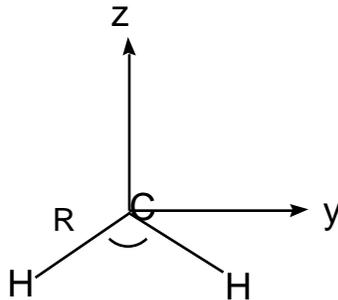
$$E(J = 2) = 1851.3 \text{ cm}^{-1}$$

$$E(J = 3) = 1847.7 \text{ cm}^{-1}$$

3. The  $C_2H_2Cl_2$  molecule has a  $\sigma_h$  plane of symmetry (plane of molecule), a  $C_2$  axis (to plane), and inversion symmetry, this results in  $C_{2h}$  symmetry. Using  $C_{2h}$  symmetry labels the modes can be labeled as follows: 1, 2, 3, 4, and 5 are  $a_g$ , 6 and 7 are  $a_u$ , 8 is  $b_g$ , and 9, 10, 11, and 12 are  $b_u$ .

Problems:

1.



Molecule I  
 $R_{CH} = 1.121 \text{ \AA}$

$$\text{HCH} = 104^\circ$$

$$y_H = R \sin(\theta/2) = \pm 0.8834$$

$$z_H = R \cos(\theta/2) = -0.6902$$

Center of Mass(COM):

Molecule II  
 $R_{CH} = 1.076 \text{ \AA}$

$$\text{HCH} = 136^\circ$$

$$y_H = \pm 0.9976$$

$$z_H = -0.4031$$

clearly,  $X = Y = 0$ ,

$$Z = \frac{12(0) - 2RCos(\pi/2)}{14} = -0.0986 \qquad Z = -0.0576$$

a. 
$$I_{xx} = \sum_j m_j(y_j^2 + z_j^2) - M(Y^2 + Z^2)$$

$$I_{xy} = - \sum_j m_j x_j y_j - MXY$$

$$I_{xx} = 2(1.121)^2 - 14(-0.0986)^2$$

$$= 2.377$$

$$I_{xx} = 2(1.076)^2 - 14(-0.0576)^2$$

$$= 2.269$$

$$I_{yy} = 2(0.6902)^2 - 14(-0.0986)^2$$

$$= 0.8167$$

$$I_{yy} = 2(0.4031)^2 - 14(-0.0576)^2$$

$$= 0.2786$$

$$I_{zz} = 2(0.8834)^2$$

$$= 1.561$$

$$I_{zz} = 2(0.9976)^2$$

$$= 1.990$$

$$I_{xz} = I_{yz} = I_{xy} = 0$$

b. Since the moment of inertia tensor is already diagonal, the principal moments of inertia have already been determined to be

$$(I_a < I_b < I_c):$$

$$I_{yy} < I_{zz} < I_{xx}$$

$$0.8167 < 1.561 < 2.377$$

$$I_{yy} < I_{zz} < I_{xx}$$

$$0.2786 < 1.990 < 2.269$$

Using the formula: 
$$A = \frac{h}{8 \sqrt{2cI_a}} = \frac{6.626 \times 10^{-27}}{8 \sqrt{2(3 \times 10^{10})I_a}} \times \frac{6.02 \times 10^{23}}{(1 \times 10^{-8})^2}$$

$$A = \frac{16.84}{I_a} \text{ cm}^{-1}$$

similarly,  $B = \frac{16.84}{I_b} \text{ cm}^{-1}$ , and  $C = \frac{16.84}{I_c} \text{ cm}^{-1}$ .

So,

Molecule I

y  $A = 20.62$

z  $B = 10.79$

x  $C = 7.08$

Molecule II

y  $A = 60.45$

z  $B = 8.46$

x  $C = 7.42$

c. Averaging B + C:

$$B = (B + C)/2 = 8.94$$

$$A - B = 11.68$$

$$B = (B + C)/2 = 7.94$$

$$A - B = 52.51$$

Using the prolate top formula:

$$E = (A - B) K^2 + B J(J + 1),$$

Molecule I

$$E = 11.68K^2 + 8.94J(J + 1)$$

Levels:  $J = 0, 1, 2, \dots$  and  $K = 0, 1, \dots, J$

For a given level defined by J and K, there are  $M_J$  degeneracies given by:  $(2J + 1) \times$

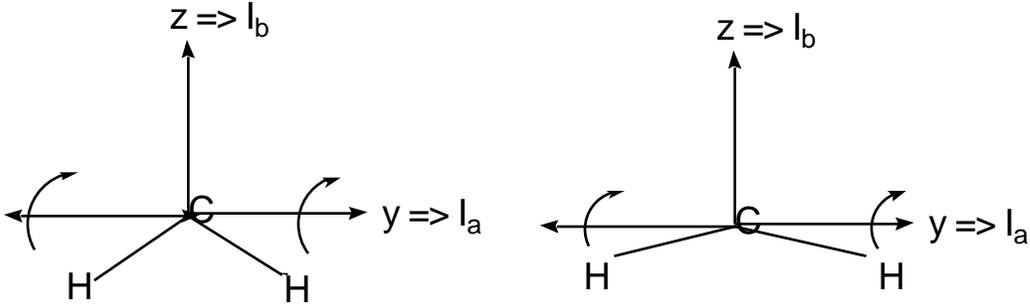
1 for  $K = 0$

2 for  $K = 1$

d.

Molecule I

Molecule II



e. Since  $\mu$  is along Y,  $K = 0$  since  $K$  describes rotation about the y axis.

Therefore  $J = \pm 1$

f. Assume molecule I is  $\text{CH}_2^-$  and molecule II is  $\text{CH}_2$ . Then,

$$E = E_{J_j}(\text{CH}_2) - E_{J_i}(\text{CH}_2^-), \text{ where:}$$

$$E(\text{CH}_2) = 52.51K^2 + 7.94J(J + 1), \text{ and } E(\text{CH}_2^-) = 11.68K^2 + 8.94J(J + 1)$$

For R-branches:  $J_j = J_i + 1, K = 0$ :

$$\begin{aligned} E_R &= E_{J_j}(\text{CH}_2) - E_{J_i}(\text{CH}_2^-) \\ &= 7.94(J_i + 1)(J_i + 1 + 1) - 8.94J_i(J_i + 1) \\ &= (J_i + 1)\{7.94(J_i + 1 + 1) - 8.94J_i\} \\ &= (J_i + 1)\{(7.94 - 8.94)J_i + 2(7.94)\} \\ &= (J_i + 1)\{-J_i + 15.88\} \end{aligned}$$

For P-branches:  $J_j = J_i - 1, K = 0$ :

$$\begin{aligned} E_P &= E_{J_j}(\text{CH}_2) - E_{J_i}(\text{CH}_2^-) \\ &= 7.94(J_i - 1)(J_i - 1 + 1) - 8.94J_i(J_i + 1) \\ &= J_i\{7.94(J_i - 1) - 8.94(J_i + 1)\} \\ &= J_i\{(7.94 - 8.94)J_i - 7.94 - 8.94\} \\ &= J_i\{-J_i - 16.88\} \end{aligned}$$

This indicates that the R branch lines occur at energies which grow closer and closer together as  $J$  increases (since the  $15.88 - J_i$  term will cancel). The P branch lines occur at energies which lie more and more negative (i.e. to the left of the origin). So, you can predict that if molecule I is  $\text{CH}_2^-$  and molecule II is  $\text{CH}_2$  then the R-branch has a band head and the P-branch does not. This is observed therefore our assumption was correct:

molecule I is  $\text{CH}_2^-$  and molecule II is  $\text{CH}_2$ .

g. The band head occurs when  $\frac{d(E_R)}{dJ} = 0$ .

$$\begin{aligned} \frac{d(E_R)}{dJ} &= \frac{d}{dJ} [(J_i + 1)\{-J_i + 15.88\}] = 0 \\ &= \frac{d}{dJ} (-J_i^2 - J_i + 15.88J_i + 15.88) = 0 \\ &= -2J_i + 14.88 = 0 \\ J_i &= 7.44, \text{ so } J = 7 \text{ or } 8. \end{aligned}$$

At  $J = 7.44$ :

$$E_R = (J + 1)\{-J + 15.88\}$$

$$E_R = (7.44 + 1)\{-7.44 + 15.88\} = (8.44)(8.44) = 71.2 \text{ cm}^{-1} \text{ above the origin.}$$

2. a.

D <sub>6h</sub>	E	2C <sub>6</sub>	2C <sub>3</sub>	C <sub>2</sub>	3C <sub>2</sub> '	3C <sub>2</sub> "	i	2S <sub>3</sub>	2S <sub>6</sub>	h	3 d	3 v				
A <sub>1g</sub>	1	1	1	1	1	1	1	1	1	1	1	1		R <sub>z</sub>	x <sup>2</sup> +y <sup>2</sup> ,z <sup>2</sup>	
A <sub>2g</sub>	1	1	1	1	-1	-1	1	1	1	1	-1	-1				
B <sub>1g</sub>	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1				
B <sub>2g</sub>	1	-1	1	-1	-1	1	1	-1	1	-1	-1	1		(R <sub>x</sub> ,R <sub>y</sub> )	(xz,yz) (x <sup>2</sup> -y <sup>2</sup> ,xy)	
E <sub>1g</sub>	2	1	-1	-2	0	0	2	1	-1	-2	0	0				
E <sub>2g</sub>	2	-1	-1	2	0	0	2	-1	-1	2	0	0				
A <sub>1u</sub>	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1		z		
A <sub>2u</sub>	1	1	1	1	-1	-1	-1	-1	-1	-1	1	1				
B <sub>1u</sub>	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1				
B <sub>2u</sub>	1	-1	1	-1	-1	1	-1	1	-1	1	1	-1		(x,y)		
E <sub>1u</sub>	2	1	-1	-2	0	0	-2	-1	1	2	0	0				
E <sub>2u</sub>	2	-1	-1	2	0	0	-2	1	1	-2	0	0				
C-H	6	0	0	0	0	2	0	0	0	6	2	0				

b. The number of irreducible representations may be found by using the following formula:

$$n_{\text{irrep}} = \frac{1}{g} \sum_{\text{R}} \text{red}(\text{R}) \text{ irrep}(\text{R}),$$

where  $g$  = the order of the point group (24 for D<sub>6h</sub>).

$$\begin{aligned}
 n_{A_{1g}} &= \frac{1}{24} \sum_{\text{R}} \text{C-H}(\text{R}) \cdot A_{1g}(\text{R}) \\
 &= \frac{1}{24} \{ (1)(6)(1) + (2)(0)(1) + (2)(0)(1) + (1)(0)(1) \\
 &\quad + (3)(0)(1) + (3)(2)(1) + (1)(0)(1) + (2)(0)(1) \\
 &\quad + (2)(0)(1) + (1)(6)(1) + (3)(2)(1) + (3)(0)(1) \} \\
 &= 1 \\
 n_{A_{2g}} &= \frac{1}{24} \{ (1)(6)(1) + (2)(0)(1) + (2)(0)(1) + (1)(0)(1) \\
 &\quad + (3)(0)(-1) + (3)(2)(-1) + (1)(0)(1) + (2)(0)(1) \\
 &\quad + (2)(0)(1) + (1)(6)(1) + (3)(2)(-1) + (3)(0)(-1) \} \\
 &= 0 \\
 n_{B_{1g}} &= \frac{1}{24} \{ (1)(6)(1) + (2)(0)(-1) + (2)(0)(1) + (1)(0)(-1) \\
 &\quad + (3)(0)(1) + (3)(2)(-1) + (1)(0)(1) + (2)(0)(-1) \\
 &\quad + (2)(0)(1) + (1)(6)(-1) + (3)(2)(1) + (3)(0)(-1) \} \\
 &= 0
 \end{aligned}$$

$$\begin{aligned}
n_{B_{2g}} &= \frac{1}{24} \{ (1)(6)(1) + (2)(0)(-1) + (2)(0)(1) + (1)(0)(-1) \\
&\quad + (3)(0)(-1) + (3)(2)(1) + (1)(0)(1) + (2)(0)(-1) \\
&\quad + (2)(0)(1) + (1)(6)(-1) + (3)(2)(-1) + (3)(0)(1) \} \\
&= 0 \\
n_{E_{1g}} &= \frac{1}{24} \{ (1)(6)(2) + (2)(0)(1) + (2)(0)(-1) + (1)(0)(-2) \\
&\quad + (3)(0)(0) + (3)(2)(0) + (1)(0)(2) + (2)(0)(1) \\
&\quad + (2)(0)(-1) + (1)(6)(-2) + (3)(2)(0) + (3)(0)(0) \} \\
&= 0 \\
n_{E_{2g}} &= \frac{1}{24} \{ (1)(6)(2) + (2)(0)(-1) + (2)(0)(-1) + (1)(0)(2) \\
&\quad + (3)(0)(0) + (3)(2)(0) + (1)(0)(2) + (2)(0)(-1) \\
&\quad + (2)(0)(-1) + (1)(6)(2) + (3)(2)(0) + (3)(0)(0) \} \\
&= 1 \\
n_{A_{1u}} &= \frac{1}{24} \{ (1)(6)(1) + (2)(0)(1) + (2)(0)(1) + (1)(0)(1) \\
&\quad + (3)(0)(1) + (3)(2)(1) + (1)(0)(-1) + (2)(0)(-1) \\
&\quad + (2)(0)(-1) + (1)(6)(-1) + (3)(2)(-1) + (3)(0)(-1) \} \\
&= 0 \\
n_{A_{2u}} &= \frac{1}{24} \{ (1)(6)(1) + (2)(0)(1) + (2)(0)(1) + (1)(0)(1) \\
&\quad + (3)(0)(-1) + (3)(2)(-1) + (1)(0)(-1) + (2)(0)(-1) \\
&\quad + (2)(0)(-1) + (1)(6)(-1) + (3)(2)(1) + (3)(0)(1) \} \\
&= 0 \\
n_{B_{1u}} &= \frac{1}{24} \{ (1)(6)(1) + (2)(0)(-1) + (2)(0)(1) + (1)(0)(-1) \\
&\quad + (3)(0)(1) + (3)(2)(-1) + (1)(0)(-1) + (2)(0)(1) \\
&\quad + (2)(0)(-1) + (1)(6)(1) + (3)(2)(-1) + (3)(0)(1) \} \\
&= 0 \\
n_{B_{2u}} &= \frac{1}{24} \{ (1)(6)(1) + (2)(0)(-1) + (2)(0)(1) + (1)(0)(-1) \\
&\quad + (3)(0)(-1) + (3)(2)(1) + (1)(0)(-1) + (2)(0)(1) \\
&\quad + (2)(0)(-1) + (1)(6)(1) + (3)(2)(1) + (3)(0)(-1) \} \\
&= 1 \\
n_{E_{1u}} &= \frac{1}{24} \{ (1)(6)(2) + (2)(0)(1) + (2)(0)(-1) + (1)(0)(-2) \\
&\quad + (3)(0)(0) + (3)(2)(0) + (1)(0)(-2) + (2)(0)(-1) \\
&\quad + (2)(0)(1) + (1)(6)(2) + (3)(2)(0) + (3)(0)(0) \} \\
&= 1 \\
n_{E_{2u}} &= \frac{1}{24} \{ (1)(6)(2) + (2)(0)(-1) + (2)(0)(-1) + (1)(0)(2) \\
&\quad + (3)(0)(0) + (3)(2)(0) + (1)(0)(-2) + (2)(0)(1) \\
&\quad + (2)(0)(1) + (1)(6)(-2) + (3)(2)(0) + (3)(0)(0) \} \\
&= 0
\end{aligned}$$

We see that  $c-H = A_{1g} \quad E_{2g} \quad B_{2u} \quad E_{1u}$

c.  $x$  and  $y \quad E_{1u}$ ,  $z \quad A_{2u}$ , so, the ground state  $A_{1g}$  level can be excited to the degenerate  $E_{1u}$  level by coupling through the  $x$  or  $y$  transition dipoles. Therefore  $E_{1u}$  is infrared active and polarized.

d.  $(x^2 + y^2, z^2) \quad A_{1g}$ ,  $(xz, yz) \quad E_{1g}$ ,  $(x^2 - y^2, xy) \quad E_{2g}$ , so, the ground state  $A_{1g}$  level can be excited to the degenerate  $E_{2g}$  level by coupling through the  $x^2 - y^2$  or  $xy$

transitions or be excited to the degenerate  $A_{1g}$  level by coupling through the  $xz$  or  $yz$  transitions. Therefore  $A_{1g}$  and  $E_{2g}$  are Raman active..

e. The  $B_{2u}$  mode is not IR or Raman active.

$$3. \quad a. \quad \frac{d}{dr} (F r^{-1}) = F' r^{-1} - r^{-2} F$$

$$r^2 \frac{d}{dr} (F r^{-1}) = r F' - F$$

$$\frac{d}{dr} r^2 \frac{d}{dr} (F r^{-1}) = F' - F' + r F''$$

So,

$$\frac{-\hbar^2}{2\mu r^2} \frac{d}{dr} r^2 \frac{d}{dr} (F r^{-1}) = \frac{-\hbar^2}{2\mu} \frac{F''}{r}$$

Rewriting the radial Schrödinger equation with the substitution:  $R = r^{-1}F$  gives:

$$\frac{-\hbar^2}{2\mu r^2} \frac{d}{dr} r^2 \frac{d(F r^{-1})}{dr} + \frac{J(J+1)\hbar^2}{2\mu r^2} (F r^{-1}) + \frac{1}{2} k(r - r_e)^2 (F r^{-1}) = E (F r^{-1})$$

Using the above derived identity gives:

$$\frac{-\hbar^2}{2\mu} \frac{F''}{r} + \frac{J(J+1)\hbar^2}{2\mu r^2} (F r^{-1}) + \frac{1}{2} k(r - r_e)^2 (F r^{-1}) = E (F r^{-1})$$

Cancelling out an  $r^{-1}$ :

$$\frac{-\hbar^2}{2\mu} F'' + \frac{J(J+1)\hbar^2}{2\mu r^2} F + \frac{1}{2} k(r - r_e)^2 F = E F$$

$$b. \quad \frac{1}{r^2} = \frac{1}{(r_e + r)^2} = \frac{1}{r_e^2} \frac{1}{1 + \frac{r}{r_e}} = \frac{1}{r_e^2} \left( 1 - \frac{2r}{r_e} + \frac{3r^2}{r_e^2} \right)$$

So,

$$\frac{J(J+1)\hbar^2}{2\mu r^2} = \frac{J(J+1)\hbar^2}{2\mu r_e^2} \left( 1 - \frac{2r}{r_e} + \frac{3r^2}{r_e^2} \right)$$

c. Using this substitution we now have:

$$\frac{-\hbar^2}{2\mu} F'' + \frac{J(J+1)\hbar^2}{2\mu r_e^2} \left( 1 - \frac{2r}{r_e} + \frac{3r^2}{r_e^2} \right) F + \frac{1}{2} k(r - r_e)^2 F = E F$$

Now, regroup the terms which are linear and quadratic in  $r = r - r_e$ :

$$\frac{1}{2} k r^2 + \frac{J(J+1)\hbar^2}{2\mu r_e^2} \frac{3}{r_e^2} r^2 - \frac{J(J+1)\hbar^2}{2\mu r_e^2} \frac{2}{r_e} r$$

$$= \frac{1}{2} k + \frac{J(J+1)\hbar^2}{2\mu r_e^2} \frac{3}{r_e^2} r^2 - \frac{J(J+1)\hbar^2}{2\mu r_e^2} \frac{2}{r_e} r$$

Now, we must complete the square:

$$a r^2 - b r = a \left( r - \frac{b}{2a} \right)^2 - \frac{b^2}{4a}$$

So,

$$\frac{1}{2} k + \frac{J(J+1)\hbar^2}{2\mu r_e^2} \frac{3}{r_e^2} \left( r - \frac{J(J+1)\hbar^2}{2\mu r_e^2} \frac{1}{r_e} \right)^2 - \frac{\frac{J(J+1)\hbar^2}{2\mu r_e^2} \frac{1}{r_e}}{\frac{1}{2} k + \frac{J(J+1)\hbar^2}{2\mu r_e^2} \frac{3}{r_e^2}}$$

Now, redefine the first term as  $\frac{1}{2} k$ , second term as  $(r - \bar{r}_e)^2$ , and the third term as - giving:

$$\frac{1}{2} k (r - \bar{r}_e)^2 -$$

From:

$$\frac{-\hbar^2}{2\mu} F'' + \frac{J(J+1)\hbar^2}{2\mu r_e^2} \left( 1 - \frac{2}{r_e} r + \frac{3}{r_e^2} r^2 \right) F + \frac{1}{2} k (r - r_e)^2 F = E F,$$

$$\frac{-\hbar^2}{2\mu} F'' + \frac{J(J+1)\hbar^2}{2\mu r_e^2} F + \frac{J(J+1)\hbar^2}{2\mu r_e^2} \left( -\frac{2}{r_e} r + \frac{3}{r_e^2} r^2 \right) + \frac{1}{2} k r^2 F = E F,$$

and making the above substitution results in:

$$\frac{-\hbar^2}{2\mu} F'' + \frac{J(J+1)\hbar^2}{2\mu r_e^2} F + \frac{1}{2} k (r - \bar{r}_e)^2 F = E F,$$

or,

$$\frac{-\hbar^2}{2\mu} F'' + \frac{1}{2} k (r - \bar{r}_e)^2 F = E - \frac{J(J+1)\hbar^2}{2\mu r_e^2} F.$$

d. Since the above is nothing but a harmonic oscillator differential equation in  $x$  with force constant  $k$  and equilibrium bond length  $\bar{r}_e$ , we know that:

$$\frac{-\hbar^2}{2\mu} F'' + \frac{1}{2} k (r - \bar{r}_e)^2 F = E F, \text{ has energy levels:}$$

$$= \hbar \sqrt{\frac{k}{\mu}} \left( v + \frac{1}{2} \right), \quad v = 0, 1, 2, \dots$$

So,

$$E + \frac{J(J+1)\hbar^2}{2\mu r_e^2} =$$

tells us that:

$$E = \hbar \sqrt{\frac{k}{\mu}} \left( v + \frac{1}{2} \right) + \frac{J(J+1)\hbar^2}{2\mu r_e^2}.$$

As  $J$  increases,  $\bar{r}_e$  increases because of the centrifugal force pushing the two atoms apart. On the other hand  $k$  increases which indicates that the molecule finds it more difficult to

stretch against both the centrifugal and Hooke's law (spring) Harmonic force field. The total energy level (labeled by J and v) will equal a rigid rotor component  $\frac{J(J + 1)\hbar^2}{2\mu r_e^2}$  plus a

Harmonic oscillator part  $\hbar\sqrt{\frac{k}{\mu}} v + \frac{1}{2}$  (which has a force constant k which increases with J) .

## Section 5 Exercises, Problems, and Solutions

### Exercises:

1. Time dependent perturbation theory provides an expression for the radiative lifetime of an excited electronic state, given by  $\tau_R$ :

$$\tau_R = \frac{3\hbar^4 c^3}{4(E_i - E_f)^3 |\mu_{fi}|^2},$$

where  $i$  refers to the excited state,  $f$  refers to the lower state, and  $\mu_{fi}$  is the transition dipole.

a. Evaluate the z-component of the transition dipole for the  $2p_z \rightarrow 1s$  transition in a hydrogenic atom of nuclear charge  $Z$ , given:

$$\psi_{1s} = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0}\right)^{3/2} e^{-Zr/a_0}, \text{ and } \psi_{2p_z} = \frac{1}{4\sqrt{2}} \left(\frac{Z}{a_0}\right)^{5/2} r \cos\theta e^{-Zr/2a_0}.$$

Express your answer in units of  $ea_0$ .

b. Use symmetry to demonstrate that the x- and y-components of  $\mu_{fi}$  are zero, i.e.

$$\langle 2p_z | e_x | 1s \rangle = \langle 2p_z | e_y | 1s \rangle = 0.$$

c. Calculate the radiative lifetime  $\tau_R$  of a hydrogenlike atom in its  $2p_z$  state. Use the relation  $e^2 = \frac{\hbar^2}{m_e a_0}$  to simplify your results.

2. Consider a case in which the complete set of states  $\{ |k\rangle \}$  for a Hamiltonian is known.

a. If the system is initially in the state  $m$  at time  $t=0$  when a constant perturbation  $V$  is suddenly turned on, find the probability amplitudes  $C_k^{(2)}(t)$  and  $C_m^{(2)}(t)$ , to second order in  $V$ , that describe the system being in a different state  $k$  or the same state  $m$  at time  $t$ .

b. If the perturbation is turned on adiabatically, what are  $C_k^{(2)}(t)$  and  $C_m^{(2)}(t)$ ?

Here, consider that the initial time is  $t_0 \rightarrow 0^-$ , and the potential is  $V e^{-t/\tau}$ , where the positive parameter  $\tau$  is allowed to approach zero  $\rightarrow 0$  in order to describe the adiabatically (i.e., slowly) turned on perturbation.

c. Compare the results of parts a. and b. and explain any differences.

d. Ignore first order contributions (assume they vanish) and evaluate the transition rates  $\frac{d}{dt} |C_k^{(2)}(t)|^2$  for the results of part b. by taking the limit  $\tau \rightarrow 0^+$ , to obtain the adiabatic results.

3. If a system is initially in a state  $m$ , conservation of probability requires that the total probability of transitions out of state  $m$  be obtainable from the decrease in the probability of being in state  $m$ . Prove this to the lowest order by using the results of exercise 2, i.e.

$$\text{show that: } |C_m|^2 = 1 - \sum_k |C_k|^2.$$

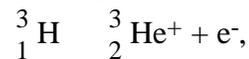
$$\sum_k |C_k|^2$$

### Problems:

1. Consider an interaction or perturbation which is carried out suddenly (instantaneously, e.g., within an interval of time  $t$  which is small compared to the natural period  $\omega_{nm}^{-1}$  corresponding to the transition from state  $m$  to state  $n$ ), and after that is turned off adiabatically (i.e., extremely slowly as  $V \rightarrow 0$ ). The transition probability in this case is given as:

$$T_{nm} = \frac{|\langle n|V|m\rangle|^2}{\hbar^2 \omega_{nm}^2}$$

where  $V$  corresponds to the maximum value of the interaction when it is turned on. This formula allows one to calculate the transition probabilities under the action of sudden perturbations which are small in absolute value whenever perturbation theory is applicable. Let's use this "sudden approximation" to calculate the probability of excitation of an electron under a sudden change of the charge of the nucleus. Consider the reaction:



and assume the tritium atom has its electron initially in a  $1s$  orbital. a. Calculate the transition probability for the transition  $1s \rightarrow 2s$  for this reaction using the above formula for the transition probability.

b. Suppose that at time  $t = 0$  the system is in a state which corresponds to the wavefunction  $\psi_m$ , which is an eigenfunction of the operator  $H_0$ . At  $t = 0$ , the sudden change of the Hamiltonian occurs (now denoted as  $H$  and remains unchanged). Calculate the same  $1s \rightarrow 2s$  transition probability as in part a., only this time as the square of the magnitude of the coefficient,  $A_{1s,2s}$  using the expansion:

$$\psi(r,0) = \psi_m(r) = \sum_n A_{nm} \psi_n(r), \text{ where } A_{nm} = \int \psi_m(r) \psi_n^*(r) d^3r$$

Note, that the eigenfunctions of  $H$  are  $\psi_n$  with eigenvalues  $E_n$ . Compare this "exact" value with that obtained by perturbation theory in part a.

2. The methyl iodide molecule is studied using microwave (pure rotational) spectroscopy. The following integral governs the rotational selection rules for transitions labeled  $J, M, K \rightarrow J', M', K'$ :

$$I = \langle D_{M'K'}^{J'} | \mu \cdot \hat{r} | D_{MK}^J \rangle.$$

The dipole moment  $\mu$  lies along the molecule's  $C_3$  symmetry axis. Let the electric field of the light define the lab-fixed  $Z$ -direction.

a. Using the fact that  $\cos \theta = D_{00}^{1*}$ , show that

$$I = 8 \mu^2 (-1)^{M+K} \begin{matrix} J' & 1 & J \\ M & 0 & M \end{matrix} \begin{matrix} J' & 1 & J \\ K & 0 & K \end{matrix} \begin{matrix} M'M \\ K'K \end{matrix}$$

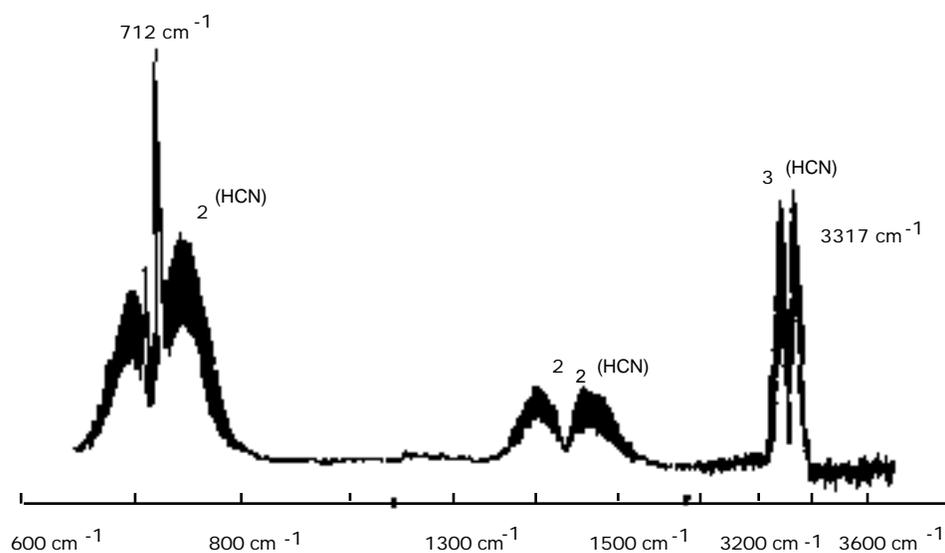
b. What restrictions does this result place on  $J = J' - J$ ? Explain physically why the  $K$  quantum number can not change.

3. Consider the molecule  $\text{BO}$ .

a. What are the total number of possible electronic states which can be formed by combination of ground state  $B$  and  $O$  atoms?

- b. What electron configurations of the molecule are likely to be low in energy? Consider all reasonable orderings of the molecular orbitals. What are the states corresponding to these configurations?
- c. What are the bond orders in each of these states?
- d. The true ground state of BO is  $^2$ . Specify the +/- and u/g symmetries for this state.
- e. Which of the excited states you derived above will radiate to the  $^2$  ground state? Consider electric dipole, magnetic dipole, and electric quadrupole radiation.
- f. Does ionization of the molecule to form a cation lead to a stronger, weaker, or equivalent bond strength?
- g. Assuming that the energies of the molecular orbitals do not change upon ionization, what are the ground state, the first excited state, and the second excited state of the positive ion?
- h. Considering only these states, predict the structure of the photoelectron spectrum you would obtain for ionization of BO.

4.



The above figure shows part of the infrared absorption spectrum of HCN gas. The molecule has a CH stretching vibration, a bending vibration, and a CN stretching vibration.

- Are any of the vibrations of linear HCN degenerate?
- To which vibration does the group of peaks between 600  $\text{cm}^{-1}$  and 800  $\text{cm}^{-1}$  belong?
- To which vibration does the group of peaks between 3200  $\text{cm}^{-1}$  and 3400  $\text{cm}^{-1}$  belong?
- What are the symmetries ( , , ) of the CH stretch, CN stretch, and bending vibrational motions?
- Starting with HCN in its 0,0,0 vibrational level, which fundamental transitions would be infrared active under parallel polarized light (i.e., z-axis polarization):

- i. 000 001?
- ii. 000 100?
- iii. 000 010?

f. Which transitions would be active when perpendicular polarized light is used?  
 g. Why does the 712 cm<sup>-1</sup> transition have a Q-branch, whereas that near 3317 cm<sup>-1</sup> has only P- and R-branches?

## Solutions

### Exercises:

1. a. Evaluate the z-component of  $\mu_{fi}$ :

$$\mu_{fi} = \langle 2p_z | e r \cos \theta | 1s \rangle, \text{ where } 1s = \frac{1}{\sqrt{\pi}} \left( \frac{Z}{a_0} \right)^{3/2} e^{-Zr/a_0}, \text{ and } 2p_z = \frac{1}{4\sqrt{2}} \left( \frac{Z}{a_0} \right)^{5/2} r \cos \theta e^{-Zr/2a_0}.$$

$$\begin{aligned} \mu_{fi} &= \frac{1}{4\sqrt{2}} \left( \frac{Z}{a_0} \right)^{5/2} \frac{1}{\sqrt{\pi}} \left( \frac{Z}{a_0} \right)^{3/2} \langle r \cos \theta e^{-Zr/2a_0} | e r \cos \theta e^{-Zr/a_0} \rangle \\ &= \frac{1}{4\sqrt{2}} \left( \frac{Z}{a_0} \right)^4 \langle r \cos \theta e^{-Zr/2a_0} | e r \cos \theta e^{-Zr/a_0} \rangle \\ &= \frac{e}{4\sqrt{2}} \left( \frac{Z}{a_0} \right)^4 \int_0^\infty r^2 dr \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi r^2 e^{-Zr/2a_0} e^{-Zr/a_0} \cos^2 \theta \\ &= \frac{e}{4\sqrt{2}} \left( \frac{Z}{a_0} \right)^4 \int_0^\infty r^4 e^{-3Zr/2a_0} dr \int_0^\pi \sin \theta \cos^2 \theta d\theta \int_0^{2\pi} d\phi \end{aligned}$$

Using integral equation 4 to integrate over r and equation 17 to integrate over  $\theta$  we obtain:

$$\begin{aligned} &= \frac{e}{4\sqrt{2}} \left( \frac{Z}{a_0} \right)^4 \frac{4!}{3Z^5} \frac{-1}{3} \int_0^{2\pi} \cos^3 \theta d\theta \\ &= \frac{e}{4\sqrt{2}} \left( \frac{Z}{a_0} \right)^4 \frac{2^5 a_0^5 4!}{3^5 Z^5} \frac{-1}{3} ((-1)^3 - (1)^3) \\ &= \frac{e}{\sqrt{2}} \frac{2^8 a_0}{3^5 Z} = \frac{ea_0}{Z} \frac{2^8}{\sqrt{2} 3^5} = 0.7449 \frac{ea_0}{Z} \end{aligned}$$

b. Examine the symmetry of the integrands for  $\langle 2p_z | e^x | 1s \rangle$  and  $\langle 2p_z | e^y | 1s \rangle$ . Consider reflection in the xy plane:

Function	Symmetry
$2p_z$	-1
x	+1
1s	+1
y	+1

Under this operation the integrand of  $\langle 2p_z | e^x | 1s \rangle$  is  $(-1)(1)(1) = -1$  (it is antisymmetric) and hence  $\langle 2p_z | e^x | 1s \rangle = 0$ .

Similarly, under this operation the integrand of  $\langle 2p_z | e^y | 1s \rangle$  is  $(-1)(1)(1) = -1$  (it is also antisymmetric) and hence  $\langle 2p_z | e^y | 1s \rangle = 0$ .

$$c. \quad R = \frac{3\hbar^4 c^3}{4(E_i - E_f)^3 |\mu_{fi}|^2},$$

$$E_i = E_{2p_z} = -\frac{1}{4} Z^2 \frac{e^2}{2a_0}$$

$$E_f = E_{1s} = -Z^2 \frac{e^2}{2a_0}$$

$$E_i - E_f = \frac{3}{8} \frac{e^2}{a_0} Z^2$$

Making the substitutions for  $E_i - E_f$  and  $|\mu_{fi}|$  in the expression for  $R$  we obtain:

$$R = \frac{3\hbar^4 c^3}{4 \left( \frac{3}{8} \frac{e^2}{a_0} Z^2 \right)^3 \left( \frac{ea_0}{Z} \frac{2^8}{\sqrt{2} 3^5} \right)^2},$$

$$= \frac{3\hbar^4 c^3}{4 \frac{3^3}{8^3} \frac{e^6}{a_0^3} Z^6 \frac{e^2 a_0^2}{Z^2} \frac{2^{16}}{(2)3^{10}}},$$

$$= \frac{\hbar^4 c^3 3^8 a_0}{e^8 Z^4 2^8},$$

Inserting  $e^2 = \frac{\hbar^2}{m_e a_0}$  we obtain:

$$R = \frac{\hbar^4 c^3 3^8 a_0 m_e^4 a_0^4}{\hbar^8 Z^4 2^8} = \frac{3^8 c^3 a_0^5 m_e^4}{2^8 \hbar^4 Z^4}$$

$$= 25.6289 \frac{c^3 a_0^5 m_e^4}{\hbar^4 Z^4}$$

$$= 25,6289 \frac{1}{Z^4} \times \frac{(2.998 \times 10^{10} \text{ cm sec}^{-1})^3 (0.529177 \times 10^{-8} \text{ cm})^5 (9.109 \times 10^{-28} \text{ g})^4}{(1.0546 \times 10^{-27} \text{ g cm}^2 \text{ sec}^{-1})^4}$$

$$= 1.595 \times 10^{-9} \text{ sec} \times \frac{1}{Z^4}$$

So, for example:

Atom	R
H	1.595 ns
He <sup>+</sup>	99.7 ps
Li <sup>+2</sup>	19.7 ps
Be <sup>+3</sup>	6.23 ps
Ne <sup>+9</sup>	159 fs

2. a.  $H = H_0 + H'(t)$ ,  $H'(t) = V(t)$ ,  $H_0 \psi_k = E_k \psi_k$ ,  $\psi_k = E_k/\hbar$

$$i\hbar \frac{d}{dt} \psi = H \psi$$

let  $\psi(r,t) = \sum_j c_j(t) \psi_j e^{-i E_j t/\hbar}$  and insert into the above expression:

$$i\hbar \sum_j \dot{c}_j \psi_j e^{-i E_j t/\hbar} - \sum_j c_j E_j \psi_j e^{-i E_j t/\hbar} = i\hbar \sum_j \dot{c}_j \psi_j e^{-i E_j t/\hbar} (H_0 + H'(t)) \psi_j$$

$$i\hbar \sum_j \dot{c}_j \psi_j e^{-i E_j t/\hbar} + \sum_j c_j E_j \psi_j e^{-i E_j t/\hbar} - \sum_j c_j E_j \psi_j e^{-i E_j t/\hbar} - \sum_j c_j H' \psi_j e^{-i E_j t/\hbar} = 0$$

$$i\hbar \sum_j \dot{c}_j \langle m | \psi_j \rangle - \sum_j c_j \langle m | H' | \psi_j \rangle e^{-i E_j t/\hbar} = 0$$

$$i\hbar \dot{c}_m e^{-i E_m t/\hbar} = \sum_j c_j H'_{mj} e^{-i E_j t/\hbar}$$

So,

$$c_m = \frac{1}{i\hbar} \sum_j c_j H'_{mj} e^{-i(E_j - E_m)t/\hbar}$$

Going back a few equations and multiplying from the left by  $\psi_k$  instead of  $\psi_m$  we obtain:

$$i\hbar \sum_j \dot{c}_j \langle k | \psi_j \rangle - \sum_j c_j \langle k | H' | \psi_j \rangle e^{-i E_j t/\hbar} = 0$$

$$i\hbar \dot{c}_k e^{-i E_k t/\hbar} = \sum_j c_j H'_{kj} e^{-i E_j t/\hbar}$$

So,

$$c_k = \frac{1}{i\hbar} \sum_j c_j H'_{kj} e^{-i(jk)t}$$

Now, let:

$$c_m = c_m^{(0)} + c_m^{(1)} + c_m^{(2)} + \dots$$

$$c_k = c_k^{(0)} + c_k^{(1)} + c_k^{(2)} + \dots$$

and substituting into above we obtain:

$$c_m^{(0)} + c_m^{(1)} + c_m^{(2)} + \dots = \frac{1}{i\hbar} \sum_j [c_j^{(0)} + c_j^{(1)} + c_j^{(2)} + \dots] H'_{mj} e^{-i(jm)t}$$

first order:

$$c_m^{(0)} = 0 \quad c_m^{(0)} = 1$$

second order:

$$c_m^{(1)} = \frac{1}{i\hbar} \sum_j c_j^{(0)} H'_{mj} e^{-i(jm)t}$$

(n+1)<sup>st</sup> order:

$$c_m^{(n)} = \frac{1}{i\hbar} \sum_j c_j^{(n-1)} H'_{mj} e^{-i(jm)t}$$

Similarly:

first order:

$$c_k^{(0)} = 0 \quad c_k^{(0)} = 0$$

second order:

$$c_k^{(1)} = \frac{1}{i\hbar} \sum_j c_j^{(0)} H'_{kj} e^{-i(jk)t}$$

(n+1)<sup>st</sup> order:

$$c_k^{(n)} = \frac{1}{i\hbar} \sum_j c_j^{(n-1)} H'_{kj} e^{-i(jk)t}$$

So,

$$c_m^{(1)} = \frac{1}{i\hbar} c_m^{(0)} H'_{mm} e^{-i(mm)t} = \frac{1}{i\hbar} H'_{mm}$$

$$c_m^{(1)}(t) = \frac{1}{i\hbar} \int_0^t dt' V_{mm} = \frac{V_{mm}t}{i\hbar}$$

and similarly,

$$c_k^{(1)} = \frac{1}{i\hbar} c_m^{(0)} H'_{km} e^{-i(km)t} = \frac{1}{i\hbar} H'_{km} e^{-i(km)t}$$

$$c_k^{(1)}(t) = \frac{1}{i\hbar} \int_0^t dt' V_{km} e^{-i(km)t'} = \frac{V_{km}}{i\hbar} [e^{-i(km)t} - 1]$$

$$c_m^{(2)} = \frac{1}{i\hbar} \sum_j c_j^{(1)} H'_{mj} e^{-i(\epsilon_j - \epsilon_m)t}$$

$$c_m^{(2)} = \sum_j \frac{1}{i\hbar} \frac{V_{jm}}{\hbar} [e^{-i(\epsilon_j - \epsilon_m)t} - 1] H'_{mj} e^{-i(\epsilon_j - \epsilon_m)t} + \frac{1}{i\hbar} \frac{V_{mm}t}{\hbar} H'_{mm}$$

$$c_m^{(2)} = \sum_j \frac{1}{i\hbar} \frac{V_{jm}V_{mj}}{\hbar} \int_0^t dt' e^{-i(\epsilon_j - \epsilon_m)t'} [e^{-i(\epsilon_j - \epsilon_m)t'} - 1] - \frac{V_{mm}V_{mm}}{\hbar^2} \int_0^t t' dt'$$

$$= \sum_j \frac{V_{jm}V_{mj}}{\hbar^2} \int_0^t dt' [1 - e^{-i(\epsilon_j - \epsilon_m)t'}] - \frac{|V_{mm}|^2}{\hbar^2} \frac{t^2}{2}$$

$$= \sum_j \frac{V_{jm}V_{mj}}{\hbar^2} t - \frac{e^{-i(\epsilon_j - \epsilon_m)t} - 1}{-i(\epsilon_j - \epsilon_m)} - \frac{|V_{mm}|^2}{\hbar^2} \frac{t^2}{2}$$

$$= \sum_j \frac{V_{jm}V_{mj}}{\hbar^2} (e^{-i(\epsilon_j - \epsilon_m)t} - 1) + \sum_j \frac{V_{jm}V_{mj}}{\hbar^2} t - \frac{|V_{mm}|^2}{2\hbar^2} t^2$$

Similarly,

$$c_k^{(2)} = \frac{1}{i\hbar} \sum_j c_j^{(1)} H'_{kj} e^{-i(\epsilon_j - \epsilon_k)t}$$

$$= \sum_j \frac{1}{i\hbar} \frac{V_{jm}}{\hbar} [e^{-i(\epsilon_j - \epsilon_m)t} - 1] H'_{kj} e^{-i(\epsilon_j - \epsilon_k)t} +$$

$$\frac{1}{i\hbar} \frac{V_{mm}t}{\hbar} H'_{km} e^{-i(\epsilon_m - \epsilon_k)t}$$

$$c_k^{(2)}(t) = \sum_j \frac{V_{jm}V_{kj}}{\hbar^2} \int_0^t dt' e^{-i(\epsilon_j - \epsilon_k)t'} [e^{-i(\epsilon_j - \epsilon_m)t'} - 1]$$

$$- \frac{V_{mm}V_{km}}{\hbar^2} \int_0^t t' dt' e^{-i(\epsilon_m - \epsilon_k)t'}$$

$$= \sum_j \frac{V_{jm}V_{kj}}{\hbar^2} \frac{e^{-i(\epsilon_j - \epsilon_m)t} - 1}{-i(\epsilon_j - \epsilon_m)} - \frac{e^{-i(\epsilon_m - \epsilon_k)t} - 1}{-i(\epsilon_m - \epsilon_k)}$$

$$- \frac{V_{mm}V_{km}}{\hbar^2} e^{-i(\epsilon_m - \epsilon_k)t} \frac{t}{-i(\epsilon_m - \epsilon_k)} - \frac{1}{-(i(\epsilon_m - \epsilon_k))^2} t$$

$$= \sum_j \frac{V_{jm}V_{kj}}{\hbar^2} \frac{e^{-i(\epsilon_j - \epsilon_m)t} - 1}{\epsilon_j - \epsilon_m} - \frac{e^{-i(\epsilon_m - \epsilon_k)t} - 1}{\epsilon_m - \epsilon_k}$$

$$\begin{aligned}
& + \frac{V_{mm}V_{km}}{\hbar^2} e^{-i(mk)t} \frac{t}{i} - \frac{1}{mk} \frac{t}{0} \\
= & \sum_j \frac{V_{jm}V_{kj}}{E_m - E_j} \frac{e^{-i(mk)t} - 1}{E_m - E_k} - \frac{e^{-i(jk)t} - 1}{E_j - E_k} \\
& + \frac{V_{mm}V_{km}}{\hbar(E_m - E_k)} e^{-i(mk)t} \frac{t}{i} - \frac{1}{mk} + \frac{1}{mk}
\end{aligned}$$

So, the overall amplitudes  $c_m$ , and  $c_k$ , to second order are:

$$\begin{aligned}
c_m(t) = & 1 + \frac{V_{mm}t}{i\hbar} + \sum_j \frac{V_{jm}V_{mj}}{i\hbar(E_m - E_j)} t + \\
& \sum_j \frac{V_{jm}V_{mj}}{\hbar^2(E_m - E_j)^2} (e^{-i(jm)t} - 1) - \frac{|V_{mm}|^2 t^2}{2\hbar^2} \\
c_k(t) = & \frac{V_{km}}{(E_m - E_k)} [e^{-i(mk)t} - 1] + \\
& \frac{V_{mm}V_{km}}{(E_m - E_k)^2} [1 - e^{-i(mk)t}] + \frac{V_{mm}V_{km}}{(E_m - E_k)} \frac{t}{i\hbar} e^{-i(mk)t} + \\
& \sum_j \frac{V_{jm}V_{kj}}{E_m - E_j} \frac{e^{-i(mk)t} - 1}{E_m - E_k} - \frac{e^{-i(jk)t} - 1}{E_j - E_k}
\end{aligned}$$

b. The perturbation equations still hold:

$$c_m^{(n)} = \frac{1}{i\hbar} \sum_j c_j^{(n-1)} H'_{mj} e^{-i(jm)t} ; c_k^{(n)} = \frac{1}{i\hbar} \sum_j c_j^{(n-1)} H'_{kj} e^{-i(jk)t}$$

So,  $c_m^{(0)} = 1$  and  $c_k^{(0)} = 0$

$$c_m^{(1)} = \frac{1}{i\hbar} H'_{mm} t$$

$$c_m^{(1)} = \frac{1}{i\hbar} \int_0^t V_{mm} dt' e^{-i(mm)t'} = \frac{V_{mm}e^{-i(mm)t}}{i\hbar}$$

$$c_k^{(1)} = \frac{1}{i\hbar} H'_{km} e^{-i(mk)t}$$

$$\begin{aligned}
c_k^{(1)} = & \frac{1}{i\hbar} \int_0^t V_{km} dt' e^{-i(mk)t'} = \frac{V_{km}}{i\hbar(-i(mk))} [e^{-i(mk)t} - 1] \\
= & \frac{V_{km}}{E_m - E_k + i\hbar} [e^{-i(mk)t} - 1]
\end{aligned}$$

$$c_m^{(2)} = \sum_j \frac{1}{i\hbar} \frac{V_{jm}}{E_m - E_j + i\hbar} e^{-i(mj)t} V_{mj} e^{-i(jm)t} +$$

$$c_m^{(2)} = \int_j^m \frac{1}{i\hbar} \frac{V_{jm}V_{mj}}{E_m - E_j + i\hbar} e^{2t} dt' - \frac{|V_{mm}|^2}{\hbar^2} e^{2t} dt'$$

$$= \int_j^m \frac{V_{jm}V_{mj}}{i\hbar^2 (E_m - E_j + i\hbar)} e^{2t} dt' - \frac{|V_{mm}|^2}{2\hbar^2} e^{2t}$$

$$c_k^{(2)} = \int_j^m \frac{1}{i\hbar} \frac{V_{jm}}{E_m - E_j + i\hbar} e^{-i(mj+)t} H'_{kj} e^{-i(jk)t} + \frac{1}{i\hbar} \frac{V_{mm} e^t}{\hbar} H'_{km} e^{-i(mk)t}$$

$$c_k^{(2)} = \int_j^m \frac{1}{i\hbar} \frac{V_{jm}V_{kj}}{E_m - E_j + i\hbar} e^{-i(mk+2)t} dt' - \frac{V_{mm}V_{km}}{\hbar^2} e^{-i(mk+2)t} dt'$$

$$= \int_j^m \frac{V_{jm}V_{kj} e^{-i(mk+2)t}}{(E_m - E_j + i\hbar)(E_m - E_k + 2i\hbar)} - \frac{V_{mm}V_{km} e^{-i(mk+2)t}}{i\hbar (E_m - E_k + 2i\hbar)}$$

Therefore, to second order:

$$c_m(t) = 1 + \frac{V_{mm}e^t}{i\hbar} + \int_j^m \frac{V_{jm}V_{mj}}{i\hbar^2 (E_m - E_j + i\hbar)} e^{2t}$$

$$c_k(t) = \frac{V_{km}}{i\hbar(-i(mk+))} [e^{-i(mk+)t}] + \int_j^m \frac{V_{jm}V_{kj} e^{-i(mk+2)t}}{(E_m - E_j + i\hbar)(E_m - E_k + 2i\hbar)}$$

c. In part a. the  $c^{(2)}(t)$  grow linearly with time (for  $V_{mm} = 0$ ) while in part b. they remain finite for  $t > 0$ . The result in part a. is due to the sudden turning on of the field.

$$d. |c_k(t)|^2 = \int_{jj'} \frac{V_{kj}V_{kj'}V_{jm}V_{j'm} e^{-i(mk+2)t} e^{i(mk+2)t}}{(E_m - E_j + i\hbar)(E_m - E_j' - i\hbar)(E_m - E_k + 2i\hbar)(E_m - E_k - 2i\hbar)}$$

$$= \frac{V_{kj}V_{kj}^*V_{jm}V_{jm}^* e^{4t}}{jj' [(E_m-E_j)(E_m-E_j')+i\hbar (E_j-E_j')+ \hbar^2 - 2][(E_m-E_k)^2+4\hbar^2 - 2]}$$

$$\frac{d}{dt} |c_k(t)|^2 = \frac{4 V_{kj}V_{kj}^*V_{jm}V_{jm}^*}{jj' [(E_m-E_j)(E_m-E_j')+i\hbar (E_j-E_j')+ \hbar^2 - 2][(E_m-E_k)^2+4\hbar^2 - 2]}$$

Now, look at the limit as  $\hbar \rightarrow 0^+$ :

$$\frac{d}{dt} |c_k(t)|^2 \rightarrow 0 \text{ when } E_m = E_k$$

$$\lim_{0^+} \frac{4}{((E_m-E_k)^2+4\hbar^2 - 2)} (E_m-E_k)$$

So, the final result is the 2<sup>nd</sup> order golden rule expression:

$$\frac{d}{dt} |c_k(t)|^2 \frac{2}{\hbar} (E_m-E_k) \lim_{0^+} \frac{V_{jm}V_{kj}}{j (E_j - E_m - i\hbar)}$$

3. For the sudden perturbation case:

$$|c_m(t)|^2 = 1 + \sum_j \frac{V_{jm}V_{mj}}{(E_m - E_j)^2} [e^{-i(jm)t} - 1 + e^{i(jm)t} - 1] + O(V^3)$$

$$|c_m(t)|^2 = 1 + \sum_j \frac{V_{jm}V_{mj}}{(E_m - E_j)^2} [e^{-i(jm)t} + e^{i(jm)t} - 2] + O(V^3)$$

$$|c_k(t)|^2 = \frac{V_{km}V_{mk}}{(E_m - E_k)^2} [-e^{-i(km)t} - e^{i(km)t} + 2] + O(V^3)$$

$$1 - \sum_k |c_k(t)|^2 = 1 - \sum_k \frac{V_{km}V_{mk}}{(E_m - E_k)^2} [-e^{-i(km)t} - e^{i(km)t} + 2] + O(V^3)$$

$$= 1 + \sum_k \frac{V_{km}V_{mk}}{(E_m - E_k)^2} [e^{-i(km)t} + e^{i(km)t} - 2] + O(V^3)$$

to order  $V^2$ ,  $|c_m(t)|^2 = 1 - \sum_k |c_k(t)|^2$ , with no assumptions made regarding  $V_{mm}$ .

k

For the adiabatic perturbation case:

$$|c_m(t)|^2 = 1 + \sum_j \frac{V_{jm}V_{mj}e^{2t}}{i\hbar^2 (E_m - E_j + i\hbar)} + \frac{V_{jm}V_{mj}e^{2t}}{-i\hbar^2 (E_m - E_j - i\hbar)} + O(V^3)$$

$$= 1 + \sum_j \frac{1}{i\hbar^2} \frac{1}{(E_m-E_j+i\hbar)} - \frac{1}{(E_m-E_j-i\hbar)} V_{jm}V_{mj}e^{2t} + O(V^3)$$

$$= 1 + \sum_j \frac{1}{i\hbar^2} \frac{-2i\hbar}{(E_m-E_j)^2+\hbar^2} V_{jm}V_{mj}e^{2t} + O(V^3)$$

$$= 1 - \sum_{j \neq m} \frac{V_{jm} V_{mj} e^{2i t}}{(E_m - E_j)^2 + \hbar^2} + O(V^3)$$

$$|c_k(t)|^2 = \frac{V_{km} V_{mk}}{(E_m - E_k)^2 + \hbar^2} e^{2i t} + O(V^3)$$

to order  $V^2$ ,  $|c_m(t)|^2 = 1 - \sum_k |c_k(t)|^2$ , with no assumptions made regarding  $V_{mm}$  for this case as well.

Problems:

1. a.  $T_{nm} = \frac{|\langle n|V|m\rangle|^2}{\hbar^2 \omega_{nm}^2}$

evaluating  $\langle 1s|V|2s\rangle$  (using only the radial portions of the 1s and 2s wavefunctions since the spherical harmonics will integrate to unity) where  $V = (e^2/r)$ :

$$\langle 1s|V|2s\rangle = \int_0^\infty 2 \frac{Z}{a_0} \frac{3}{2} e^{-Zr/a_0} \frac{1}{r} \frac{1}{\sqrt{2}} \frac{Z}{a_0} \frac{3}{2} \left(1 - \frac{Zr}{2a_0}\right) e^{-Zr/2a_0} r^2 dr$$

$$\langle 1s|V|2s\rangle = \frac{2}{\sqrt{2}} \frac{Z}{a_0} \int_0^\infty r e^{-3Zr/2a_0} dr - \frac{Zr^2}{2a_0} e^{-3Zr/2a_0} dr$$

Using integral equation 4 for the two integrations we obtain:

$$\langle 1s|V|2s\rangle = \frac{2}{\sqrt{2}} \frac{Z}{a_0} \left[ \frac{1}{3Z/2} - \frac{Z}{2a_0} \frac{2}{3Z/3} \right]$$

$$\langle 1s|V|2s\rangle = \frac{2}{\sqrt{2}} \frac{Z}{a_0} \left[ \frac{2^2 a_0^2}{3^2 Z^2} - \frac{2^3 a_0^2}{3^3 Z^2} \right]$$

$$\langle 1s|V|2s\rangle = \frac{2}{\sqrt{2}} \frac{Z}{a_0} \frac{(3)2^2 a_0^2 - 2^3 a_0^2}{3^3 Z^2} = \frac{8Z}{\sqrt{2} 27 a_0}$$

Now,

$$E_n = -\frac{Z^2 e^2}{n^2 a_0}, E_{1s} = -\frac{Z^2 e^2}{a_0}, E_{2s} = -\frac{Z^2 e^2}{8a_0}, E_{2s} - E_{1s} = \frac{3Z^2 e^2}{8a_0}$$

So,

$$T_{nm} = \frac{\left(\frac{8Z}{\sqrt{2} 27 a_0}\right)^2}{\frac{3Z^2}{8a_0} \frac{2}{2}} = \frac{2^6 Z^2 2^6 a_0^2}{(2) 3^8 a_0^2 Z^4} = \frac{2^{11}}{3^8 Z^2} = 0.312 \text{ (for } Z = 1)$$

b.  $m(r) = 1s = 2 \frac{Z}{a_0} \frac{3}{2} e^{-Zr/a_0} Y_{00}$

The orthogonality of the spherical harmonics results in only s-states having non-zero values for  $A_{nm}$ . We can then drop the  $Y_{00}$  (integrating this term will only result in unity) in determining the value of  $A_{1s,2s}$ .

$$\psi_{2s}(r) = \frac{1}{\sqrt{2}} \left( \frac{Z}{a_0} \right)^{3/2} \left( 1 - \frac{Zr}{2a_0} \right) e^{-\frac{Zr}{2a_0}}$$

Remember for  $1s$   $Z = 1$  and for  $2s$   $Z = 2$

$$A_{nm} = \int_0^\infty \frac{1}{\sqrt{2}} \left( \frac{Z}{a_0} \right)^{3/2} e^{-\frac{Zr}{a_0}} \frac{1}{\sqrt{2}} \left( \frac{Z+1}{a_0} \right)^{3/2} \left( 1 - \frac{(Z+1)r}{2a_0} \right) e^{-\frac{(Z+1)r}{2a_0}} r^2 dr$$

$$A_{nm} = \frac{2}{\sqrt{2}} \left( \frac{Z}{a_0} \right)^{3/2} \left( \frac{Z+1}{a_0} \right)^{3/2} \int_0^\infty e^{-\frac{(3Z+1)r}{2a_0}} \left( 1 - \frac{(Z+1)r}{2a_0} \right) r^2 dr$$

$$A_{nm} = \frac{2}{\sqrt{2}} \left( \frac{Z}{a_0} \right)^{3/2} \left( \frac{Z+1}{a_0} \right)^{3/2} \int_0^\infty r^2 e^{-\frac{(3Z+1)r}{2a_0}} dr - \frac{(Z+1)r^3}{2a_0} e^{-\frac{(3Z+1)r}{2a_0}} dr$$

Evaluating these integrals using integral equation 4 we obtain:

$$A_{nm} = \frac{2}{\sqrt{2}} \left( \frac{Z}{a_0} \right)^{3/2} \left( \frac{Z+1}{a_0} \right)^{3/2} \left[ \frac{2}{\frac{(3Z+1)^3}{2a_0}} - \frac{Z+1}{2a_0} \frac{(3)(2)}{\frac{(3Z+1)^4}{2a_0}} \right]$$

$$A_{nm} = \frac{2}{\sqrt{2}} \left( \frac{Z}{a_0} \right)^{3/2} \left( \frac{Z+1}{a_0} \right)^{3/2} \left[ \frac{2^4 a_0^3}{(3Z+1)^3} - (Z+1) \frac{(3)2^4 a_0^3}{(3Z+1)^4} \right]$$

$$A_{nm} = \frac{2}{\sqrt{2}} \left( \frac{Z}{a_0} \right)^{3/2} \left( \frac{Z+1}{a_0} \right)^{3/2} \frac{-2^5 a_0^3}{(3Z+1)^4}$$

$$A_{nm} = -2 \frac{[2^3 Z(Z+1)]^{3/2}}{(3Z+1)^4}$$

The transition probability is the square of this amplitude:

$$T_{nm} = -2 \frac{[2^3 Z(Z+1)]^{3/2}}{(3Z+1)^4} = \frac{2^{11} Z^3 (Z+1)^3}{(3Z+1)^8} = 0.25 \text{ (for } Z = 1\text{)}.$$

The difference in these two results (parts a. and b.) will become negligible at large values of  $Z$  when the perturbation becomes less significant as in the case of  $Z = 1$ .

2.  $\hat{\mu}$  is along  $Z$  (lab fixed), and  $\mu$  is along  $z$  (the C-I molecule fixed bond). The angle between  $Z$  and  $z$  is  $\theta$ :

$$\hat{\mu} \cdot \mu = \mu \cos \theta = \mu D_{00}^{1*}(\theta)$$

So,

$$\begin{aligned} I &= \langle D_{M'K'}^{J'} | \hat{\mu} \cdot \mu | D_{MK}^J \rangle = \langle D_{M'K'}^{J'} | \mu D_{MK}^J \sin \theta | D_{MK}^J \rangle \\ &= \mu \langle D_{M'K'}^{J'} | D_{00}^{1*} | D_{MK}^J \rangle \sin \theta \int d\Omega \end{aligned}$$

Now use:

$$D_{M'n'}^{J'*} D_{00}^{1*} = \sum_{jmn} \langle J'M'10|jm\rangle^* D_{mn}^{j*} \langle jn|J'K'10\rangle^*$$

to obtain:

$$I = \mu \sum_{jmn} \langle J'M'10|jm\rangle^* \langle jn|J'K'10\rangle^* D_{mn}^{j*} D_{MK}^J \sin^2 \theta \cos^2 \theta$$

Now use:

$$D_{mn}^{j*} D_{MK}^J \sin^2 \theta \cos^2 \theta = \frac{8}{2J+1} J_j M_m n,$$

to obtain:

$$\begin{aligned} I &= \mu \frac{8}{2J+1} \sum_{jmn} \langle J'M'10|jm\rangle^* \langle jn|J'K'10\rangle^* J_j M_m n \\ &= \mu \frac{8}{2J+1} \langle J'M'10|JM\rangle \langle JK|J'K'10\rangle. \end{aligned}$$

We use:

$$\langle JK|J'K'10\rangle = \sqrt{2J+1} (-i)^{(J'-1+K)} \begin{matrix} J' & 1 & J \\ & K' & 0 & K \end{matrix}$$

and,

$$\langle J'M'10|JM\rangle = \sqrt{2J+1} (-i)^{(J'-1+M)} \begin{matrix} J' & 1 & J \\ & M' & 0 & M \end{matrix}$$

to give:

$$\begin{aligned} I &= \mu \frac{8}{2J+1} \sqrt{2J+1} (-i)^{(J'-1+M)} \begin{matrix} J' & 1 & J \\ & M' & 0 & M \end{matrix} \sqrt{2J+1} (-i)^{(J'-1+K)} \begin{matrix} J' & 1 & J \\ & K' & 0 & K \end{matrix} \\ &= \mu 8^2 (-i)^{(J'-1+M+J'-1+K)} \begin{matrix} J' & 1 & J \\ & M' & 0 & M \end{matrix} \begin{matrix} J' & 1 & J \\ & K' & 0 & K \end{matrix} \\ &= \mu 8^2 (-i)^{(M+K)} \begin{matrix} J' & 1 & J \\ & M' & 0 & M \end{matrix} \begin{matrix} J' & 1 & J \\ & K' & 0 & K \end{matrix} \end{aligned}$$

The 3-J symbols vanish unless:  $K' + 0 = K$  and  $M' + 0 = M$ .

So,

$$I = \mu 8^2 (-i)^{(M+K)} \begin{matrix} J' & 1 & J \\ & M & 0 & M \end{matrix} \begin{matrix} J' & 1 & J \\ & K & 0 & K \end{matrix} M'M K'K.$$

b.  $\begin{matrix} J' & 1 & J \\ & M & 0 & M \end{matrix}$  and  $\begin{matrix} J' & 1 & J \\ & K & 0 & K \end{matrix}$  vanish unless  $J' = J + 1, J, J - 1$   
 $J = \pm 1, 0$

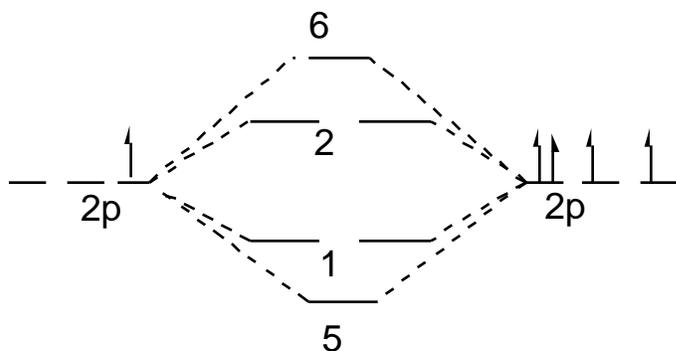
The K quantum number can not change because the dipole moment lies along the molecule's  $C_3$  axis and the light's electric field thus can exert no torque that twists the molecule about this axis. As a result, the light can not induce transitions that excite the molecule's spinning motion about this axis.

3. a. B atom:  $1s^2 2s^2 2p^1$ ,  $^2P$  ground state  $L = 1, S = \frac{1}{2}$ , gives a degeneracy  $((2L+1)(2S+1))$  of 6.

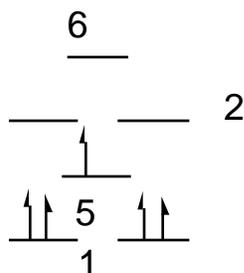
O atom:  $1s^2 2s^2 2p^4$ ,  $^3P$  ground state  $L = 1, S = 1$ , gives a degeneracy  $((2L+1)(2S+1))$  of 9.

The total number of states formed is then  $(6)(9) = 54$ .

b. We need only consider the p orbitals to find the low lying molecular states:



Which, in reality look like this:



This is the correct ordering to give a  $^2 +$  ground state. The only low-lying electron configurations are  $1^3 5^2$  or  $1^4 5^1$ . These lead to  $^2 -$  and  $^2 +$  states, respectively.

c. The bond orders in both states are  $2\frac{1}{2}$ .

d. The  $^2 -$  is + and g/u cannot be specified since this is a heteronuclear molecule.

e. Only one excited state, the  $^2 -$ , is spin-allowed to radiate to the  $^2 +$ . Consider symmetries of transition moment operators that arise in the E1, E2 and M1 contributions to the transition rate

Electric dipole allowed:  $z^+, x, y^-$ , the  $^2 -$   $^2 +$  is electric dipole allowed via a perpendicular band.

Magnetic dipole allowed:  $R_z^-, R_{x,y}^+$ , the  $^2 -$   $^2 +$  is magnetic dipole allowed.

Electric quadrupole allowed:  $x^2+y^2, z^2^+, xy, yz^-, x^2-y^2, xy^2$  the  $^2 -$   $^2 +$  is electric quadrupole allowed as well.

f. Since ionization will remove a bonding electron, the  $BO^+$  bond is weaker than the BO bond.

g. The ground state  $BO^+$  is  $1^+$  corresponding to a  $1^4$  electron configuration. An electron configuration of  $1^3 5^1$  leads to a  $^3 -$  and a  $^1 -$  state. The  $^3 -$  will be lower in energy. A  $1^2 5^2$  configuration will lead to higher lying states of  $^3 -$ ,  $^1 -$ , and  $^1 +$ .

h. There should be 3 bands corresponding to formation of  $BO^+$  in the  $1^+, ^3 -$ , and  $^1 -$  states. Since each of these involves removing a bonding electron, the Franck-Condon integrals will be appreciable for several vibrational levels, and thus a vibrational progression should be observed.

4. a. The bending ( $\nu_2$ ) vibration is degenerate.

b.  $\text{H}-\text{C} \equiv \text{N}$

bending fundamental

c.  $\text{H}-\text{C} \equiv \text{N}$

stretching fundamental

d. CH stretch ( $\nu_3$  in figure) is  $\nu_3$ , CN stretch is  $\nu_1$ , and HCN ( $\nu_2$  in figure) bend is  $\nu_2$ .

e. Under z ( $\nu_2$ ) light the CN stretch and the CH stretch can be excited, since  $\nu_0 = 0$ ,  $\nu_1 = 1$  and z provides coupling.

f. Under x,y ( $\nu_2$ ) light the HCN bend can be excited, since  $\nu_0 = 0$ ,  $\nu_1 = 1$  and x,y provides coupling.

g. The bending vibration is active under (x,y) perpendicular polarized light.  $J = 0, \pm 1$  are the selection rules for  $\perp$  transitions. The CH stretching vibration is active under (z)  $\parallel$  polarized light.  $J = \pm 1$  are the selection rules for  $\parallel$  transitions.

## Section 6 Exercises, Problems, and Solutions

### Review Exercises:

1. Contrast Slater type orbitals (STOs) with Gaussian type orbitals (GTOs).

### Exercises:

1. By expanding the molecular orbitals  $\{\psi_k\}$  as linear combinations of atomic orbitals  $\{\phi_\mu\}$ ,

$$\psi_k = \sum_{\mu} c_{\mu k} \phi_{\mu}$$

show how the canonical Hartree-Fock (HF) equations:

$$F \psi_i = \epsilon_i \psi_i$$

reduce to the matrix eigenvalue-type equation of the form given in the text:

$$F_{\mu} C_i = \epsilon_i S_{\mu} C_i$$

where:

$$F_{\mu} = \langle \phi_{\mu} | h | \phi_{\mu} \rangle + \sum_{\nu} \langle \phi_{\mu} | g | \phi_{\nu} \rangle - \sum_{\nu}^{\text{ex}} \langle \phi_{\mu} | g | \phi_{\nu} \rangle,$$

$$S_{\mu} = \langle \phi_{\mu} | \phi_{\mu} \rangle, \quad S_{\mu\nu} = \sum_{i=\text{occ}} C_i C_i,$$

$$\text{and } \sum_{i=\text{occ and same spin}}^{\text{ex}} C_i C_i.$$

Note that the sum over  $i$  in  $S_{\mu\nu}$  and  $\sum_{i=\text{occ and same spin}}^{\text{ex}}$  is a sum over spin orbitals. In addition, show that this Fock matrix can be further reduced for the closed shell case to:

$$F_{\mu} = \langle \phi_{\mu} | h | \phi_{\mu} \rangle + P \langle \phi_{\mu} | g | \phi_{\mu} \rangle - \frac{1}{2} \sum_{\nu} \langle \phi_{\mu} | g | \phi_{\nu} \rangle,$$

where the charge bond order matrix,  $P$ , is defined to be:

$$P = \sum_{i=\text{occ}} 2 C_i C_i,$$

where the sum over  $i$  here is a sum over orbitals not spin orbitals.

2. Show that the HF total energy for a closed-shell system may be written in terms of integrals over the orthonormal HF orbitals as:

$$E(\text{SCF}) = 2 \sum_k^{\text{occ}} \langle k|h|k \rangle + \sum_{kl}^{\text{occ}} [2\langle kl|gkl \rangle - \langle kl|g|lk \rangle] + \sum_{\mu} \frac{Z_{\mu}Z}{R_{\mu}}$$

3. Show that the HF total energy may alternatively be expressed as:

$$E(\text{SCF}) = \sum_k^{\text{occ}} \epsilon_k + \sum_{\mu} \frac{Z_{\mu}Z}{R_{\mu}},$$

where the  $\epsilon_k$  refer to the HF orbital energies.

### Problems:

1. This problem will be concerned with carrying out an SCF calculation for the  $\text{HeH}^+$  molecule in the  $1^1_g+(1^2)$  ground state. The one- and two-electron integrals (in atomic units) needed to carry out this SCF calculation at  $R = 1.4$  a.u. using Slater type orbitals with orbital exponents of 1.6875 and 1.0 for the He and H, respectively are:

$$\begin{aligned} S_{11} &= 1.0, & S_{22} &= 1.0, & S_{12} &= 0.5784, \\ h_{11} &= -2.6442, & h_{22} &= -1.7201, & h_{12} &= -1.5113, \\ g_{1111} &= 1.0547, & g_{1121} &= 0.4744, & g_{1212} &= 0.5664, \\ g_{2211} &= 0.2469, & g_{2221} &= 0.3504, & g_{2222} &= 0.6250, \end{aligned}$$

where 1 refers to  $1s_{\text{He}}$  and 2 to  $1s_{\text{H}}$ . Note that the two-electron integrals are given in Dirac notation. Parts a. - d. should be done by hand. Any subsequent parts can make use of the QMIC software provided.

a. Using  $\phi_1 = 1s_{\text{He}}$  for the initial guess of the occupied molecular orbital, form a  $2 \times 2$  Fock matrix. Use the equation derived above in question 1 for  $F_{\mu}$ .

b. Solve the Fock matrix eigenvalue equations given above to obtain the orbital energies and an improved occupied molecular orbital. In so doing, note that  $\langle \phi_1 | \phi_1 \rangle = 1 = C_1^T S C_1$  gives the needed normalization condition for the expansion coefficients of the  $\phi_1$  in the atomic orbital basis.

c. Determine the total SCF energy using the result of exercise 3 above at this step of the iterative procedure. When will this energy agree with that obtained by using the alternative expression for  $E(\text{SCF})$  given in exercise 2?

d. Obtain the new molecular orbital,  $\phi_1$ , from the solution of the matrix eigenvalue problem (part b).

e. A new Fock matrix and related total energy can be obtained with this improved choice of molecular orbital,  $\phi_1$ . This process can be continued until a convergence criterion has been satisfied. Typical convergence criteria include: no significant change in the molecular orbitals or the total energy (or both) from one iteration to the next. Perform this iterative procedure for the  $\text{HeH}^+$  system until the difference in total energy between two successive iterations is less than  $10^{-5}$  a.u.

f. Show, by comparing the difference between the SCF total energy at one iteration and the converged SCF total energy, that the convergence of the above SCF approach is primarily linear (or first order).

g. Is the SCF total energy calculated at each iteration of the above SCF procedure (via exercise 3) an upper bound to the exact ground-state total energy?

h. Using the converged self-consistent set of molecular orbitals,  $\phi_1$  and  $\phi_2$ , calculate the one- and two-electron integrals in the molecular orbital basis. Using the equations for  $E(\text{SCF})$  in exercises 2 and 3 calculate the converged values of the orbital energies making use of these integrals in the mo basis.

i. Does this SCF wavefunction give rise (at  $R = 1.4$ ) to proper dissociation products?

2. This problem will continue to address the same  $\text{HeH}^+$  molecular system as above, extending the analysis to include "correlation effects." We will use the one- and two-electron integrals (same geometry) in the converged (to  $10^{-5}$  au) SCF molecular orbital basis which we would have obtained after 7 iterations above. The converged mos you would have obtained in problem 1 are:

$$\begin{array}{l} \epsilon_1 = -0.89997792 \quad \epsilon_2 = -0.83233180 \\ \epsilon_1 = -0.15843012 \quad \epsilon_2 = 1.21558030 \end{array}$$

a. Carry out a two configuration CI calculation using the  $1^2$  and  $2^2$  configurations first by obtaining an expression for the CI matrix elements  $H_{ij}$  ( $i, j = 1^2, 2^2$ ) in terms of one- and two-electron integrals, and secondly by showing that the resultant CI matrix is (ignoring the nuclear repulsion term):

$$\begin{pmatrix} -4.2720 & 0.1261 \\ 0.1261 & -2.0149 \end{pmatrix}$$

b. Obtain the two CI energies and eigenvectors for the matrix found in part a.

c. Show that the lowest energy CI wavefunction is equivalent to the following two-determinant (single configuration) wavefunction:

$$\frac{1}{2} \left( \frac{1}{a^2} \phi_1 + \frac{1}{b^2} \phi_2 \right) + \frac{1}{2} \left( \frac{1}{a^2} \phi_1 - \frac{1}{b^2} \phi_2 \right)$$

involving the polarized orbitals:  $\frac{1}{a^2} \phi_1 \pm \frac{1}{b^2} \phi_2$ , where  $a = 0.9984$  and  $b = 0.0556$ .

d. Expand the CI list to 3 configurations by adding the  $1^2$  to the original  $1^2$  and  $2^2$  configurations of part a above. First, express the proper singlet spin-coupled  $1^2$  configuration as a combination of Slater determinants and then compute all elements of this  $3 \times 3$  matrix.

e. Obtain all eigenenergies and corresponding normalized eigenvectors for this CI problem.

f. Determine the excitation energies and transition moments for  $\text{HeH}^+$  using the full CI result of part e above. The nonvanishing matrix elements of the dipole operator  $\mathbf{r}(x, y, z)$  in the atomic basis are:

$$\langle 1s_H | z | 1s_{He} \rangle = 0.2854 \quad \text{and} \quad \langle 1s_H | z | 1s_H \rangle = 1.4.$$

First determine the matrix elements of  $\mathbf{r}$  in the SCF orbital basis then determine the excitation energies and transition moments from the ground state to the two excited singlet states of  $\text{HeH}^+$ .

g. Now turning to perturbation theory, carry out a RSPT calculation of the first-order wavefunction  $|1^2\rangle^{(1)}$  for the case in which the zeroth-order wavefunction is taken to be the  $1^2$  Slater determinant. Show that the first-order wavefunction is given by:

$$|1^2 \rangle^{(1)} = -0.0442 |2^2 \rangle.$$

- h. Why does the  $|1^2 \rangle$  configuration not enter into the first-order wavefunction?
- i. Normalize the resultant wavefunction that contains zeroth- plus first-order parts and compare it to the wavefunction obtained in the two-configuration CI study of part b.
- j. Show that the second-order RSPT correlation energy,  $E^{(2)}$ , of  $\text{HeH}^+$  is -0.0056 a.u. How does this compare with the correlation energy obtained from the two-configuration CI study of part b?
3. Using the QMIC programs, calculate the SCF energy of  $\text{HeH}^+$  using the same geometry as in problem 1 and the STO3G basis set provided in the QMIC basis set library. How does this energy compare to that found in problem 1? Run the calculation again with the 3-21G basis set provided. How does this energy compare to the STO3G and the energy found using STOs in problem 1?
4. Generate SCF potential energy surfaces for  $\text{HeH}^+$  and  $\text{H}_2$  using the QMIC software provided. Use the 3-21G basis set and generate points for geometries of  $R = 1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 2.5, \text{ and } 10.0$ . Plot the energies vs. geometry for each system. Which system dissociates properly?
5. Generate CI potential energy surfaces for the 4 states of  $\text{H}_2$  resulting from a CAS calculation with 2 electrons in the lowest 2 SCF orbitals ( $1_g$  and  $1_u$ ). Use the same geometries and basis set as in problem 4. Plot the energies vs. geometry for each system. Properly label and characterize each of the states (e.g., repulsive, dissociate properly, etc.).

## Solutions

### Review Exercises:

1. Slater type orbitals (STOs) are "hydrogen-like" in that they have a normalized form of:

$$\frac{2}{a_0} n^{+\frac{1}{2}} \frac{1}{(2n)!} \frac{1}{2} r^{n-1} e^{-\frac{r}{a_0}} Y_{l,m}(\theta, \phi),$$

whereas gaussian type orbitals GTOs have the form:

$$N r^l e^{(-r^2)} Y_{l,m}(\theta, \phi),$$

although in most quantum chemistry computer programs they are specified in so-called "cartesian" form as:

$$N' x^a y^b z^c e^{(-r^2)},$$

where  $a, b,$  and  $c$  are quantum numbers each ranging from zero upward in unit steps. So, STOs give "better" overall energies and properties that depend on the shape of the wavefunction near the nuclei (e.g., Fermi contact ESR hyperfine constants) but they are more difficult to use (two-electron integrals are more difficult to evaluate; especially the 4-center variety which have to be integrated numerically). GTOs on the other hand are easier to use (more easily integrable) but improperly describe the wavefunction near the nuclear centers because of the so-called cusp condition (they have zero slope at  $R = 0$ , whereas 1s STOs have non-zero slopes there).

### Exercises:

$$1. \quad F_{ij} = h_{ij} + \sum_k [J_k - K_{kj}]_{ij}$$

Let the closed shell Fock potential be written as:

$$V_{ij} = \sum_k (2\langle ik|jk\rangle - \langle ik|kj\rangle), \text{ and the } 1e^- \text{ component as:}$$

$$h_{ij} = \left\langle i \left| -\frac{1}{2} \nabla^2 - \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} \right| j \right\rangle, \text{ and the delta as:}$$

$$\delta_{ij} = \langle ij \rangle, \text{ so that: } h_{ij} + V_{ij} = \delta_{ij}.$$

using:  $i = \sum_{\mu} C_{\mu i} \mu$ ,  $j = \sum_{\nu} C_{\nu j}$ , and  $k = \sum_{\kappa} C_{\kappa k}$ , and transforming from the mo to ao

basis we obtain:

$$\begin{aligned} V_{ij} &= \sum_{\mu} C_{\mu i} C_{\mu j} \sum_{\kappa} C_{\kappa k} C_{\kappa k} (2\langle \mu | \kappa \rangle - \langle \mu | \kappa \rangle) \\ &= \sum_{\mu} (C_{\mu i} C_{\mu j}) \sum_{\kappa} C_{\kappa k} C_{\kappa k} (2\langle \mu | \kappa \rangle - \langle \mu | \kappa \rangle) \\ &= \sum_{\mu} (C_{\mu i} C_{\mu j}) V_{\mu} \text{ where,} \end{aligned}$$

$$V_{\mu} = \sum_{\kappa} P_{\mu\kappa} (2\langle \mu | \kappa \rangle - \langle \mu | \kappa \rangle), \text{ and } P_{\mu\kappa} = \sum_k (C_{\kappa k} C_{\kappa k}),$$

$$h_{ij} = \sum_{\mu} (C_{\mu i} C_{\mu j}) h_{\mu}, \text{ where}$$

$$h_{\mu} = \left\langle \mu \left| -\frac{1}{2} \nabla^2 - \frac{Z_A}{|\mathbf{r} - \mathbf{R}_A|} \right| \mu \right\rangle, \text{ and}$$

$$\delta_{ij} = \langle ij \rangle = \sum_{\mu} (C_{\mu i} S_{\mu} C_{\mu j}).$$

So,  $h_{ij} + V_{ij} = \delta_{ij}$  becomes:

$$\sum_{\mu} (C_{\mu i} C_{\mu j}) h_{\mu} + \sum_{\mu} (C_{\mu i} C_{\mu j}) V_{\mu} = \sum_{\mu} (C_{\mu i} S_{\mu} C_{\mu j}),$$

$$\sum_{\mu} (C_{\mu i} S_{\mu} C_{\mu j}) - \sum_{\mu} (C_{\mu i} C_{\mu j}) h_{\mu} - \sum_{\mu} (C_{\mu i} C_{\mu j}) V_{\mu} = 0 \text{ for all } i, j$$

$$\sum_{\mu} C_{\mu i} S_{\mu} - h_{\mu} - V_{\mu} C_{\mu j} = 0 \text{ for all } i, j$$

Therefore,

$$h_{\mu} + V_{\mu} - \sum_j S_{\mu j} C_j = 0$$

This is FC = SCE.

2. The Slater Condon rule for zero (spin orbital) difference with N electrons in N spin orbitals is:

$$\begin{aligned} E = \langle H + G \rangle &= \sum_i^N \langle i|h|i\rangle + \sum_{i>j}^N \langle i|j|g|i\rangle - \sum_{i>j}^N \langle i|j|g|j\rangle \\ &= \sum_i h_{ii} + \sum_{i>j} (g_{ijij} - g_{ijji}) \\ &= \sum_i h_{ii} + \frac{1}{2} \sum_{ij} (g_{ijij} - g_{ijji}) \end{aligned}$$

If all orbitals are doubly occupied and we carry out the spin integration we obtain:

$$E = 2 \sum_i^{\text{occ}} h_{ii} + \sum_{ij}^{\text{occ}} (2g_{ijij} - g_{ijji}),$$

where i and j now refer to orbitals (not spin-orbitals).

3. If the occupied orbitals obey  $F_k = \epsilon_k C_k$ , then the expression for E in problem 2 above can be rewritten as.

$$E = \sum_i^{\text{occ}} h_{ii} + \sum_j^{\text{occ}} (2g_{ijij} - g_{ijji}) + \sum_i^{\text{occ}} h_{ii}$$

We recognize the closed shell Fock operator expression and rewrite this as:

$$E = \sum_i^{\text{occ}} F_{ii} + \sum_i^{\text{occ}} h_{ii} = \sum_i^{\text{occ}} (\epsilon_i + h_{ii})$$

### Problems:

1. We will use the QMOC software to do this problem. Lets just start from the beginning. Get the starting "guess" mo coefficients on disk. Using the program MOCOEFs it asks us for the first and second mo vectors. We input 1, 0 for the first mo (this means that the first mo is 1.0 times the He 1s orbital plus 0.0 times the H 1s orbital; this bonding mo is more likely to be heavily weighted on the atom having the higher nuclear charge) and 0, 1 for the

second. Our beginning mo-ao array looks like:  $\begin{matrix} 1.0 & 0.0 \\ 0.0 & 1.0 \end{matrix}$  and is placed on disk in a file we

choose to call "mocoefs.dat". We also put the ao integrals on disk using the program RW\_INTS. It asks for the unique one- and two- electron integrals and places a canonical list of these on disk in a file we choose to call "ao\_integrals.dat". At this point it is useful for us to step back and look at the set of equations which we wish to solve: FC = SCE. The QMOC software does not provide us with a so-called generalized eigenvalue solver (one that contains an overlap matrix; or metric), so in order to use the diagonalization program that is provided we must transform this equation (FC = SCE) to one that looks like (F'C' = C'E). We do that in the following manner:

Since  $S$  is symmetric and positive definite we can find an  $S^{-\frac{1}{2}}$  such that  $S^{-\frac{1}{2}} S^{+\frac{1}{2}} = 1$ ,  $S^{-\frac{1}{2}} S^{+\frac{1}{2}} = S^{+\frac{1}{2}}$ , etc.

rewrite  $FC = SCE$  by inserting unity between  $FC$  and multiplying the whole equation on the left by  $S^{-\frac{1}{2}}$ . This gives:

$$S^{-\frac{1}{2}} FS^{-\frac{1}{2}} S^{+\frac{1}{2}} C = S^{-\frac{1}{2}} SCE = S^{+\frac{1}{2}} CE.$$

Letting:  $F' = S^{-\frac{1}{2}} FS^{-\frac{1}{2}}$

$$C' = S^{+\frac{1}{2}} C, \text{ and inserting these expressions above give:}$$

$$F'C' = C'E$$

Note, that to get the next iterations mo coefficients we must calculate  $C$  from  $C'$ :

$C' = S^{+\frac{1}{2}} C$ , so, multiplying through on the left by  $S^{-\frac{1}{2}}$  gives:

$$S^{-\frac{1}{2}} C' = S^{-\frac{1}{2}} S^{+\frac{1}{2}} C = C$$

This will be the method we will use to solve our fock equations.

Find  $S^{-\frac{1}{2}}$  by using the program FUNCT\_MAT (this program generates a function of a matrix). This program will ask for the elements of the  $S$  array and write to disk a file

(name of your choice ... a good name might be "shalf") containing the  $S^{-\frac{1}{2}}$  array. Now we are ready to begin the iterative Fock procedure.

a. Calculate the Fock matrix,  $F$ , using program FOCK which reads in the mo coefficients from "mocoefs.dat" and the integrals from "ao\_integrals.dat" and writes the resulting Fock matrix to a user specified file (a good filename to use might be something like "fock1").

b. Calculate  $F' = S^{-\frac{1}{2}} FS^{-\frac{1}{2}}$  using the program UTMATU which reads in  $F$  and  $S^{-\frac{1}{2}}$  from files on the disk and writes  $F'$  to a user specified file (a good filename to use might be something like "fock1p"). Diagonalize  $F'$  using the program DIAG. This program reads in the matrix to be diagonalized from a user specified filename and writes the resulting eigenvectors to disk using a user specified filename (a good filename to use might be something like "coef1p"). You may wish to choose the option to write the eigenvalues (Fock orbital energies) to disk in order to use them at a later time in program FENERGY.

Calculate  $C$  by back transforming e.g.  $C = S^{-\frac{1}{2}} C'$ . This is accomplished by using the program MATXMAT which reads in two matrices to be multiplied from user specified files and writes the product to disk using a user specified filename (a good filename to use might be something like "mocoefs.dat").

c. The QMIC program FENERGY calculates the total energy, using the result of exercises 2 and 3;

$$2\langle k|h|k\rangle + 2\langle kl|kl\rangle - \langle kl|lk\rangle + \frac{Z_{\mu}Z}{R_{\mu}}, \quad \text{and}$$

$\mu >$

$$E_k = \langle k|h|k\rangle + \frac{Z_\mu Z_\mu}{R_\mu}$$

This is the conclusion of one iteration of the Fock procedure ... you may continue by going back to part a. and proceeding onward.

d. and e. Results for the successful convergence of this system using the supplied QMIC software is as follows (this is a lot of bloody detail but will give the user assurance that they are on the right track; alternatively one could switch to the QMIC program SCF and allow that program to iteratively converge the Fock equations):

The one-electron AO integrals:

-2.644200	-1.511300
-1.511300	-1.720100

The two-electron AO integrals:

1	1	1	1	1.054700
2	1	1	1	0.4744000
2	1	2	1	0.5664000
2	2	1	1	0.2469000
2	2	2	1	0.3504000
2	2	2	2	0.6250000

The "initial" MO-AO coefficients:

1.000000	0.000000
0.000000	1.000000

AO overlap matrix (S):

1.000000	0.578400
0.578400	1.000000

$S^{-\frac{1}{2}}$

1.168032	-0.3720709
-0.3720709	1.168031

\*\*\*\*\*  
 ITERATION 1  
 \*\*\*\*\*

The charge bond order matrix:

1.000000	0.0000000
0.0000000	0.0000000

The Fock matrix (F):

-1.589500	-1.036900
-1.036900	-0.8342001

$S^{-\frac{1}{2}} F S^{-\frac{1}{2}}$

-1.382781	-0.5048679
-0.5048678	-0.4568883

The eigenvalues of this matrix (Fock orbital energies) are:

[ -1.604825 -0.2348450 ]

Their corresponding eigenvectors ( $C' = S^{+\frac{1}{2}} * C$ ) are:

-0.9153809 -0.4025888  
-0.4025888 0.9153810

The "new" MO-AO coefficients ( $C = S^{-\frac{1}{2}} * C'$ ):

-0.9194022 -0.8108231  
-0.1296498 1.218985

The one-electron MO integrals:

-2.624352 -0.1644336  
-0.1644336 -1.306845

The two-electron MO integrals:

1 1 1 1 0.9779331  
2 1 1 1 0.1924623  
2 1 2 1 0.5972075  
2 2 1 1 0.1170838  
2 2 2 1 -0.0007945194  
2 2 2 2 0.6157323

The closed shell Fock energy from formula:

$$2 \sum_{kl} \langle k|h|k \rangle + 2 \sum_{kl} \langle kl|kl \rangle - \sum_{\mu} \langle kl|lk \rangle + \frac{Z_{\mu}Z}{R_{\mu}} = -2.84219933$$

from formula:

$$\sum_k \langle k|h|k \rangle + \sum_{\mu} \frac{Z_{\mu}Z}{R_{\mu}} = -2.80060530$$

the difference is: -0.04159403

\*\*\*\*\*  
ITERATION 2  
\*\*\*\*\*

The charge bond order matrix: 0.8453005 0.1192003  
0.1192003 0.01680906

The Fock matrix:

	-1.624673	-1.083623
	-1.083623	-0.8772071

$S^{-\frac{1}{2}} F S^{-\frac{1}{2}}$

	-1.396111	-0.5411037
	-0.5411037	-0.4798213

The eigenvalues of this matrix (Fock orbital energies) are:

[ -1.646972   -0.2289599 ]

Their corresponding eigenvectors ( $C' = S^{+\frac{1}{2}} * C$ ) are:

-0.9072427	-0.4206074
-0.4206074	0.9072427

The "new" MO-AO coefficients ( $C = S^{-\frac{1}{2}} * C'$ ):

-0.9031923	-0.8288413
-0.1537240	1.216184

The one-electron MO integrals:

-2.617336	-0.1903475
-0.1903475	-1.313861

The two-electron MO integrals:

1 1 1 1	0.9626070
2 1 1 1	0.1949828
2 1 2 1	0.6048143
2 2 1 1	0.1246907
2 2 2 1	0.003694540
2 2 2 2	0.6158437

The closed shell Fock energy from formula:

$$2\sum_{kl} \langle k|h|k\rangle + 2\sum_{kl} \langle kl|kl\rangle - \sum_{\mu} \langle kl|lk\rangle + \sum_{\mu} \frac{Z_{\mu}Z}{R_{\mu}} = -2.84349298$$

from formula:

$$\sum_k \langle k|h|k\rangle + \sum_{\mu} \frac{Z_{\mu}Z}{R_{\mu}} = -2.83573675$$



$$2\langle k|h|k\rangle + 2\langle kl|kl\rangle - \langle kl|lk\rangle + \frac{Z_{\mu}Z}{R_{\mu}} = -2.84353018$$

from formula:

$$\langle k|h|k\rangle + \frac{Z_{\mu}Z}{R_{\mu}} = -2.84225941$$

the difference is: -0.00127077

\*\*\*\*\*

ITERATION 4

\*\*\*\*\*

The charge bond order matrix:

0.8108885	0.1419869
0.1419869	0.02486194

The Fock matrix:

-1.632213	-1.093155
-1.093155	-0.8866909

$S^{-\frac{1}{2}} F S^{-\frac{1}{2}}$

-1.399426	-0.5480287
-0.5480287	-0.4855191

The eigenvalues of this matrix (Fock orbital energies) are:

[ -1.656015   -0.2289308 ]

Their corresponding eigenvectors ( $C' = S^{+\frac{1}{2}} * C$ ) are:

-0.9056494	-0.4240271
-0.4240271	0.9056495

The "new" MO-AO coefficients ( $C = S^{-\frac{1}{2}} * C'$ ):

-0.9000589	-0.8322428
-0.1583111	1.215595

The one-electron MO integrals:

-2.615881	-0.1952594
-0.1952594	-1.315315

The two-electron MO integrals:

1	1	1	1	0.9596615
2	1	1	1	0.1953781
2	1	2	1	0.6062557
2	2	1	1	0.1261321
2	2	2	1	0.004601604
2	2	2	2	0.6159065

The closed shell Fock energy from formula:

$$2\sum_{kl} \langle k|h|k\rangle + 2\sum_{kl} \langle kl|kl\rangle - \sum_{\mu} \frac{Z_{\mu}Z}{R_{\mu}} = -2.84352922$$

from formula:

$$\sum_k \langle k|h|k\rangle + \sum_{\mu} \frac{Z_{\mu}Z}{R_{\mu}} = -2.84332418$$

the difference is: -0.00020504

\*\*\*\*\*  
 ITERATION 5  
 \*\*\*\*\*

The charge bond order matrix:

0.8101060	0.1424893
0.1424893	0.02506241

The Fock matrix:

-1.632385	-1.093368
-1.093368	-0.8869066

$S^{-\frac{1}{2}} F S^{-\frac{1}{2}}$

-1.399504	-0.5481812
-0.5481813	-0.4856516

The eigenvalues of this matrix (Fock orbital energies) are:

[ -1.656219 -0.2289360 ]

Their corresponding eigenvectors ( $C' = S^{-\frac{1}{2}} * C$ ) are:

-0.9056138	-0.4241026
-0.4241028	0.9056141

The "new" MO-AO coefficients ( $C = S^{-\frac{1}{2}} * C'$ ):

-0.8999892 -0.8323179  
 -0.1584127 1.215582

The one-electron MO integrals:

-2.615847 -0.1953674  
 -0.1953674 -1.315348

The two-electron MO integrals:

1 1 1 1 0.9595956  
 2 1 1 1 0.1953862  
 2 1 2 1 0.6062872  
 2 2 1 1 0.1261639  
 2 2 2 1 0.004621811  
 2 2 2 2 0.6159078

The closed shell Fock energy from formula:

$$2 \sum_{kl} \langle k|h|k \rangle + 2 \sum_{kl} \langle kl|kl \rangle - \sum_{\mu} \frac{Z_{\mu}Z}{R_{\mu}} = -2.84352779$$

from formula:

$$\sum_k \langle k|h|k \rangle + \sum_{\mu} \frac{Z_{\mu}Z}{R_{\mu}} = -2.84349489$$

the difference is: -0.00003290

\*\*\*\*\*  
 ITERATION 6  
 \*\*\*\*\*

The charge bond order matrix: 0.8099805 0.1425698  
 0.1425698 0.02509460

The Fock matrix: -1.632412 -1.093402  
 -1.093402 -0.8869413

$S^{-\frac{1}{2}} F S^{-\frac{1}{2}}$  -1.399517 -0.5482056  
 -0.5482056 -0.4856730

The eigenvalues of this matrix (Fock orbital energies) are:

[ -1.656253 -0.2289375 ]

Their corresponding eigenvectors ( $C' = S^{-\frac{1}{2}} * C$ ) are:

-0.9056085 -0.4241144  
 -0.4241144 0.9056086

The "new" MO-AO coefficients ( $C = S^{\frac{1}{2}} * C'$ ):

-0.8999786 -0.8323296  
 -0.1584283 1.215580

The one-electron MO integrals:

-2.615843 -0.1953846  
 -0.1953846 -1.315353

The two-electron MO integrals:

1 1 1 1 0.9595859  
 2 1 1 1 0.1953878  
 2 1 2 1 0.6062925  
 2 2 1 1 0.1261690  
 2 2 2 1 0.004625196  
 2 2 2 2 0.6159083

The closed shell Fock energy from formula:

$$2 \sum_{kl} \langle k|h|k\rangle + 2 \sum_{kl} \langle kl|kl\rangle - \sum_{\mu} \langle k||l\rangle + \frac{Z_{\mu}Z}{R_{\mu}} = -2.84352827$$

from formula:

$$\sum_k \langle k|h|k\rangle + \sum_{\mu} \frac{Z_{\mu}Z}{R_{\mu}} = -2.84352398$$

the difference is: -0.00000429

\*\*\*\*\*  
 ITERATION 7  
 \*\*\*\*\*

The charge bond order matrix: 0.8099616 0.1425821  
 0.1425821 0.02509952

The Fock matrix:

$$\begin{matrix} & & -1.632416 & -1.093407 \\ & & -1.093407 & -0.8869464 \\ S^{\frac{1}{2}} F S^{\frac{1}{2}} & & -1.399519 & -0.5482093 \\ & & -0.5482092 & -0.4856761 \end{matrix}$$

The eigenvalues of this matrix (Fock orbital energies) are:

$$[ -1.656257 \quad -0.2289374 ]$$

Their corresponding eigenvectors ( $C' = S^{\frac{1}{2}} * C$ ) are:

$$\begin{matrix} -0.9056076 & -0.4241164 \\ -0.4241164 & 0.9056077 \end{matrix}$$

The "new" MO-AO coefficients ( $C = S^{-\frac{1}{2}} * C'$ ):

$$\begin{matrix} -0.8999770 & -0.8323317 \\ -0.1584310 & 1.215580 \end{matrix}$$

The one-electron MO integrals:

$$\begin{matrix} -2.615843 & -0.1953876 \\ -0.1953876 & -1.315354 \end{matrix}$$

The two-electron MO integrals:

$$\begin{matrix} 1 & 1 & 1 & 1 & 0.9595849 \\ 2 & 1 & 1 & 1 & 0.1953881 \\ 2 & 1 & 2 & 1 & 0.6062936 \\ 2 & 2 & 1 & 1 & 0.1261697 \\ 2 & 2 & 2 & 1 & 0.004625696 \\ 2 & 2 & 2 & 2 & 0.6159083 \end{matrix}$$

The closed shell Fock energy from formula:

$$2 \sum_{kl} \langle k|h|k \rangle + 2 \sum_{kl} \langle kl|kl \rangle - \sum_{\mu} \frac{Z_{\mu} Z}{R_{\mu}} = -2.84352922$$

from formula:

$$\sum_k \langle k|h|k \rangle + \sum_{\mu} \frac{Z_{\mu} Z}{R_{\mu}} = -2.84352827$$

the difference is:

-0.00000095

\*\*\*\*\*

ITERATION 8

\*\*\*\*\*

The charge bond order matrix:                    0.8099585    0.1425842  
    0.1425842    0.02510037

The Fock matrix:                                    -1.632416   -1.093408  
   -1.093408   -0.8869470

$S^{-\frac{1}{2}} F S^{-\frac{1}{2}}$                                     -1.399518   -0.5482103  
   -0.5482102   -0.4856761

The eigenvalues of this matrix (Fock orbital energies) are:

[ -1.656258   -0.2289368 ]

Their corresponding eigenvectors ( $C' = S^{+\frac{1}{2}} * C$ ) are:

-0.9056074   -0.4241168  
-0.4241168   0.9056075

The "new" MO-AO coefficients ( $C = S^{-\frac{1}{2}} * C'$ ):

-0.8999765   -0.8323320  
-0.1584315   1.215579

The one-electron MO integrals:

-2.615842   -0.1953882  
-0.1953882   -1.315354

The two-electron MO integrals:

1 1 1 1    0.9595841  
2 1 1 1    0.1953881  
2 1 2 1    0.6062934  
2 2 1 1    0.1261700  
2 2 2 1    0.004625901  
2 2 2 2    0.6159081

The closed shell Fock energy from formula:

$$2\langle k|h|k\rangle + 2\langle kl|kl\rangle - \langle kl|lk\rangle + \frac{Z_\mu Z}{R_\mu} = -2.84352827$$

from formula:

$$E_k + \langle k|h|k\rangle + \frac{Z_\mu Z}{R_\mu} = -2.84352827$$

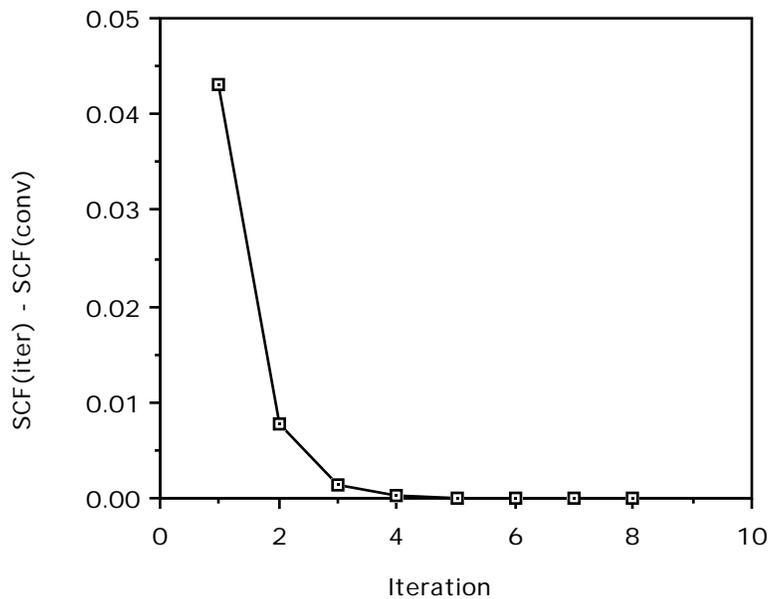
the difference is: 0.00000000

f. In looking at the energy convergence we see the following:

Iter	Formula 1	Formula 2
1	-2.84219933	-2.80060530
2	-2.84349298	-2.83573675
3	-2.84353018	-2.84225941
4	-2.84352922	-2.84332418
5	-2.84352779	-2.84349489
6	-2.84352827	-2.84352398
7	-2.84352922	-2.84352827
8	-2.84352827	-2.84352827

f. If you look at the energy differences (SCF at iteration n - SCF converged) and plot this data versus iteration number, and do a 5th order polynomial fit, we see the following:

$$y = 0.144 - 0.153x + 0.063x^2 - 0.013x^3 + 0.001x^4 \quad R = 1.00$$



In looking at the polynomial fit we see that the convergence is primarily linear since the coefficient of the linear term is much larger than those of the cubic and higher terms.

g. The converged SCF total energy calculated using the result of exercise 3 is an upper bound to the ground state energy, but, during the iterative procedure it is not. At

convergence, the expectation value of the Hamiltonian for the Hartree Fock determinant is given by the equation in exercise 3.

h. The one- and two- electron integrals in the MO basis are given above (see part e iteration 8). The orbital energies are found using the result of exercise 2 and 3 to be:

$$E(\text{SCF}) = \sum_k h_k + \sum_{\mu} \frac{Z_{\mu}Z}{R_{\mu}}$$

$$E(\text{SCF}) = \sum_{kl} 2\langle k|h|k\rangle + 2\langle kl|kl\rangle - \langle kl|lk\rangle + \sum_{\mu} \frac{Z_{\mu}Z}{R_{\mu}}$$

so, 
$$h_k = \langle k|h|k\rangle + \frac{1}{\text{occ}} (2\langle kl|kl\rangle - \langle kl|lk\rangle)$$

$$\begin{aligned} 1 &= h_{11} + 2\langle 11|11\rangle - \langle 11|11\rangle \\ &= -2.615842 + 0.9595841 \\ &= -1.656258 \end{aligned}$$

$$\begin{aligned} 2 &= h_{22} + 2\langle 21|21\rangle - \langle 21|12\rangle \\ &= -1.315354 + 2*0.6062934 - 0.1261700 \\ &= -0.2289372 \end{aligned}$$

i. Yes, the  $1^2$  configuration does dissociate properly because at  $R \rightarrow \infty$  the lowest energy state is  $\text{He} + \text{H}^+$ , which also has a  $1^2$  orbital occupancy (i.e.,  $1s^2$  on He and  $1s^0$  on  $\text{H}^+$ ).

2. At convergence the mo coefficients are:

$$\begin{aligned} 1 &= \begin{matrix} -0.8999765 & -0.8323320 \\ -0.1584315 & 1.215579 \end{matrix} \\ 2 &= \end{aligned}$$

and the integrals in this MO basis are:

$$\begin{aligned} h_{11} &= -2.615842 & h_{21} &= -0.1953882 & h_{22} &= -1.315354 \\ g_{1111} &= 0.9595841 & g_{2111} &= 0.1953881 & g_{2121} &= 0.6062934 \\ g_{2211} &= 0.1261700 & g_{2221} &= 0.04625901 & g_{2222} &= 0.6159081 \end{aligned}$$

$$\begin{aligned} \text{a. } H &= \begin{matrix} \langle 1^2|H|1^2\rangle & \langle 1^2|H|2^2\rangle \\ \langle 2^2|H|1^2\rangle & \langle 2^2|H|2^2\rangle \end{matrix} = \begin{matrix} 2h_{11} + g_{1111} & g_{1122} \\ g_{1122} & 2h_{22} + g_{2222} \end{matrix} \\ &= \begin{matrix} 2*-2.615842 + 0.9595841 & 0.1261700 \\ 0.1261700 & 2*-1.315354 + 0.6159081 \end{matrix} \\ &= \begin{matrix} -4.272100 & 0.126170 \\ 0.126170 & -2.014800 \end{matrix} \end{aligned}$$

b. The eigenvalues are  $E_1 = -4.279131$  and  $E_2 = -2.007770$ . The corresponding eigenvectors are:

$$C_1 = \begin{matrix} -0.99845123 & 0.05563438 \\ 0.05563439 & 0.99845140 \end{matrix}, C_2 =$$

c.

$$\frac{1}{2} \left( \frac{1}{a^2} \psi_1 + \frac{1}{b^2} \psi_2 \right) + \frac{1}{2} \left( \frac{1}{a^2} \psi_1 - \frac{1}{b^2} \psi_2 \right) + \frac{1}{2} \left( \frac{1}{a^2} \psi_1 - \frac{1}{b^2} \psi_2 \right) + \frac{1}{2} \left( \frac{1}{a^2} \psi_1 + \frac{1}{b^2} \psi_2 \right)$$

$$= \frac{1}{2\sqrt{2}} \left( a^2 \frac{1}{1} + b^2 \frac{1}{2} - a^2 \frac{1}{1} - b^2 \frac{1}{2} + a^2 \frac{1}{1} - b^2 \frac{1}{2} + a^2 \frac{1}{1} + b^2 \frac{1}{2} \right) \left( \begin{array}{c} - \\ - \end{array} \right)$$

$$= \frac{1}{\sqrt{2}} (a \frac{1}{1} - b \frac{1}{2}) \left( \begin{array}{c} - \\ - \end{array} \right)$$

$$= a | \frac{1}{1} \frac{1}{1} | - b | \frac{1}{2} \frac{1}{2} | .$$

(note from part b.  $a = 0.9984$  and  $b = 0.0556$ )

d. The third configuration  $| \frac{1}{2} \frac{1}{2} | = \frac{1}{\sqrt{2}} [ | \frac{1}{2} \frac{1}{2} | - | \frac{1}{2} \frac{1}{2} | ] ,$

Adding this configuration to the previous 2x2 CI results in the following 3x3 'full' CI:

$$H = \begin{matrix} \langle 1 \frac{1}{2} | H | 1 \frac{1}{2} \rangle & \langle 1 \frac{1}{2} | H | 2 \frac{1}{2} \rangle & \langle 1 \frac{1}{2} | H | 1 \frac{1}{2} \rangle \\ \langle 2 \frac{1}{2} | H | 1 \frac{1}{2} \rangle & \langle 2 \frac{1}{2} | H | 2 \frac{1}{2} \rangle & \langle 2 \frac{1}{2} | H | 1 \frac{1}{2} \rangle \\ \langle 1 \frac{1}{2} | H | 1 \frac{1}{2} \rangle & \langle 2 \frac{1}{2} | H | 1 \frac{1}{2} \rangle & \langle 1 \frac{1}{2} | H | 1 \frac{1}{2} \rangle \end{matrix}$$

$$= \begin{matrix} 2h_{11} + g_{1111} & g_{1122} & \frac{1}{\sqrt{2}} [ 2h_{12} + 2g_{2111} ] \\ g_{1122} & 2h_{22} + g_{2222} & \frac{1}{\sqrt{2}} [ 2h_{12} + 2g_{2221} ] \\ \frac{1}{\sqrt{2}} [ 2h_{12} + 2g_{2111} ] & \frac{1}{\sqrt{2}} [ 2h_{12} + 2g_{2221} ] & h_{11} + h_{22} + g_{2121} + g_{2211} \end{matrix}$$

Evaluating the new matrix elements:

$$H_{13} = H_{31} = \sqrt{2} * (-0.1953882 + 0.1953881) = 0.0$$

$$H_{23} = H_{32} = \sqrt{2} * (-0.1953882 + 0.004626) = -0.269778$$

$$H_{33} = -2.615842 - 1.315354 + 0.606293 + 0.126170$$

$$= -3.198733$$

$$= \begin{matrix} -4.272100 & 0.126170 & 0.0 \\ 0.126170 & -2.014800 & -0.269778 \\ 0.0 & -0.269778 & -3.198733 \end{matrix}$$

e. The eigenvalues are  $E_1 = -4.279345$ ,  $E_2 = -3.256612$  and  $E_3 = -1.949678$ . The corresponding eigenvectors are:

$$C_1 = \begin{matrix} -0.99825280 & -0.02605343 & -0.05302767 \\ 0.05732290 & -0.20969283 & -0.97608540 \\ 0.01431085 & -0.97742000 & 0.21082004 \end{matrix} , C_2 = \begin{matrix} -0.02605343 & -0.97608540 & 0.21082004 \\ -0.97608540 & -0.02605343 & -0.99825280 \\ 0.21082004 & 0.21082004 & -0.97608540 \end{matrix} , C_3 = \begin{matrix} -0.05302767 & -0.97608540 & 0.21082004 \\ -0.97608540 & -0.02605343 & -0.99825280 \\ 0.21082004 & -0.97608540 & -0.05302767 \end{matrix}$$

f. We need the non-vanishing matrix elements of the dipole operator in the mo basis. These can be obtained by calculating them by hand. They are more easily obtained by using the TRANS program. Put the 1e- ao integrals on disk by running the program RW\_INTS. In this case you are inserting  $z_{11} = 0.0$ ,  $z_{21} = 0.2854$ , and  $z_{22} = 1.4$  (insert 0.0 for all the 2e- integrals) ... call the output file "ao\_dipole.ints" for example. The converged MO-AO coefficients should be in a file ("mocoefs.dat" is fine). The

transformed integrals can be written to a file (name of your choice) for example "mo\_dipole.ints". These matrix elements are:

$$z_{11} = 0.11652690, z_{21} = -0.54420990, z_{22} = 1.49117320$$

The excitation energies are  $E_2 - E_1 = -3.256612 - -4.279345 = 1.022733$ , and  $E_3 - E_1 = -1.949678 - -4.279345 = 2.329667$ .

Using the Slater-Condon rules to obtain the matrix elements between configurations we get:

$$\begin{aligned} & \langle 1^2 | z | 1^2 \rangle \quad \langle 1^2 | z | 2^2 \rangle \quad \langle 1^2 | z | 1^2 2 \rangle \\ H_z = & \langle 2^2 | z | 1^2 \rangle \quad \langle 2^2 | z | 2^2 \rangle \quad \langle 2^2 | z | 1^2 2 \rangle \\ & \langle 1^2 2 | z | 1^2 2 \rangle \quad \langle 2^2 | z | 1^2 2 \rangle \quad \langle 1^2 2 | z | 1^2 2 \rangle \\ & 2z_{11} \quad 0 \quad \frac{1}{\sqrt{2}}[2z_{12}] \\ = & 0 \quad 2z_{22} \quad \frac{1}{\sqrt{2}}[2z_{12}] \\ & \frac{1}{\sqrt{2}}[2z_{12}] \quad \frac{1}{\sqrt{2}}[2z_{12}] \quad z_{11} + z_{22} \\ & 0.233054 \quad 0 \quad -0.769629 \\ = & 0 \quad 2.982346 \quad -0.769629 \\ & -0.769629 \quad -0.769629 \quad 1.607700 \end{aligned}$$

Now,  $\langle 1^2 | z | 2^2 \rangle = C_1^T H_z C_2$ , (this can be accomplished with the program UTMATU)

$$\begin{aligned} & -0.99825280 \quad T \quad 0.233054 \quad 0 \quad -0.769629 \quad -0.02605343 \\ = & 0.05732290 \quad 0 \quad 2.982346 \quad -0.769629 \quad -0.20969283 \\ & 0.01431085 \quad -0.769629 \quad -0.769629 \quad 1.607700 \quad -0.97742000 \\ = & -.757494 \end{aligned}$$

and,  $\langle 1^2 | z | 3^2 \rangle = C_1^T H_z C_3$

$$\begin{aligned} & -0.99825280 \quad T \quad 0.233054 \quad 0 \quad -0.769629 \quad -0.05302767 \\ = & 0.05732290 \quad 0 \quad 2.982346 \quad -0.769629 \quad -0.97608540 \\ & 0.01431085 \quad -0.769629 \quad -0.769629 \quad 1.607700 \quad 0.21082004 \\ = & 0.014322 \end{aligned}$$

g. Using the converged coefficients the orbital energies obtained from solving the Fock equations are  $\epsilon_1 = -1.656258$  and  $\epsilon_2 = -0.228938$ . The resulting expression for the RSPT first-order wavefunction becomes:

$$\begin{aligned} |1^2 \rangle^{(1)} &= -\frac{g_{2211}}{2(\epsilon_2 - \epsilon_1)} |2^2 \rangle \\ |1^2 \rangle^{(1)} &= -\frac{0.126170}{2(-0.228938 + 1.656258)} |2^2 \rangle \\ |1^2 \rangle^{(1)} &= -0.0441982 |2^2 \rangle \end{aligned}$$

h. As you can see from part c., the matrix element  $\langle 1^2 | H | 1^2 \rangle = 0$  (this is also a result of the Brillouin theorem) and hence this configuration does not enter into the first-order wavefunction.

i.  $|0\rangle = |1^2\rangle - 0.0441982|2^2\rangle$ . To normalize we divide by:

$$\sqrt{[1 + (0.0441982)^2]} = 1.0009762$$

$$|0\rangle = 0.999025|1^2\rangle - 0.044155|2^2\rangle$$

In the 2x2 CI we obtained:

$$|0\rangle = 0.99845123|1^2\rangle - 0.05563439|2^2\rangle$$

j. The expression for the 2<sup>nd</sup> order RSPT is:

$$E^{(2)} = -\frac{|g_{2211}|^2}{2(\epsilon_2 - \epsilon_1)} = -\frac{0.126170^2}{2(-0.228938 + 1.656258)}$$

$$= -0.005576 \text{ au}$$

Comparing the 2x2 CI energy obtained to the SCF result we have:

$$-4.279131 - (-4.272102) = -0.007029 \text{ au}$$

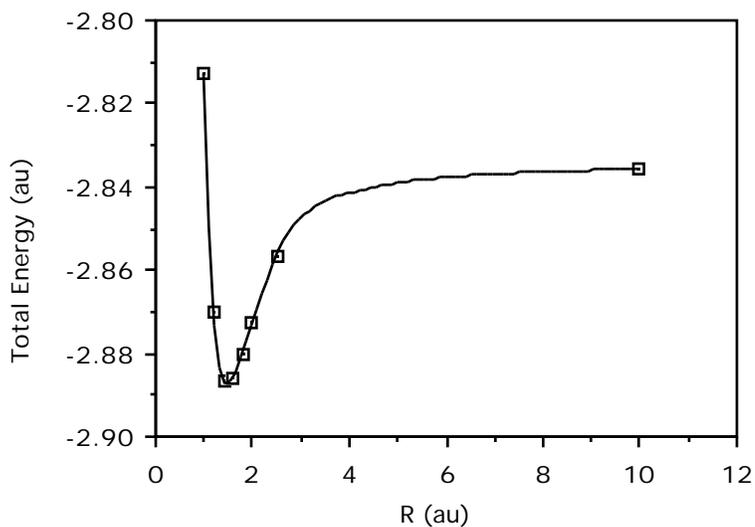
3. STO total energy: -2.8435283  
 STO3G total energy -2.8340561  
 3-21G total energy -2.8864405

The STO3G orbitals were generated as a best fit of 3 primitive gaussians (giving 1 CGTO) to the STO. So, STO3G can at best reproduce the STO result. The 3-21G orbitals are more flexible since there are 2 CGTOs per atom. This gives 4 orbitals (more parameters to optimize) and a lower total energy.

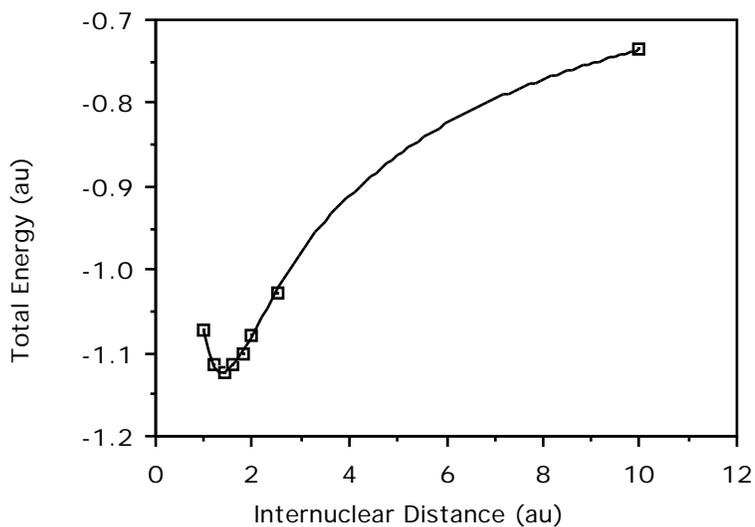
4.

R	HeH <sup>+</sup> Energy	H <sub>2</sub> Energy
1.0	-2.812787056	-1.071953297
1.2	-2.870357513	-1.113775015
1.4	-2.886440516	-1.122933507
1.6	-2.886063576	-1.115567684
1.8	-2.880080938	-1.099872589
2.0	-2.872805595	-1.080269098
2.5	-2.856760263	-1.026927710
10.0	-2.835679293	-0.7361705303

Plotting total energy vs. geometry for HeH<sup>+</sup>:



Plotting total energy vs. geometry for H<sub>2</sub>:



For HeH<sup>+</sup> at R = 10.0 au, the eigenvalues of the converged Fock matrix and the corresponding converged MO-AO coefficients are:

-.1003571E+01	-.4961988E+00	.5864846E+00	.1981702E+01
.4579189E+00	-.8245406E-05	.1532163E-04	.1157140E+01
.6572777E+00	-.4580946E-05	-.6822942E-05	-.1056716E+01
-.1415438E-05	.3734069E+00	.1255539E+01	-.1669342E-04
.1112778E-04	.7173244E+00	-.1096019E+01	.2031348E-04

Notice that this indicates that orbital 1 is a combination of the s functions on He only (dissociating properly to He + H<sup>+</sup>).

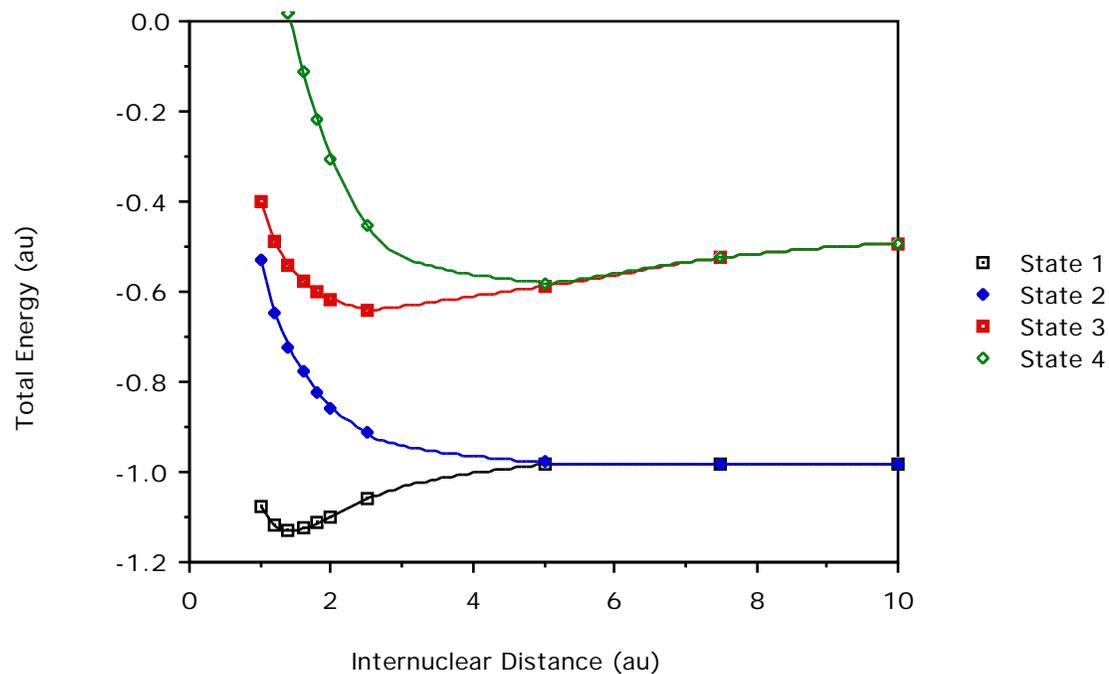
For H<sub>2</sub> at R = 10.0 au, the eigenvalues of the converged Fock matrix and the corresponding converged MO-AO coefficients are:

-.2458041E+00	-.1456223E+00	.1137235E+01	.1137825E+01
.1977649E+00	-.1978204E+00	.1006458E+01	-.7903225E+00
.5632566E+00	-.5628273E+00	-.8179120E+00	.6424941E+00
.1976312E+00	.1979216E+00	.7902887E+00	.1006491E+01
.5629326E+00	.5631776E+00	-.6421731E+00	-.8181460E+00

Notice that this indicates that orbital 1 is a combination of the s functions on both H atoms (dissociating improperly; equal probabilities of H<sub>2</sub> dissociating to two neutral atoms or to a proton plus hydride ion).

5. The H<sub>2</sub> CI result:

R	1 <sub>g<sup>+</sup></sub>	3 <sub>u<sup>+</sup></sub>	1 <sub>u<sup>+</sup></sub>	1 <sub>g<sup>+</sup></sub>
1.0	-1.074970	-0.5323429	-0.3997412	0.3841676
1.2	-1.118442	-0.6450778	-0.4898805	0.1763018
1.4	-1.129904	-0.7221781	-0.5440346	0.0151913
1.6	-1.125582	-0.7787328	-0.5784428	-0.1140074
1.8	-1.113702	-0.8221166	-0.6013855	-0.2190144
2.0	-1.098676	-0.8562555	-0.6172761	-0.3044956
2.5	-1.060052	-0.9141968	-0.6384557	-0.4530645
5.0	-0.9835886	-0.9790545	-0.5879662	-0.5802447
7.5	-0.9806238	-0.9805795	-0.5247415	-0.5246646
10.0	-0.980598	-0.9805982	-0.4914058	-0.4913532



For  $H_2$  at  $R = 1.4$  au, the eigenvalues of the Hamiltonian matrix and the corresponding determinant amplitudes are:

determinant	-1.129904	-0.722178	-0.544035	0.015191
$ 1_g 1_g $	0.99695	0.00000	0.00000	0.07802

$ 1_g 1_u $		0.00000	0.70711	0.70711	0.00000
-------------	--	---------	---------	---------	---------

$$\begin{array}{l}
 |1_g 1_u| \\
 |1_u 1_u|
 \end{array}
 \begin{array}{l}
 | \\
 |
 \end{array}
 \begin{array}{cccc}
 0.00000 & 0.70711 & -0.70711 & 0.00000 \\
 -0.07802 & 0.00000 & 0.00000 & 0.99695
 \end{array}$$

This shows, as expected, the mixing of the first  $1_g^+$  ( $1_g^2$ ) and the 2nd  $1_g^+$  ( $1_u^2$ ) determinants, the

$$^3 u^+ = \left( \frac{1}{\sqrt{2}} (|1_g 1_u| + |1_g 1_u|) \right),$$

$$\text{and the } ^1 u^+ = \left( \frac{1}{\sqrt{2}} (|1_g 1_u| - |1_g 1_u|) \right).$$

Also notice that the first  $1_g^+$  state is the bonding (0.99695 - 0.07802) combination (note specifically the + - combination) and the second  $1_g^+$  state is the antibonding combination (note specifically the ++ combination). The ++ combination always gives a higher energy than the + - combination. Also notice that the 1st and 2nd states ( $1_g^+$  and  $^3 u^+$ ) are dissociating to two neutral atoms and the 3rd and 4th states ( $1_g^+$  and  $^3 u^+$ ) are dissociating to proton/anion combinations. The difference in these energies is the ionization potential of H minus the electron affinity of H.

## Mathematics Review A

### I. Matrices

#### A. Vectors

A vector is an object having  $n$ -components

$$\mathbf{x} = (x_1, x_2, \dots, x_n).$$

These components may represent, for example, the cartesian coordinates of a particle (in which case,  $n=3$ ) or the cartesian coordinates of  $N$  particles (in which case,  $n=3N$ ). Alternatively, the vector components may have nothing what so ever to do with cartesian or other coordinate-system positions.

The numbers  $x_i$  are called the components of the vector  $\mathbf{x}$  in the directions of some  $n$  elementary unit vectors:

$$\begin{aligned} \mathbf{x} &= x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + x_3 \mathbf{e}_3 + \dots + x_n \mathbf{e}_n \\ &= x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ \vdots \\ \vdots \\ \vdots \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 1 \end{pmatrix} \end{aligned}$$

The unit vectors  $\mathbf{e}_i$ , whose exact definition, meaning and interpretation depend on the particular application at hand, are called basis vectors and form the elements of a basis. They are particularly simple to work with because they are orthogonal. This means that their dot products vanish  $\mathbf{e}_i \cdot \mathbf{e}_j = 0$ , unless  $i = j$ . If  $i = j$ , then the scalar or dot product is unity (it is usually convenient, but not necessary, to use bases that are normalized so  $\mathbf{e}_i \cdot \mathbf{e}_i = 1$ ). The shorthand way of representing this information is to write

$$\mathbf{e}_i \cdot \mathbf{e}_j = \langle \mathbf{e}_i \mathbf{e}_j \rangle = \delta_{ij},$$

where  $\delta_{ij}$  is called the Kronecker delta function defined by:

$$\delta_{ij} = 0, \text{ if } i \neq j, \text{ and}$$

$$\delta_{ij} = 1 \text{ if } i = j.$$

The above equation for  $\mathbf{x}$  provides an example of expressing a vector as a linear combination of other vectors (in this case, the basis vectors). The vector  $\mathbf{x}$  is expressed as

a linear combination of the unit vectors  $\mathbf{e}_i$ , and the numbers  $x_i$  are the coefficients in the linear combination. Another way of writing this within the summation notation is:

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i .$$

The idea of a linear combination is an important idea that will be encountered when we discuss how a matrix operator affects a linear combination of vectors.

## B. Products of Matrices and Vectors

If  $\mathbf{M}$  is an  $n \times n$  matrix with elements  $M_{ij}$ , (the first subscript specifies the row number and the second subscript specifies the column number), then the product  $\mathbf{M} \mathbf{x} = \mathbf{y}$  is a vector whose components (when subscripts  $i, j, k$ , etc. appear, they can take any value  $1, 2, \dots, n$  unless otherwise specified) are defined as follows:

$$y_k = \sum_{j=1}^n M_{kj} x_j$$

The vector components  $y_k$  can be understood as either the components of a new vector  $\mathbf{y}$  in the directions of the original basis  $\mathbf{e}_i$  ( $i=1, 2, \dots, n$ ) or as the components of the old vector  $\mathbf{x}$  in the directions of new bases.

There are always these two ways to view a matrix acting on a vector:

1. The operation can be thought of as transforming the vector into a different vector. This view is called the active view (vector in a different place), and is the interpretation we will use most often.

2. The operation can be thought of as expressing the same vector in terms of a different coordinate system or basis. This view is called the passive view.

Some examples may help to clarify these perspectives:

For the matrix-vector product

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

the active interpretation states that the vector is scaled in the  $x$  direction by an amount  $a$ . In the passive interpretation, the original vector is written in terms of new bases  $(a^{-1}, 0)$  and  $(0, 1)$ :

$$\begin{pmatrix} x \\ y \end{pmatrix} = ax \begin{pmatrix} a^{-1} \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \end{pmatrix} .$$

As another example, consider the following matrix multiplication:

$$\mathbf{M} \mathbf{x} = \begin{pmatrix} \cos & -\sin \\ \sin & \cos \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \cos - y \sin \\ x \sin + y \cos \end{pmatrix}$$

In the active interpretation, the vector whose cartesian and polar representations are:

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r\cos \\ r\sin \end{pmatrix},$$

is rotated by an angle to obtain:

$$\begin{aligned} \underline{\mathbf{M}}\mathbf{x} &= \begin{pmatrix} \cos & -\sin \\ \sin & \cos \end{pmatrix} \begin{pmatrix} r\cos \\ r\sin \end{pmatrix} \\ &= \begin{pmatrix} (r\cos \cos - r\sin \sin) \\ (r\cos \sin + r\sin \cos) \end{pmatrix} \\ &= \begin{pmatrix} r\cos(\cos^2 + \sin^2) \\ r\sin(\sin \cos + \cos \sin) \end{pmatrix}. \end{aligned}$$

In the passive interpretation, the original vector  $\mathbf{x}$  is expressed in terms of a new coordinate

system with axes rotated by  $-\theta$  with new bases  $\begin{pmatrix} \cos \\ -\sin \end{pmatrix}$  and  $\begin{pmatrix} \sin \\ \cos \end{pmatrix}$ .

$$\begin{aligned} \begin{pmatrix} x \\ y \end{pmatrix} &= (x\cos - y\sin) \begin{pmatrix} \cos \\ -\sin \end{pmatrix} + (x\sin + y\cos) \begin{pmatrix} \sin \\ \cos \end{pmatrix} \\ &= \begin{pmatrix} x(\cos^2 + \sin^2) + y(\sin \cos - \sin \cos) \\ y(\cos^2 + \sin^2) + x(\sin \cos - \sin \cos) \end{pmatrix} \\ &= \begin{pmatrix} x \\ y \end{pmatrix} \end{aligned}$$

As a general rule, active transformations and passive transformations are inverses of each other; you can do something to a vector or else do the reverse to the coordinate system. The two pictures can be summarized by the following two equations:

(i.)  $\underline{\mathbf{M}}\mathbf{x} = \mathbf{y}$  states the active picture, and

(ii.)  $\mathbf{x} = \underline{\mathbf{M}}^{-1}\mathbf{y}$  states the passive picture.

### C. Matrices as Linear Operators

Matrices are examples of linear operators for which

$$\underline{M}(\underline{ax} + \underline{by}) = a\underline{Mx} + b\underline{My},$$

which can easily be demonstrated by examining the components:

$$\begin{aligned} [\underline{M}(\underline{ax} + \underline{by})]_i &= \sum_k M_{ik}(ax_k + by_k) \\ &= a \sum_k M_{ik}x_k + b \sum_k M_{ik}y_k \\ &= a(\underline{Mx})_i + b(\underline{My})_i. \end{aligned}$$

One can also see that this property holds for a linear combination of many vectors rather than for the two considered above.

We can visualize how the action of a matrix on arbitrary vectors can be expressed if one knows its action on the elementary basis vectors. Given the expansion of  $\underline{x}$  in the  $\underline{e}_i$ ,

$$\underline{x} = \sum_i x_i \underline{e}_i,$$

one can write

$$\underline{Mx} = \sum_i x_i \underline{Me}_i.$$

Using the fact that all of the components of  $\underline{e}_i$  are zero except one,  $(\underline{e}_i)_i = 1$ , we see that

$$(\underline{Me}_i)_k = \sum_j M_{kj}(\underline{e}_i)_j = M_{ki}$$

This equation tells us that the  $i$ -th column of a matrix,  $\underline{M}$ , contains the result of operating on the  $i$ -th unit vector  $\underline{e}_i$  with the matrix. More specifically, the element  $M_{ki}$  in the  $k$ -th row and  $i$ -th column is the component of  $\underline{Me}_i$  in the direction of the  $\underline{e}_k$  unit vector. As a generalization, we can construct any matrix by first deciding how the matrix affects the elementary unit vectors and then placing the resulting vectors as the columns of the matrix.

## II. Properties of General $n \times n$ (Square) Matrices

The following operations involving square matrices each of the same dimension are useful to express in terms of the individual matrix elements:

1. Sum of matrices;  $\underline{A} + \underline{B} = \underline{C}$  if  $A_{ij} + B_{ij} = C_{ij}$
2. Scalar multiplication;  $c\underline{M} = \underline{N}$  if  $c M_{ij} = N_{ij}$

3. Matrix multiplication;  $\underline{A}\underline{B} = \underline{C}$  if 
$$C_{ij} = \sum_k A_{ik} B_{kj}$$

4. Determinant of a matrix;  
The determinant is defined inductively

for  $n = 1$  
$$A = \det(A) = A_{11}$$

for  $n > 1$  
$$A = \det(A) = \sum_j A_{ij} \det(a_{ij}) (-1)^{(i+j)}$$
 ; where 
$$j=1,2,\dots,n$$
 and  $a_{ij}$  is the minor matrix obtained by deleting the  $i$ th row and  $j$ th column.

5. There are certain special matrices that are important to know:

A. The zero matrix; 
$$0_{ij} = 0 \text{ for } i = 1,2,\dots,n \text{ and } j = 1,2,\dots,n$$

B. The identity matrix; 
$$I_{ij} = \delta_{ij}$$

(Note  $(\underline{I} \underline{M})_{ij} = \sum_k \delta_{ik} M_{kj} = M_{ij}$ , so  $\underline{I} \underline{M} = \underline{M} \underline{I} = \underline{M}$ )

6. Transpose of a matrix;  $(M^T)_{ij} = M_{ji}$

7. Complex Conjugate of a Matrix;  $(M^*)_{ij} = M^*_{ij}$

8. Adjoint of a Matrix;  $(M^\dagger)_{ij} = M^*_{ji} = (M^T)^*_{ij}$

9. Inverse of a Matrix; if  $\underline{N} \underline{M} = \underline{M} \underline{N} = \underline{I}$  then  $\underline{N} = \underline{M}^{-1}$

10. Trace (or Character) of Matrix; 
$$\text{Tr}(\underline{M}) = \sum_i M_{ii}$$
  
(sum of diagonal elements)

### III. Special Kinds of Square Matrices

If a matrix obeys certain conditions, we give it one of several special names. These names include the following:

A. Diagonal Matrix:  $D_{ij} = d_i \delta_{ij} = d_j \delta_{ij}$

B. Real Matrix:  $\underline{M} = \underline{M}^*$  or  $M_{ij} = M_{ij}^*$  (real elements)

C. Symmetric Matrix:  $\underline{M} = \underline{M}^T$  or  $M_{ij} = M_{ji}$  (symmetric about main diagonal)

D. Hermitian Matrix:  $\underline{M} = \underline{M}^*$  or  $M_{ij} = M_{ji}^*$

E. Unitary Matrix:  $\underline{M} = \underline{M}^{-1}$

F. Real Orthogonal  $\underline{M} = \underline{M}^T = \underline{M}^{-1}$

#### IV. Eigenvalues and Eigenvectors of a Square Matrix

An eigenvector of a matrix,  $\underline{M}$ , is a vector such that

$$\underline{M}\mathbf{v} = \lambda \mathbf{v}$$

where  $\lambda$  is called the eigenvalue. An eigenvector thus has the property that when it is multiplied by the matrix, the direction of the resultant vector is unchanged. The length, however, is altered by a factor  $\lambda$ . Note that any multiple of an eigenvector is also an eigenvector, which we demonstrate as follows:

$$\underline{M}(a\mathbf{v}) = a\underline{M}\mathbf{v} = a \lambda \mathbf{v} = \lambda(a\mathbf{v}).$$

Hence, an eigenvector can be thought of as defining a direction in n-dimensional space. The length (normalization) of the eigenvector is arbitrary; we are free to choose the length to be anything we please. Usually we choose the length to be unity (because, in quantum mechanics, our vectors usually represent some wavefunction that we wish to obey a normalization condition).

The basic eigenvalue equation can be rewritten as

$$(\underline{M} - \lambda \underline{I})\mathbf{v} = 0$$

or, in an element-by-element manner, as

$$\begin{aligned} (M_{11} - \lambda)v_1 + M_{12}v_2 + M_{13}v_3 + \dots + M_{1n}v_n &= 0 \\ M_{21}v_1 + (M_{22} - \lambda)v_2 + M_{23}v_3 + \dots + M_{2n}v_n &= 0 \end{aligned}$$

$$M_{n1}v_1 + M_{n2}v_2 + M_{n3}v_3 + \dots + (M_{nn} - \lambda)v_n = 0.$$

If you try to solve these n equations for all of the elements of the  $\mathbf{v}$  vector ( $v_1 \dots v_n$ ), you can eliminate one variable using one equation, a second variable using a second equation, etc., because the equations are linear. For example you could solve for  $v_1$  using the first equation and then substitute for  $v_1$  in the second equation as you solve for  $v_2$ , etc. Then when you come to the nth equation, you would have n-1 of the variables expressed in terms of the one remaining variable,  $v_n$ .

However, you find that you cannot use the remaining equation to solve for the value of  $v_n$ ; the last equation is found to appear in the form

$$(C - \lambda) v_n = 0$$

once the  $v_1, v_2, v_3$ , and  $v_{n-1}$  are expressed in terms of  $v_n$ . We should not really have expected to solve for  $v_n$  since, as we saw above, the length of the vector  $\mathbf{v}$  is not determined from the set of eigenvalue equations. You find that the only solution is  $v_n = 0$ , which then implies that all of the other  $v_k = 0$  because you expressed them in terms of  $v_n$ , unless the eigenvalue  $\lambda$  is chosen to obey  $\lambda = C$ .

Upon analyzing what has gone into the  $C$  element, one finds that the  $v_k$  ( $k=1,2,3,\dots,n-1$ ) were eliminated in favor of  $v_n$  by successively combining rows of the  $(\underline{M} - \lambda \underline{I})$  matrix. Thus,  $(C - \lambda)$  can vanish if and only if the last row of  $(\underline{M} - \lambda \underline{I})$  is a linear combination of the other  $n-1$  rows. A theorem dealing with determinants states that the rows of a matrix are linearly dependent (i.e., one row is a linear combination of the rest) if and only if the determinant of the matrix is zero. We can therefore make the eigenvalue equation have a solution  $\mathbf{v}$  by adjusting  $\lambda$  so the determinant of  $(\underline{M} - \lambda \underline{I})$  vanishes.

### A. Finding the Eigenvalues

In summary, to solve an eigenvalue equation, we first solve the determinantal equation:

$$\det(\underline{M} - \lambda \underline{I}) = 0.$$

Using the definition of a determinant, one can see that expanding the determinant results in an  $n^{\text{th}}$  order polynomial in

$$a_n(\lambda) \lambda^n + a_{n-1}(\lambda) \lambda^{n-1} + \dots + a_1(\lambda) \lambda + a_0(\lambda) = 0,$$

where the coefficients  $a_i(\lambda)$  depend on the matrix elements of  $\underline{M}$ . A theorem in algebra shows that such an equation always has  $n$  roots some or all of which may be complex (e.g., a quadratic equation has 2 roots). Thus there are  $n$  different ways of adjusting the parameter  $\lambda$  so the determinant vanishes. Each of these solutions is a candidate for use as  $\lambda$  in subsequently solving for the  $\mathbf{v}$  vector's coefficients. That is, each  $\lambda$  value has its own vector  $\mathbf{v}$ .

### B. Finding the Eigenvectors

One can substitute each of these  $n$  values ( $\lambda_k, k = 1,2,\dots,n$ ) back into the eigenvalue equation, one at a time, and solve for the  $n$  eigenvectors  $\mathbf{v}(k)$ . By using one of the  $\lambda_k$  values, the  $n^{\text{th}}$  equation is guaranteed to be equal to zero, so you use  $n-1$  of the equations to solve for  $n-1$  of the components in terms of the  $n^{\text{th}}$  one. The eigenvectors will then be determined up to a multiplicative factor which we then fix by requiring normalization:

$$v_i^*(k)v_i(k) = \langle \mathbf{v}(k) | \mathbf{v}(k) \rangle = 1.$$

This expression now defines the dot or inner product (Hermitian inner product) for vectors which can have complex valued components. We use this definition so the dot product of a complex valued vector with itself is real.

In summary, we have shown how to solve the n equations:

$$M_{ij}v_j(k) = \lambda_k v_i(k); k = 1, 2, \dots, n.$$

for the eigenvalues of the matrix  $M_{ij}$  and the corresponding normalized eigenvectors  $\mathbf{v}(k)$ ,  $k = 1, 2, \dots, n$ . Now let us work an example that is chosen to illustrate the concepts we have learned as well as an additional complication, that of eigenvalue degeneracy.

### C. Examples

Consider the following real symmetric matrix:

$$\mathbf{M} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & \frac{5}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{5}{2} \end{pmatrix}$$

The set of eigenvalue-eigenvector equations has non-trivial ( $\mathbf{v}(k) \neq 0$  is "trivial") solutions if

$$\mathbf{M} - \lambda \mathbf{I} = 0.$$

In our example, this amounts to

$$\begin{pmatrix} 3 - \lambda & 0 & 0 \\ 0 & \frac{5}{2} - \lambda & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{5}{2} - \lambda \end{pmatrix} = 0,$$

or

$$(3 - \lambda) \begin{vmatrix} \frac{5}{2} - \lambda & \frac{1}{2} \\ \frac{1}{2} & \frac{5}{2} - \lambda \end{vmatrix} = 0,$$

or,

$$(3 - \lambda) \left( \frac{5}{2} - \lambda \right)^2 - \frac{1}{4} = 0,$$

or,

$$(3 - k) [ k^2 - 5k + 6 ] = 0 = (3 - k)(k - 3)(k - 2).$$

There are three real solutions to this cubic equation (why all the solutions are real in this case for which the M matrix is real and symmetric will be made clear later):

$$1) \quad \lambda_1 = 3, 2) \quad \lambda_2 = 3, \text{ and } 3) \quad \lambda_3 = 2.$$

Notice that the eigenvalue  $\lambda_k = 3$  appears twice; we say that the eigenvalue  $\lambda_k = 3$  is doubly degenerate.  $\lambda_k = 2$  is a non-degenerate eigenvalue.

The eigenvectors  $\mathbf{v}(k)$  are found by plugging the above values for  $\lambda_k$  into the basic eigenvalue equation

$$\underline{\mathbf{M}} \mathbf{v}(k) = \lambda_k \mathbf{v}(k)$$

For the non-degenerate eigenvalue ( $\lambda_3 = 2$ ), this gives:

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & \frac{5}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{5}{2} \end{pmatrix} \begin{pmatrix} v_1(3) \\ v_2(3) \\ v_3(3) \end{pmatrix} = 2 \begin{pmatrix} v_1(3) \\ v_2(3) \\ v_3(3) \end{pmatrix} .$$

The following algebraic steps can then be followed:

- i.  $3v_1(3) = 2v_1(3)$  implies that  $v_1(3) = 0$ ,
- ii.  $\frac{5}{2} v_2(3) + \frac{1}{2} v_3(3) = 2v_2(3)$ , and
- iii.  $\frac{1}{2} v_2(3) + \frac{5}{2} v_3(3) = 2v_3(3)$ .

The last two equations can not be solved for both  $v_2(3)$  and  $v_3(3)$ . To see the trouble, multiply equation iii. by 5 and subtract it from equation ii. to obtain

$$-12 v_3(3) = 2v_2(3) - 10v_3(3),$$

which implies that  $v_3(3) = -v_2(3)$ . Now, substitute this result into the equation ii. to obtain

$$\frac{5}{2} v_2(3) + \frac{1}{2} (-v_2(3)) = 2v_2(3),$$

or,

$$2v_2(3) = 2v_2(3).$$

This is a trivial identity; it does not allow you to solve for  $v_2(3)$ .

Hence, for this non-degenerate root, one is able to solve for all of the  $v_j$  elements in terms of one element that is yet undetermined.

As in all matrix eigenvalue problems, we are able to express (n-1) elements of the eigenvector  $\mathbf{v}(k)$  in terms of one remaining element. However, we can never solve for this one last element. So, for convenience, we impose one more constraint (equation to be

obeyed) which allows us to solve for the remaining element of  $\mathbf{v}(\mathbf{k})$ . We require that the eigenvectors be normalized:

$$\langle \mathbf{v}(\mathbf{k}) | \mathbf{v}(\mathbf{k}) \rangle = \sum_a \mathbf{v}_a^*(\mathbf{k}) \mathbf{v}_a(\mathbf{k}) = 1.$$

In our example, this means that

$$v_1^2(3) + v_2^2(3) + v_3^2(3) = 1,$$

or,

$$0^2 + v_2^2(3) + (-v_2(3))^2 = 1,$$

which implies that  $v_2(3) = \pm \frac{1}{\sqrt{2}}$ .

So,

$$v_3(3) = \pm \frac{1}{\sqrt{2}},$$

and finally the vector is given by:

$$\mathbf{v}(3) = \pm \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}.$$

Note that even after requiring normalization, there is still an indeterminacy in the sign of  $\mathbf{v}(3)$ . The eigenvalue equation, as we recall, only specifies a direction in space. The sense or sign is not determined. We can choose either sign we prefer.

Finding the first eigenvector was not too difficult. The degenerate eigenvectors are more difficult to find. For  $\lambda_1 = \lambda_2 = 3$ ,

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & \frac{5}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{5}{2} \end{pmatrix} \begin{pmatrix} v_1(1) \\ v_2(1) \\ v_3(1) \end{pmatrix} = 3 \begin{pmatrix} v_1(1) \\ v_2(1) \\ v_3(1) \end{pmatrix}.$$

Again, the algebraic equations can be written as follows:

- i.  $3v_1(1) = 3v_1(1)$ ; this tells us nothing!
- ii.  $\frac{5}{2} v_2(1) + \frac{1}{2} v_3(1) = 3v_2(1)$ , and

$$\text{iii. } \frac{1}{2} v_2(1) + \frac{5}{2} v_3(1) = 3v_3(1).$$

If we multiply equation iii. by 5 and subtract it from equation ii., we obtain:

$$-12v_3(1) = -15v_3(1) + 3v_2(1),$$

or

$$3v_3(1) = 3v_2(1),$$

which implies that  $v_3(1) = v_2(1)$ . So far, all we have is  $v_3(1) = v_2(1)$ ; we don't know  $v_1(1)$  nor do we know either  $v_3(1)$  or  $v_2(1)$ , and we have used all three equations.

Normalization provides one more equation  $v_1^2(1) + v_2^2(1) + (v_2(1))^2 = 1$ , but we are still in a situation with more unknowns (2) than equations (1).

One might think that by restricting our eigenvectors to be orthogonal as well as normalized, we might overcome this difficulty (however, such is not the case, as we now show).

For our vectors, the constraint that the nondegenerate vector  $\mathbf{v}(3)$  be orthogonal to the one we are trying to find  $\mathbf{v}(1)$ , amounts to

$$\langle \mathbf{v}(3) \mathbf{v}(1) \rangle = 0$$

$$v_1(3)^* v_1(1) + v_2(3)^* v_2(1) + v_3(3)^* v_3(1) = 0$$

$$0 v_1(1) \pm \frac{1}{\sqrt{2}} v_2(1) - \frac{1}{\sqrt{2}} (v_2(1)) = 0.$$

We see that  $\mathbf{v}(3)$  and  $\mathbf{v}(1)$  are already orthogonal regardless of how  $v_2(1)$  and  $v_3(1)$  turn out. This is shown below to be guaranteed because  $\mathbf{v}(1)$  and  $\mathbf{v}(3)$  have different eigenvalues (i.e., two eigenvectors belonging to different eigenvalues of any symmetric or hermitian matrix must be orthonormal). Hence, this first attempt at finding additional equations to use has failed.

What about the two degenerate eigenvectors  $\mathbf{v}(1)$  and  $\mathbf{v}(2)$ ? Are they also orthonormal? So far, we know that these two eigenvectors have the structure

$$\mathbf{v}(1) = \begin{pmatrix} v_1(1) \\ v_2(1) \\ v_3(1) \end{pmatrix}, \text{ with } 1 = v_1^2(1) + 2v_2^2(1).$$

If we go through all of the above steps for  $\mathbf{v}(2)$  with  $\ell = 2$ , we will find that this vector obeys exactly the same set of equations

$$\mathbf{v}(2) = \begin{pmatrix} v_1(2) \\ v_2(2) \\ v_3(2) \end{pmatrix}, \text{ with } 1 = v_1^2(2) + 2v_2^2(2).$$

We showed above that  $\langle \mathbf{v}(1) \mathbf{v}(3) \rangle = 0$ , and it is easy to show that  $\langle \mathbf{v}(2) \mathbf{v}(3) \rangle = 0$  because the elements of  $\mathbf{v}(2)$ , thus far, obey the same equations as  $\mathbf{v}(1)$ .

If we also wish to make the two degenerate eigenvectors orthogonal to one another

$$\langle \mathbf{v}(1) \mathbf{v}(2) \rangle = 0,$$

then we obtain additional relationships among our yet-undetermined vector amplitudes. In particular, we obtain

$$v_1(1)v_1(2) + v_2(1)v_2(2) + v_3(1)v_3(2) = 0,$$

or,

$$v_1(1)v_1(2) + 2v_2(1)v_2(2) = 0.$$

Although progress has been made, we still have four unknowns  $v_1(1), v_2(1); v_1(2), v_2(2)$  and only three equations:

$$0 = v_1(1) v_1(2) + 2v_2(1) v_2(2) = 0,$$

$$1 = v_1(1) v_1(1) + 2v_2(1) v_2(1) = 1, \text{ and}$$

$$1 = v_1(2) v_1(2) + 2v_2(2) v_2(2) = 1.$$

It appears as though we are stuck again. We are; but for good reasons. We are trying to find two vectors  $\mathbf{v}(1)$  and  $\mathbf{v}(2)$  that are orthonormal and are eigenvectors of  $\underline{\mathbf{M}}$  having eigenvalue equal to 3. Suppose that we do find two such vectors. Because  $\underline{\mathbf{M}}$  is a linear operator, any two vectors generated by taking linear combinations of these two vectors would also be eigenvectors of  $\underline{\mathbf{M}}$ . There is a degree of freedom, that of recombining  $\mathbf{v}(1)$  and  $\mathbf{v}(2)$ , which can not be determined by insisting that the two vectors be eigenfunctions. Thus, in this degenerate-eigenvalue case our requirements do not give a unique pair of eigenvectors. They just tell us the two-dimensional space in which the acceptable eigenvectors lie (this difficulty does not arise in nondegenerate cases because one-dimensional spaces have no flexibility other than a sign.).

So to find an acceptable pair of vectors, we are free to make an additional choice. For example, we can choose one of the four unknown components of  $\mathbf{v}(1)$  and  $\mathbf{v}(2)$  equal to zero. Let us make the choice  $v_1(1) = 0$ . Then the above equations can be solved for the other elements of  $\mathbf{v}(1)$  to give  $v_2(1) = \pm \frac{1}{\sqrt{2}} = v_3(1)$ . The orthogonality between  $\mathbf{v}(1)$  and  $\mathbf{v}(2)$  then gives  $0 = 2 \pm \frac{1}{\sqrt{2}} v_2(2)$ , which implies that  $v_2(2) = v_3(2) = 0$ ; the remaining equation involving  $\mathbf{v}(2)$  then gives  $v_1(2) = \pm 1$ .

In summary, we have now found a specific solution once the choice  $v_1(1) = 0$  is made:

$$\mathbf{v}(1) = \pm \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}, \mathbf{v}(2) = \pm \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \text{ and } \mathbf{v}(3) = \pm \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Other choices for  $\mathbf{v}(1)$  will yield different specific solutions.

## V. Properties of Eigenvalues and Eigenvectors of Hermitian Matrices

The above example illustrates many of the properties of the matrices that we will most commonly encounter in quantum mechanics. It is important to examine these properties in more detail and to learn about other characteristics that Hermitian matrices have.

### A. Outer product

Given any vector  $\mathbf{v}$ , we can form a square matrix denoted  $|\mathbf{v}(i)\rangle\langle\mathbf{v}(i)|$ , whose elements are defined as follows:

$$|\mathbf{v}(i)\rangle\langle\mathbf{v}(i)| = \begin{pmatrix} \mathbf{v}_1^*(i)\mathbf{v}_1(i) & \mathbf{v}_1^*(i)\mathbf{v}_2(i) & \dots & \mathbf{v}_1^*(i)\mathbf{v}_n(i) \\ \mathbf{v}_2^*(i)\mathbf{v}_1(i) & \mathbf{v}_2^*(i)\mathbf{v}_2(i) & \dots & \mathbf{v}_2^*(i)\mathbf{v}_n(i) \\ \dots & \dots & \dots & \dots \\ \mathbf{v}_n^*(i)\mathbf{v}_1(i) & \mathbf{v}_n^*(i)\mathbf{v}_2(i) & \dots & \mathbf{v}_n^*(i)\mathbf{v}_n(i) \end{pmatrix}$$

We can use this matrix to project onto the component of a vector in the  $\mathbf{v}(i)$  direction. For the example we have been considering, if we form the projector onto the  $\mathbf{v}(1)$  vector, we obtain

$$\begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix},$$

for  $\mathbf{v}(2)$ , we get

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

and for  $\mathbf{v}(3)$  we find

$$\begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} .$$

These three projection matrices play important roles in what follows.

### B. Completeness Relation or Resolution of the Identity

The set of eigenvectors of any Hermitian matrix form a complete set over the space they span in the sense that the sum of the projection matrices constructed from these eigenvectors gives an exact representation of the identity matrix.

$$\sum_i |\mathbf{v}(i)\rangle\langle\mathbf{v}(i)| = \mathbf{I}.$$

For the specific matrix we have been using as an example, this relation reads as follows:

$$\begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} [1 \ 0 \ 0] + \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} =$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} .$$

Physically, this means that when you project onto the components of a vector in these three directions, you don't lose any of the vector. This happens because our vectors are orthogonal and complete. The completeness relation means that any vector in this three-dimensional space can be written in terms of  $\mathbf{v}(1)$ ,  $\mathbf{v}(2)$ , and  $\mathbf{v}(3)$  (i.e., we can use  $\mathbf{v}(1), \mathbf{v}(2), \mathbf{v}(3)$  as a new set of bases instead of  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ ).

Let us consider an example in which the following vector is expanded or written in terms of our three eigenvectors:

$$\mathbf{f} = \begin{pmatrix} 7 \\ -9 \\ 12 \end{pmatrix} = a_1 \mathbf{v}(1) + a_2 \mathbf{v}(2) + a_3 \mathbf{v}(3)$$

The task at hand is to determine the expansion coefficients, the  $a_j$ . These coefficients are the projections of the given  $\mathbf{f}$  vector onto each of the  $\mathbf{v}(i)$  directions:

$$\langle \mathbf{v}(1) \mathbf{f} \rangle = a_1 ,$$

since,

$$\langle \mathbf{v}(1) \mathbf{v}(2) \rangle = \langle \mathbf{v}(1) \mathbf{v}(3) \rangle = 0.$$

Using our three  $\mathbf{v}_i$  vectors and the above  $\mathbf{f}$  vector, the following three expansion coefficients are obtained:

$$a_1 = \begin{matrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 7 & -9 & 12 \end{matrix} = \frac{3}{\sqrt{2}}$$

$$\langle \mathbf{v}(2) \mathbf{f} \rangle = 7$$

$$\langle \mathbf{v}(3) \mathbf{f} \rangle = -\frac{21}{\sqrt{2}}$$

Therefore,  $\mathbf{f}$  can be written as:

$$\mathbf{f} = \begin{matrix} 7 \\ -9 \\ 12 \end{matrix} = \frac{3}{\sqrt{2}} \begin{matrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{matrix} + 7 \begin{matrix} 1 \\ 0 \\ 0 \end{matrix} - \frac{21}{\sqrt{2}} \begin{matrix} 0 \\ \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{matrix}$$

This works for any vector  $\mathbf{f}$ , and we could write the process in general in terms of the resolution of the identity as

$$\mathbf{f} = \underline{\mathbf{I}}\mathbf{f} = \sum_k |\mathbf{v}(k)\rangle \langle \mathbf{v}(k)| \mathbf{f}\rangle = \sum_k |\mathbf{v}(k)\rangle a_k ,$$

This is how we will most commonly make use of the completeness relation as it pertains to the eigenvectors of Hermitian matrices.

### C. Spectral Resolution of M

It turns out that not only the identity matrix  $\underline{\mathbf{I}}$  but also the matrix  $\underline{\mathbf{M}}$  itself can be expressed in terms of the eigenvalues and eigenvectors. In the so-called spectral representation of  $\underline{\mathbf{M}}$ , we have

$$\underline{M} = \sum_k |\mathbf{v}(k)\rangle \langle \mathbf{v}(k)|$$

In the example we have been using, the three terms in this sum read

$$\begin{array}{ccccccc}
 & 0 & 0 & 0 & & 0 & 0 & 0 & & 3 & 0 & 0 \\
 & & & & 1 & 0 & 0 & & & & & \\
 3 & 0 & \frac{1}{2} & \frac{1}{2} & + & 3 & 0 & 0 & 0 & + & 2 & 0 & \frac{1}{2} & -\frac{1}{2} & = & 0 & \frac{5}{2} & \frac{1}{2} & = \underline{M} \\
 & 0 & \frac{1}{2} & \frac{1}{2} & & 0 & 0 & 0 & & 0 & -\frac{1}{2} & \frac{1}{2} & & 0 & \frac{1}{2} & \frac{5}{2}
 \end{array}$$

This means that a matrix is totally determined if we know its eigenvalues and eigenvectors.

#### D. Eigenvalues of Hermitian Matrices are Real Numbers

A matrix can be expressed in terms of any complete set of vectors. For  $n \times n$  matrices, a complete set is any  $n$  linearly independent vectors. For a set of vectors  $|\mathbf{k}\rangle$ ,  $k = 1, 2, \dots, n$ , the matrix elements of  $\underline{M}$  are denoted  $M_{jk}$  or  $\langle \mathbf{j} | \underline{M} | \mathbf{k} \rangle$ . If the matrix is Hermitian then

$$\langle \mathbf{j} | \underline{M} | \mathbf{k} \rangle = \langle \mathbf{k} | \underline{M} | \mathbf{j} \rangle^*$$

If the vectors  $|\mathbf{k}\rangle$  are eigenvectors of  $\underline{M}$ , that is, if  $\underline{M} |\mathbf{k}\rangle = \lambda_k |\mathbf{k}\rangle$ , then the eigenvalues are real. This can be shown as follows:

$$\langle \mathbf{k} | \underline{M} | \mathbf{k} \rangle = \lambda_k \langle \mathbf{k} | \mathbf{k} \rangle = \langle \mathbf{k} | \underline{M} | \mathbf{k} \rangle^* = \lambda_k^* \langle \mathbf{k} | \mathbf{k} \rangle,$$

so  $\lambda_k = \lambda_k^*$ . This is a very important result because it forms the basis of the use of Hermitian operators in quantum mechanics; such operators are used because experiments yield real results, so to connect with experimental reality, only hermitian operators must appear.

#### E. Nondegenerate Eigenvectors of Hermitian Matrices are Orthogonal

If two eigenvalues are different,  $\lambda_k \neq \lambda_j$ , then

$$\langle \mathbf{k} | \underline{M} | \mathbf{j} \rangle = \lambda_j \langle \mathbf{k} | \mathbf{j} \rangle = \langle \mathbf{j} | \underline{M} | \mathbf{k} \rangle^* = \lambda_k^* \langle \mathbf{j} | \mathbf{k} \rangle^* = \lambda_k \langle \mathbf{k} | \mathbf{j} \rangle,$$

which implies that  $(\lambda_k - \lambda_j) \langle \mathbf{k} | \mathbf{j} \rangle = 0$ . Since, by assumption,  $\lambda_k \neq \lambda_j$ , it must be that  $\langle \mathbf{k} | \mathbf{j} \rangle = 0$ . In other words, the eigenvectors are orthogonal. We saw this earlier in our example when we "discovered" that  $\mathbf{v}(3)$  was automatically orthogonal to  $\mathbf{v}(1)$  and to  $\mathbf{v}(2)$ .

If one has degenerate eigenvalues,  $\lambda_k = \lambda_j$  for some  $k$  and  $j$ , then the corresponding eigenvectors are not automatically orthogonal to one another (they are orthogonal to other eigenvectors), but the degenerate eigenvectors can always be chosen to be orthogonal. We also encountered this in our earlier example.

In all cases then, one can find  $n$  orthonormal eigenvectors (remember we required  $\langle \mathbf{k} | \mathbf{k} \rangle = 1$  as an additional condition so that our amplitudes could be interpreted in terms of probabilities). Since any vector in an  $n$ -dimensional space can be expressed as a linear combination of  $n$  orthonormal vectors, the eigenvectors form a complete basis set. This is why the so-called resolution of the identity, in which the unit matrix can be expressed in terms of the  $n$  eigenvectors of  $\underline{M}$ , holds for all hermitian matrices.

#### F. Diagonalizing a Matrix using its Eigenvectors

The eigenvectors of  $\underline{M}$  can be used to form a matrix that diagonalizes  $\underline{M}$ . This matrix  $\underline{S}$  is defined such that the  $k^{\text{th}}$  column of  $\underline{S}$  contains the elements of  $\mathbf{v}(k)$

$$\underline{S} = \begin{pmatrix} v_1(1) & v_1(2) & \dots & v_1(n) \\ v_2(1) & v_2(2) & \dots & v_2(n) \\ v_3(1) & v_3(2) & \dots & v_3(n) \\ \vdots & \vdots & \ddots & \vdots \\ v_n(1) & v_n(2) & \dots & v_n(n) \end{pmatrix} .$$

Then using the above equation one can write the action of  $\underline{M}$  on  $\underline{S}$  as

$$M_{ij} S_{jk} = \sum_j M_{ij} v_j(k) = \sum_j v_i(k) = S_{ik} \epsilon_k .$$

Now consider another matrix  $\underline{\epsilon}$  which is a diagonal matrix with diagonal elements  $\epsilon_k$ , i.e.

$$\epsilon_{ik} = \epsilon_k \delta_{ik} .$$

One can easily show using the  $\delta_{jk}$  matrix that

$$\sum_j S_{ij} \delta_{jk} \epsilon_k = S_{ik} \epsilon_k ,$$

since the only non-zero term in the sum is the one in which  $j = k$ . Thus, comparing with the previous result for the action of  $\underline{M}$  on  $\underline{S}$ ,

$$M_{ij} S_{jk} = \sum_j S_{ij} \delta_{jk} \epsilon_k .$$

These are just the  $i,k$ th matrix elements of the matrix equation

$$\underline{M} \underline{S} = \underline{S} \underline{\epsilon} .$$

Let us assume that an inverse  $\underline{S}^{-1}$  of  $\underline{S}$  exists. Then multiply the above equation on the left by  $\underline{S}^{-1}$  to obtain

$$\underline{S}^{-1} \underline{M} \underline{S} = \underline{S}^{-1} \underline{S} \underline{\Lambda} = \underline{I} \underline{\Lambda} = \underline{\Lambda}.$$

This identity illustrates a so-called similarity transform of  $\underline{M}$  using  $\underline{S}$ . Since  $\underline{\Lambda}$  is diagonal, we say that the similarity transform  $\underline{S}$  diagonalizes  $\underline{M}$ . Notice this would still work if we had used the eigenvectors  $\mathbf{v}(k)$  in a different order. We would just get  $\underline{\Lambda}$  with the diagonal elements in a different order. Note also that the eigenvalues of  $\underline{M}$  are the same as those of  $\underline{\Lambda}$  since the eigenvalues of  $\underline{\Lambda}$  are just the diagonal elements  $\lambda_k$  with eigenvectors  $\mathbf{e}_k$  (the elementary unit vectors)

$$\underline{\Lambda} \mathbf{e}_k = \lambda_k \mathbf{e}_k, \text{ or, } (\lambda_k \mathbf{e}_k)_i = \sum_j \lambda_{kj} \mathbf{e}_j = \lambda_{ki} = \lambda_k (\mathbf{e}_k)_i.$$

### G. The Trace of a Matrix is the Sum of its Eigenvalues

Based on the above similarity transform, we can now show that the trace of a matrix (i.e., the sum of its diagonal elements) is independent of the representation in which the matrix is formed, and, in particular, the trace is equal to the sum of the eigenvalues of the matrix. The proof of this theorem proceeds as follows:

$$\begin{aligned} \sum_k \lambda_k &= \text{Tr}(\underline{\Lambda}) = \sum_k (\underline{S}^{-1} \underline{M} \underline{S})_{kk} = \sum_{kij} (\underline{S}_{ki}^{-1} \underline{M}_{ij} \underline{S}_{jk}) \\ &= \sum_{ij} \underline{M}_{ij} \sum_k \underline{S}_{jk} \underline{S}_{ki}^{-1} = \sum_{ij} \underline{M}_{ij} \delta_{ji} = \sum_i \underline{M}_{ii} = \text{Tr}(\underline{M}) \end{aligned}$$

### H. The Determinant of a Matrix is the Product of its Eigenvalues

This theorem,  $\det(\underline{M}) = \det(\underline{\Lambda})$ , can be proven, by using the theorem that  $\det(\underline{AB}) = \det(\underline{A}) \det(\underline{B})$ , inductively within the expansion by minors of the determinant

$$\begin{aligned} \lambda_1 \lambda_2 \lambda_3 \dots \lambda_n &= \det(\underline{\Lambda}) = \det(\underline{S}^{-1} \underline{M} \underline{S}) = \det(\underline{S}^{-1}) \det(\underline{M}) \det(\underline{S}) \\ &= \det(\underline{M}) \det(\underline{S}^{-1}) \det(\underline{S}) \\ &= \det(\underline{M}) \det(\underline{S}^{-1} \underline{S}) \\ &= \det(\underline{M}) \det(\underline{I}) = \det(\underline{MI}) = \det(\underline{M}). \end{aligned}$$

### I. Invariance of Matrix Identities to Similarity Transforms

We will see later that performing a similarity transform expresses a matrix in a different basis (coordinate system). If, for any matrix  $\underline{A}$ , we call  $\underline{S}^{-1} \underline{A} \underline{S} = \underline{A}'$ , we can show that performing a similarity transform on a matrix equation leaves the form of the equation unchanged. For example, if

$$\underline{A} + \underline{B} \underline{C} = \underline{P} \underline{Q} \underline{R},$$

performing a similarity transform on both sides, and remembering that  $\underline{S}^{-1} \underline{S} = \underline{I}$ , one finds

$$\begin{aligned} \underline{S}^{-1}(\underline{A} + \underline{B} \underline{C})\underline{S} &= \underline{S}^{-1} \underline{P} \underline{Q} \underline{R} \underline{S} = \\ \underline{S}^{-1} \underline{A} \underline{S} + \underline{S}^{-1} \underline{B} (\underline{S} \underline{S}^{-1}) \underline{C} \underline{S} &= \underline{S}^{-1} \underline{P} (\underline{S} \underline{S}^{-1}) \underline{Q} \underline{S} \underline{S}^{-1} \underline{R} \underline{S}, \end{aligned}$$

or,

$$\underline{A}' + \underline{B}' \underline{C}' = \underline{P}' \underline{Q}' \underline{R}'.$$

### J. How to Construct $\underline{S}^{-1}$

To form  $\underline{S}^{-1}$ , recall the eigenvalue equation we began with

$$\underline{M} \mathbf{v}(k) = \lambda_k \mathbf{v}(k).$$

If we multiply on the left by  $\underline{S}^{-1}$  and insert  $\underline{S}^{-1} \underline{S} = \underline{I}$ , we obtain

$$\underline{S}^{-1} \underline{M} (\underline{S} \underline{S}^{-1}) \mathbf{v}(k) = \lambda_k \underline{S}^{-1} \mathbf{v}(k),$$

or,

$$\mathbf{v}(k)' = \lambda_k \mathbf{v}(k)'.$$

According to the passive picture of matrix operations,  $\mathbf{v}(k)' = \underline{S}^{-1} \mathbf{v}(k)$  is the old vector  $\mathbf{v}$  expressed in terms of new basis functions which are given by the columns of  $\underline{S}$ . But the rows of  $\underline{S}^{-1}$ , or equivalently, the columns of  $\underline{S}$  have to be orthonormal vectors (using the definition of an inverse):

$$\sum_j \underline{S}^{-1}_{ij} \underline{S}_{jk} = \delta_{ik}.$$

However, the columns of  $\underline{S}$  are just the eigenvectors  $\mathbf{v}(k)$ , so the rows of  $\underline{S}^{-1}$  must also have something to do with  $\mathbf{v}(k)$ .

Now consider the matrix  $\underline{S}^{-1}$ . Its elements are  $(\underline{S}^{-1})_{kj} = \underline{S}_{jk}^*$ , so  $(\underline{S}^{-1})_{kj} = \mathbf{v}_j^*(k)$ , as a result of which we can write

$$(\underline{S}^{-1})_{ik} = \sum_j S_{ij} S_{jk} = \sum_j \mathbf{v}_j^*(i) \mathbf{v}_j(k) = \langle \mathbf{v}(i) | \mathbf{v}(k) \rangle = \delta_{ik}.$$

We have therefore found  $\underline{S}^{-1}$ , and it is  $\underline{S}^{-1}$ . We have also proven the important theorem stating that any Hermitian matrix  $\underline{M}$ , can be diagonalized by a matrix  $\underline{S}$  which is unitary

( $\underline{S}^{-1} = \underline{S}^{-1}$ ) and whose columns are the eigenvectors of  $\underline{M}$ .

## VI. Finding Inverses, Square Roots, and Other Functions of a Matrix Using its Eigenvectors and Eigenvalues

Since a matrix is defined by how it affects basis vectors upon which it acts, and since a matrix only changes the lengths, not the directions of its eigenvectors, we can form other matrices that have the same eigenvectors as the original matrix but have different eigenvalues in a straightforward manner. For example, if we want to reverse the effect of  $\underline{M}$ , we create another matrix that reduces the lengths of the eigenvectors by the same amount that  $\underline{M}$  lengthened them. This produces the inverse matrix. We illustrate this by using the same example matrix that we used earlier.

$$\underline{M}^{-1} = \sum_k k^{-1} |\mathbf{v}(i)\rangle \langle \mathbf{v}(i)| \quad (\text{this is defined only if all } k \neq 0)$$

$$\begin{pmatrix} 0 & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{10}{24} & -\frac{2}{24} \\ 0 & -\frac{2}{24} & \frac{10}{24} \end{pmatrix} = \underline{M}^{-1}$$

To show that this matrix obeys  $\underline{M}^{-1} \underline{M} = \underline{1}$ , we simply multiply the two matrices together:

$$\begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{10}{24} & -\frac{2}{24} \\ 0 & -\frac{2}{24} & \frac{10}{24} \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & \frac{5}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{5}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

An extension of this result is that a whole family of matrices, each member related to the original  $\underline{M}$  matrix, can be formed by combining the eigenvalues and eigenvectors as follows:

$$f(\underline{M}) = \sum_i |\mathbf{v}(i)\rangle \langle \mathbf{v}(i)| f(\lambda_i),$$

where  $f$  is any function for which  $f(x)$  is defined. Examples include  $\exp(\underline{M})$ ,  $\sin(\underline{M})$ ,  $\underline{M}^{1/2}$ , and  $(\underline{I}-\underline{M})^{-1}$ . The matrices so constructed (e.g.,  $\exp(\underline{M}) = \sum_i \exp(\lambda_i) |\mathbf{v}(i)\rangle \langle \mathbf{v}(i)|$ ) are proper representations of the functions of  $\underline{M}$  in the sense that they give results equal to those of the function of  $\underline{M}$  when acting on any eigenvector of  $\underline{M}$ ; because the eigenvectors form complete sets, this implies that they give identical results when acting on any vector. This equivalence can be shown most easily by carrying out a power series expansion of the function of  $\underline{M}$  (e.g., of  $\exp(\underline{M})$ ) and allowing each term in the series to act on an eigenvector.

## VII. Projectors Revisited

In hindsight, the relationships developed above for expressing functions of the original matrix in terms of its eigenvectors and eigenvalues are not unexpected because each of the matrices is given in terms of the so-called projector matrices  $|\mathbf{v}(i)\rangle \langle \mathbf{v}(i)|$ . As we saw, the matrix  $\sum_i |\mathbf{v}(i)\rangle \langle \mathbf{v}(i)|$  behaves just like  $\underline{M}$ , as demonstrated below:

$$\sum_i |\mathbf{v}(i)\rangle \langle \mathbf{v}(i)| \sum_j |\mathbf{v}(j)\rangle = \sum_i |\mathbf{v}(i)\rangle \langle \mathbf{v}(i)| \sum_j |\mathbf{v}(j)\rangle$$

The matrix  $|\mathbf{v}(i)\rangle \langle \mathbf{v}(i)|$  is called a projector onto the space of eigenvector  $|\mathbf{v}(i)\rangle$  because its action on any vector  $|\mathbf{f}\rangle$  within the class of admissible vectors (2- or 4-dimensional vectors are in a different class than 3- or 1- or 7- or 96-dimensional vectors)

$$(|\mathbf{v}(i)\rangle \langle \mathbf{v}(i)|) |\mathbf{f}\rangle = |\mathbf{v}(i)\rangle (\langle \mathbf{v}(i)|\mathbf{f}\rangle)$$

gives  $|\mathbf{v}(i)\rangle$  multiplied by the coefficient of  $|\mathbf{f}\rangle$  along  $|\mathbf{v}(i)\rangle$ .

This construction in which a vector is used to form a matrix  $|\mathbf{v}(i)\rangle \langle \mathbf{v}(i)|$  is called an "outer product". The projection matrix thus formed can be shown to be idempotent, which means that the result of applying it twice (or more times) is identical to the result of applying it once  $P P = P$ . This property is straightforward to demonstrate. Let us consider the projector  $\underline{P}_i = |\mathbf{v}(i)\rangle \langle \mathbf{v}(i)|$ , which, when applied twice yields

$$|\mathbf{v}(i)\rangle \langle \mathbf{v}(i)| |\mathbf{v}(i)\rangle \langle \mathbf{v}(i)| = |\mathbf{v}(i)\rangle \langle \mathbf{v}(i)| \langle \mathbf{v}(i)|\mathbf{v}(i)\rangle = |\mathbf{v}(i)\rangle \langle \mathbf{v}(i)|$$

Sets of projector matrices each formed from a member of an orthonormal vector set are mutually orthogonal, (i.e.,  $\underline{P}_i \underline{P}_j = 0$  if  $i \neq j$ ), which can be shown as follows:

$$\begin{aligned} |\mathbf{v}(i)\rangle \langle \mathbf{v}(i)| |\mathbf{v}(j)\rangle \langle \mathbf{v}(j)| &= |\mathbf{v}(i)\rangle \langle \mathbf{v}(i)| \langle \mathbf{v}(j)|\mathbf{v}(j)\rangle \\ &= |\mathbf{v}(i)\rangle \langle \mathbf{v}(i)| \langle \mathbf{v}(j)|\mathbf{v}(j)\rangle = 0 \quad (i \neq j) \end{aligned}$$

## VIII. Hermitian Matrices and The Turnover Rule

The eigenvalue equation:

$$\underline{M}|\mathbf{v}(i)\rangle = \lambda_i|\mathbf{v}(i)\rangle,$$

which can be expressed in terms of its indices as:

$$M_{kj}v_j(i) = \lambda_i v_k(i),$$

is equivalent to (just take the complex conjugate):

$$v_j^*(i)M_{kj}^* = \lambda_i^* v_k^*(i),$$

which, for a Hermitian matrix  $\underline{M}$ , can be rewritten as:

$$v_j^*(i)M_{jk} = \lambda_i^* v_k^*(i),$$

or, equivalently, as:

$$\langle \mathbf{v}(i) | \underline{M} = \lambda_i \langle \mathbf{v}(i) |.$$

This means that the  $\mathbf{v}(i)$ , when viewed as column vectors, obey the eigenvalue identity  $\underline{M}|\mathbf{v}(i)\rangle = \lambda_i|\mathbf{v}(i)\rangle$ . These same vectors, when viewed as row vectors (and thus complex conjugated), also obey the eigenvalue relation, but in the "turn over" form  $\langle \mathbf{v}(i) | \underline{M} = \lambda_i \langle \mathbf{v}(i) |$ . For example, in the case we have been studying, the first vector obeys

$$\begin{pmatrix} 3 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{2} & \frac{5}{2} \end{pmatrix} = \begin{pmatrix} 0 & \frac{3}{\sqrt{2}} & \frac{3}{\sqrt{2}} \\ 0 & \frac{5}{2} & \frac{1}{2} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = 3 \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

As a general rule, a hermitian matrix operating on a column vector to the right is equivalent to the matrix operating on the complex conjugate of the same vector written as a row vector to the left. For a non-Hermitian matrix, it is the adjoint matrix operating on the complex conjugate row vector that is equivalent to the original matrix acting on the column vector on the right. These two statements are consistent since, for a Hermitian matrix the adjoint is identical to the original matrix (often, Hermitian matrices are called "self-adjoint"

matrices for this reason)  $\underline{M} = \underline{M}^\dagger$ .

## IX. Connection Between Orthonormal Vectors and Orthonormal Functions.

For vectors as we have been dealing with, the scalar or dot product is defined as we have seen as follows:

$$\langle \mathbf{v}(i) \mathbf{v}(j) \rangle = \sum_k v_k^*(i) v_k(j) .$$

For functions of one or more variable (we denote the variables collectively as  $x$ ), the generalization of the vector scalar product is

$$\langle \mathbf{f}(i) \mathbf{f}(j) \rangle = \int f_i^*(x) f_j(x) dx .$$

The range of integration and the meaning of the variables  $x$  will be defined by the specific problem of interest; for example, in polar coordinates for one particle,  $x = r, \theta$ , and for  $N$  particles in a general coordinate system  $x = \mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N$ .

If the functions  $f_i(x)$  are orthonormal;

$$\langle \mathbf{f}(i) \mathbf{f}(j) \rangle = \delta_{ij} ,$$

and complete, then a resolution of the identity holds that is similar to that for vectors

$$\int_i |\mathbf{f}(i)\rangle \langle \mathbf{f}(i)| = \delta(x-x') = \int_i f_i(x) f_i^*(x') dx$$

where  $\delta(x-x')$  is called the Dirac delta function. The function is zero everywhere except at  $x'$ , but has unit area under it. In other words,

$$\begin{aligned} \int_i \delta(x-x') g(x') dx' &= g(x) = \int_i |\mathbf{f}(i)\rangle \langle \mathbf{f}(i)| \mathbf{g}\rangle = \int_i f_i(x) \int_i f_i^*(x') g(x') dx' \\ &= \int_i f_i(x) a_i , \end{aligned}$$

where  $a_i$  is the projection of  $g(x)$  along  $f_i(x)$ . The first equality can be taken as a definition of  $\delta(x-x')$ .

## X. Matrix Representations of Functions and Operators

As we saw above, if the set of functions  $\{f_i(x)\}$  is complete, any function of  $x$  can be expanded in terms of the  $\{f_i(x)\}$

$$\begin{aligned}
g(x) &= \sum_i |f_i\rangle \langle f_i| g \rangle \\
&= \sum_i f_i(x) \int f_i^*(x') g(x') dx'
\end{aligned}$$

The column vector of numbers  $a_i = \int f_i^*(x') g(x') dx'$  is called the representation of  $g(x)$  in the  $f_i(x)$  basis. Note that this vector may have an infinite number of components because there may be an infinite number of  $f_i(x)$ .

Suppose that the function  $g(x)$  obeys some operator equation (e.g., an eigenvalue equation) such as

$$\frac{dg(x)}{dx} = g(x).$$

Then, if we express  $g(x)$  in terms of the  $\{f_i(x)\}$ , the equation for  $g(x)$  can be replaced by a corresponding matrix problem for the  $a_j$  vector coefficients:

$$\sum_i \frac{df_i(x)}{dx} \int f_i^*(x') g(x') dx' = \sum_i f_i(x) \int f_i^*(x') g(x') dx'$$

If we now multiply by  $f_j^*(x)$  and integrate over  $x$ , we obtain

$$\int f_j^*(x) \frac{d}{dx} f_i(x) dx a_i = \int f_j^*(x) f_i(x) dx a_i.$$

If the  $f_i(x)$  functions are orthonormal, this result simplifies to

$$\frac{d}{dx} a_i = a_j$$

where we have defined the matrix representation of  $\frac{d}{dx}$  in the  $\{f_i\}$  basis by

$$\frac{d}{dx} a_j = \sum_i \int f_i^*(x) \frac{d}{dx} f_i(x) dx a_i.$$

Remember that  $a_i$  is the representation of  $g(x)$  in the  $\{f_i\}$  basis. So the operator eigenvalue equation is equivalent to the matrix eigenvalue problem if the functions  $\{f_i\}$  form a complete set.

Let us consider an example, that of the derivative operator in the orthonormal basis of Harmonic Oscillator functions. The fact that the solutions of the quantum Harmonic Oscillator,  $\psi_n(x)$ , are orthonormal and complete means that:

$$\int_{-\infty}^{\infty} \psi_n^*(x) \psi_m(x) dx = \delta_{nm}.$$

The lowest members of this set are given as

$$\psi_0(x) = \frac{1}{\sqrt{4}} e^{-\frac{x^2}{2}},$$

$$\psi_1(x) = \frac{1}{\sqrt{4}} \frac{1}{\sqrt{2}} x e^{-\frac{x^2}{2}},$$

$$\psi_2(x) = \frac{1}{\sqrt{4}} \frac{1}{\sqrt{8}} (4x^2 - 2) e^{-\frac{x^2}{2}}, \dots,$$

$$\psi_n(x) = A_n H_n(x) e^{-\frac{x^2}{2}}.$$

The derivatives of these functions are

$$\psi_0(x)' = \frac{1}{\sqrt{4}} (-x) e^{-\frac{x^2}{2}},$$

$$\psi_1(x)' = \frac{1}{\sqrt{4}} \frac{1}{\sqrt{2}} e^{-\frac{x^2}{2}} - x^2 e^{-\frac{x^2}{2}},$$

,

$$= \frac{1}{\sqrt{2}} (0 - 2) e^{-\frac{x^2}{2}}, \text{ etc.}$$

In general, one finds that

$$\frac{d}{dx} \psi_n(x) = \frac{1}{\sqrt{2}} \frac{1}{\sqrt{n}} \psi_{n-1}(x) - (n+1) \frac{1}{\sqrt{n+1}} \psi_{n+1}(x).$$

From this general result, it is clear that the matrix representation of the  $\frac{d}{dx}$  operator is given by

$$\underline{D} = 2^{-\frac{1}{2}} \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ -1 & 0 & \sqrt{2} & 0 & \dots \\ 0 & -\sqrt{2} & 0 & \sqrt{3} & \dots \\ 0 & 0 & -\sqrt{3} & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} .$$

The matrix  $\underline{D}$  operates on the unit vectors  $\mathbf{e}_0 = (1,0,0\dots)$ ,  $\mathbf{e}_1 = (0,1,\dots)$  etc. just like  $\frac{d}{dx}$  operates on  $\psi_n(x)$ , because these unit vectors are the representations of the  $\psi_n(x)$  in the basis of the  $\psi_n(x)$ , and  $\underline{D}$  is the representation of  $\frac{d}{dx}$  in the basis of the  $\psi_n(x)$ . Since any vector can be represented by a linear combination of the  $\mathbf{e}_i$  vectors, the matrix  $\underline{D}$  operates on any vector  $(a_0, a_1, a_2\dots)$  just like  $\frac{d}{dx}$  operates on any  $f(x)$ . Note that the matrix is not Hermitian; it is actually antisymmetric. However, if we multiply  $\underline{D}$  by  $-i\hbar$  we obtain a Hermitian matrix that represents the operator  $-i\hbar \frac{d}{dx}$ , the momentum operator.

It is easy to see that we can form the matrix representation of any linear operator for any complete basis in any space. To do so, we act on each basis function with the operator and express the resulting function as a linear combination of the original basis functions. The coefficients that arise when we express the operator acting on the functions in terms of the original functions form the the matrix representation of the operator.

It is natural to ask what the eigenvalues and eigenfunctions of the matrix you form through this process mean. If your operator is the Hamiltonian operator, then the matrix eigenvectors and eigenvalues are the representations of the solutions to the Schrodinger equation in this basis. Forming the representation of an operator reduces the solution of the operator eigenvalue equation to a matrix eigenvalue equation.

## XI. Complex Numbers, Fourier Series, Fourier Transforms, Basis Sets

One of the techniques to which chemists are frequently exposed is Fourier transforms. They are used in NMR and IR spectroscopy, quantum mechanics, and classical mechanics.

Just as we expand a function in terms of a complete set of basis function or a vector in terms of a complete set of vectors, the Fourier transform expresses a function  $f(x)$  of a continuous variable  $x$  in terms of a set of orthonormal functions that are not discretely

labeled but which are labeled by a continuous "index"  $t$ . These functions are  $(2\pi)^{-\frac{1}{2}} e^{-i t x}$ , and the "coefficients" in the expansion

$$f(x) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{+\infty} e^{-i t x} f(t) dt ,$$

are called the Fourier transform  $f(t)$  of  $f(x)$ .

The orthonormality of the  $(\frac{1}{2})^{-1} e^{-i t}$  functions will be demonstrated explicitly later. Before doing so however, it is useful to review both complex numbers and basis sets.

### A. Complex numbers

A complex number has a real part and an imaginary part and is usually denoted:

$$z = x + iy.$$

For the complex number  $z$ ,  $x$  is the real part,  $y$  is the imaginary part, and  $i = \sqrt{-1}$ . This is expressed as  $x = \text{Re}(z)$ ,  $y = \text{Im}(z)$ . For every complex number  $z$ , there is a related one called its complex conjugate  $z^* = x - iy$ .

Complex numbers can be thought of as points in a plane where  $x$  and  $y$  are the abscissa and ordinate, respectively. This point of view prompts us to introduce polar coordinates  $r$  and  $\theta$  to describe complex numbers, with

$$x = r \cos \theta \quad r = \sqrt{x^2 + y^2}$$

or,

$$y = r \sin \theta \quad \theta = \tan^{-1} \frac{y}{x} + \begin{cases} 0 & (\text{if } x > 0) \\ \pi & (\text{if } x < 0) \end{cases}$$

Another name for  $r$  is the norm of  $z$  which is denoted  $|z|$ . The angle  $\theta$  is sometimes called the argument of  $z$ ,  $\arg(z)$ , or the phase of  $z$ .

Complex numbers can be added, subtracted, multiplied and divided like real numbers. For example, the multiplication of  $z$  by  $z^*$  gives:

$$zz^* = (x + iy)(x - iy) = x^2 + ixy - ixy + y^2 = x^2 + y^2 = r^2$$

Thus  $zz^* = |z|^2$  is a real number.

An identity due to Euler is important to know and is quite useful. It states that

$$e^{i\theta} = \cos \theta + i \sin \theta.$$

It can be proven by showing that both sides of the identity obey the same differential equation. Here we will only demonstrate its plausibility by Taylor series expanding both sides:

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \dots,$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots,$$

and

$$\cos x = 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \dots$$

Therefore, the exponential  $e^{ix}$  becomes

$$\begin{aligned} e^{ix} &= 1 + i x - \frac{x^2}{2} - i \frac{x^3}{3!} + \frac{x^4}{4!} + i \frac{x^5}{5!} - \dots \\ &= \left( 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \dots \right) + i \left( x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots \right) \end{aligned}$$

The odd powers of  $i$  clearly combine to give the Sine function; the even powers give the Cosine function, so

$$e^{ix} = \cos x + i \sin x$$

is established. From this identity, it follows that  $\cos x = \frac{1}{2}(e^{ix} + e^{-ix})$  and  $\sin x = \frac{1}{2i}(e^{ix} - e^{-ix})$ .

It is now possible to express the complex number  $z$  as

$$\begin{aligned} z &= x + iy \\ &= r \cos \theta + i r \sin \theta \\ &= r(\cos \theta + i \sin \theta) \\ &= r e^{i\theta} \end{aligned}$$

This form for complex numbers is extremely useful. For instance, it can be used to easily show that

$$z z^* = r e^{i\theta} (r e^{-i\theta}) = r^2 e^{i\theta - i\theta} = r^2.$$

## B. Fourier Series

Now let us consider a function that is periodic in time with period  $T$ . Fourier's theorem states that any periodic function can be expressed in a Fourier series as a linear combination (infinite series) of Sines and Cosines whose frequencies are multiples of a fundamental frequency corresponding to the period:

$$f(t) = \sum_{n=0}^{\infty} a_n \cos(n \omega t) + \sum_{n=1}^{\infty} b_n \sin(n \omega t),$$

where  $\omega = \frac{2\pi}{T}$ . The Fourier expansion coefficients are given by projecting  $f(t)$  along each of the orthogonal Sine and Cosine functions:

$$a_0 = \frac{1}{T} \int_0^T f(t) dt,$$

$$a_n = \frac{2}{T} \int_0^T f(t) \cos(n\omega t) dt,$$

$$b_n = \frac{2}{T} \int_0^T f(t) \sin(n\omega t) dt.$$

The term in the Fourier expansion associated with  $a_0$  is a constant giving the average value (i.e., the zero-frequency or DC component) of the function. The terms with  $n = 1$  contain the fundamental frequency and higher terms contain the  $n$ th harmonics or overtones of the fundamental frequency. The coefficients,  $a_n$  and  $b_n$ , give the amplitudes of each of these frequencies. Note that if  $f(t)$  is an even function (i.e., if  $f(t) = f(-t)$ ),  $b_n = 0$  for  $n = 1, 2, 3, \dots$  so the series only has Cosine terms. If  $f(t)$  is an odd function (i.e.  $f(t) = -f(-t)$ ),  $a_n = 0$  for  $n = 0, 1, 2, 3, \dots$  and the series only has Sine terms.

The Fourier series expresses a continuous function as an infinite series of numbers  $\dots a_0, a_1, b_1, a_2, b_2, \dots$ . We say that the set of coefficients is a representation of the function in the Fourier basis. The expansion in the  $\cos n\omega t$  and  $\sin n\omega t$  basis is useful because the basis functions are orthogonal when integrated over the interval 0 to  $T$ . An alternative set of functions is sometimes used to carry out the Fourier expansion; namely the  $\frac{1}{\sqrt{T}} \exp(i n\omega t)$  functions for  $n$  from  $-\infty$  to  $+\infty$ . Their orthogonality can be proven as follows:

$$\begin{aligned} \frac{1}{T} \int_0^T \exp(i n\omega t) \exp(-i m\omega t) dt &= \frac{1}{T} \int_0^T \exp(i(m-n)\omega t) dt = 1 \text{ if } m=n, \text{ and} \\ &= \frac{1}{T} (i(m-n)\omega)^{-1} (\exp(i(m-n)\omega T) - 1) = 0 \text{ if } m \neq n. \end{aligned}$$

Let us consider the Fourier representation of  $f(t)$  in terms of the complex exponentials introduced above. For an arbitrary periodic function  $f(t)$ , we can write

$$f(t) = \sum_{n=-\infty}^{+\infty} c_n e^{i n\omega t}, \text{ where } c_n = \frac{1}{T} \int_0^T f(t) e^{-i n\omega t} dt.$$

This form of the Fourier series is entirely equivalent to the first form and the  $a_n$  and  $b_n$  can be obtained from the  $c_n$  and vice versa. For example, the  $c_n$  amplitudes are obtained by projecting  $f(t)$  onto  $\exp(in\ t)$  as:

$$c_n = \frac{1}{T} \int_0^T f(t)(\cos(n\ t) - i\sin(n\ t)) dt$$

$$= \frac{1}{T} \int_0^T f(t)\cos(n\ t)dt - \frac{1}{T} \int_0^T f(t)\sin(n\ t)dt ,$$

but these two integrals are easily recognized as the expansion in the other basis, so

$$c_n = \frac{1}{2} a_n - \frac{1}{2} i b_n = \frac{1}{2} (a_n - ib_n) .$$

By using complex exponential functions instead of trigonometric functions, we only have one family of basis functions,  $e^{in\ t}$ , but we have twice as many of them. In this form, if  $f(t)$  is even, then the  $c_n$  are real, but if  $f(t)$  is odd, then  $c_n$  are imaginary.

It is useful to consider some examples to learn this material. First, let  $f(t)$  be the odd function  $f(t) = \sin 3t$ . Then, one period has elapsed when  $3T = 2\pi$ , so  $T = \frac{2\pi}{3}$  and  $\omega = 3$ . For this function, the complex series Fourier expansion coefficients are given by

$$c_0 = \frac{3}{2} \int_0^{\frac{2\pi}{3}} \sin 3t dt = 0.$$

$$c_1 = \frac{3}{2} \int_0^{\frac{2\pi}{3}} \sin 3t e^{-i3t} dt = \frac{3}{2} \frac{(e^{i3t} - e^{-i3t})}{2i} e^{-i3t} dt$$

$$= \frac{1}{2i} \frac{3}{2} \int_0^{\frac{2\pi}{3}} (1 - e^{6it}) dt = \frac{3}{4i} \left[ \frac{2\pi}{3} - 0 \right] = \frac{1}{2i} = -\frac{i}{2}$$

Because  $\sin 3t$  is real, it is straightforward to see from the definition of  $c_n$  that  $c_{-n} = c_n^*$ , as a result of which  $c_{-1} = c_1^* = +\frac{i}{2}$ . The orthogonality of the  $\exp(in\ t)$  functions can be used to show that all of the higher  $c_n$  coefficients vanish:

$$c_n = 0 \text{ for } n \neq \pm 1.$$

Hence we see that this simple periodic function has just two terms in its Fourier series. In terms of the Sine and Cosine expansion, one finds for this same  $f(t) = \sin 3t$  that  $a_n = 0$ ,  $b_n = 0$  for  $n \neq 1$ , and  $b_1 = 1$ .

As another example, let  $f(t)$  be

$$f(t) = t, \quad -\pi < t < \pi,$$

and make  $f(t)$  periodic with period  $2\pi$  (so  $\omega = 1$ ). For this function,

$$c_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} t dt = 0$$

$$\begin{aligned} c_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} t e^{-int} dt = \frac{1}{2\pi n^2} [e^{-int}(1 + int)] \Big|_{-\pi}^{\pi} \\ &= \frac{1}{2\pi n^2} (e^{-in\pi} (1 + in\pi) - e^{in\pi} (1 - in\pi)) \\ &= \frac{1}{2\pi n^2} ((-1)^n (1 + in\pi) - (-1)^n (1 - in\pi)) \\ &= \frac{(-1)^n}{2\pi n^2} (2in\pi) = \frac{i}{n} (-1)^n \pi^{-1} \end{aligned}$$

Note that since  $t$  is an odd function,  $c_n$  is imaginary, and as  $n \rightarrow \infty$ ,  $c_n$  approaches zero.

### C. Fourier Transforms

Let us now imagine that the period  $T$  becomes very long. This means that the fundamental frequency  $\omega = \frac{2\pi}{T}$  becomes very small and the harmonic frequencies  $n\omega$  are very close together. In the limit of an infinite period one has a non-periodic function (i.e., an arbitrary function). The frequencies needed to represent such a function become continuous, and hence the Fourier sum becomes an integral.

To make this generalization more concrete, we first take the earlier Fourier expression for  $f(t)$  and multiply it by unity in the form  $\frac{T}{2\pi}$  to obtain:

$$f(t) = \sum c_n e^{in\omega t} \frac{T}{2\pi}$$

Now we replace  $n$  (which are frequencies infinitesimally close together for sufficiently long  $T$ ) by the continuous index  $\omega$ :

$$f(t) = \frac{1}{2} \sum_{n=-\infty}^{\infty} (c_n T) e^{i n \omega t} .$$

In the limit of long  $T$ , this sum becomes an integral because  $\omega$  becomes infinitesimally small ( $\omega \rightarrow d\omega$ ). As  $T$  grows, the frequency spacing between  $n$  and  $n+1$ , which is  $\frac{2\pi}{T}$ , as well as the frequency associated with a given  $n$ -value become smaller and smaller. As a result, the product  $c_n T$  evolves into a continuous function of  $\omega$  which we denote  $c(\omega)$ . Before,  $c_n$  depended on the continuous index  $n$ , and represented the contribution to the function  $f(t)$  of waves with frequency  $n$ . The function  $c(\omega)$  is the contribution per unit frequency to  $f(t)$  from waves with frequency in the range  $\omega$  to  $\omega + d\omega$ . In summary, the Fourier series expansion evolves into an integral that defines the Fourier transformation of  $f(t)$ :

$$f(t) = \frac{1}{2} \int_{-\infty}^{\infty} c(\omega) e^{-i \omega t} d\omega .$$

It is convenient to define a new function  $f(\omega) = (2\pi)^{-\frac{1}{2}} c(\omega)$  where  $f(\omega)$  is called the Fourier transform of  $f(t)$ . The Fourier transform  $f(\omega)$  is a representation of  $f(t)$  in another basis, that of the orthonormal set of oscillating functions  $e^{i \omega t} (2\pi)^{-\frac{1}{2}}$ :

$$f(t) = \frac{1}{2} \int_{-\infty}^{\infty} f(\omega) e^{-i \omega t} d\omega .$$

Earlier, for Fourier series, we had the orthogonality relation among the Fourier functions:

$$\frac{1}{T} \int_0^T n^* m dt = \delta_{nm} ,$$

but for the continuous variable  $\omega$ , we have a different kind of orthogonality

$$\int_{-\infty}^{\infty} f_1^*(\omega) f_2(\omega) d\omega = \delta(\omega_1 - \omega_2) ,$$

where  $\delta(t - t_0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega(t - t_0)} d\omega$ .

The function  $\delta(t - t_0)$ , called the Dirac delta function, is the continuous analog to  $\delta_{nm}$ . It is zero unless  $t = t_0$ . If  $t = t_0$ ,  $\delta(t - t_0)$  is infinite, but it is infinite in such a way that the area under the curve is precisely unity. Its most useful definition is that  $\delta(t - t_0)$  is the function which, for arbitrary  $f(t)$ , the following identity holds:

$$\int_{-\infty}^{\infty} f(t) \delta(t - t_0) dt = f(t_0).$$

That is, integrating  $\delta(t - t_0)$  times any function evaluated at  $t$  just gives back the value of the function at  $t_0$ .

The Dirac delta function can be expressed in a number of convenient ways, including the following two:

$$\begin{aligned} \delta(t - t_0) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega(t - t_0)} d\omega \\ &= \lim_{a \rightarrow 0} \int_0^a \sqrt{\frac{1}{a}} e^{-\frac{(t - t_0)^2}{a^2}} dt \end{aligned}$$

As an example of applying the Fourier transform method to a non-periodic function, consider the localized pulse

$$f(t) = \begin{cases} \frac{1}{T} & -\frac{T}{2} \leq t \leq \frac{T}{2} \\ 0 & t > \frac{T}{2} \end{cases}$$

For this function, the Fourier transform is

$$\begin{aligned} f(\omega) &= \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \\ &= \int_{-\frac{T}{2}}^{\frac{T}{2}} \frac{1}{T} e^{-i\omega t} dt \end{aligned}$$

$$\begin{aligned}
&= (2\pi)^{-\frac{1}{2}} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} (\cos t - i \sin t) dt \\
&= (2\pi)^{-\frac{1}{2}} \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} \cos t dt + 0 \\
&= (2\pi)^{-\frac{1}{2}} \frac{1}{T} (\sin t) \Big|_{-\frac{T}{2}}^{\frac{T}{2}} \\
&= \frac{2}{T}^{-\frac{1}{2}} \sin(\pi T/2) / T.
\end{aligned}$$

Note that  $f(\omega)$  has its maximum value of  $(2\pi)^{-\frac{1}{2}}$  for  $\omega = 0$  and that  $f(\omega)$  falls slowly in magnitude to zero as  $\omega$  increases in magnitude, while oscillating to positive and negative values. However, the primary maximum in  $f(\omega)$  near zero-frequency has a width that is inversely proportional to  $T$ . This inverse relationship between the width ( $T$ ) in  $t$ -space and the width  $\frac{2}{T}$  in  $\omega$ -space is an example of the uncertainty principle. In the present case it means that if you try to localize a wave in time, you must use a wide range of frequency components.

#### D. Fourier Series and Transforms in Space

The formalism we have developed for functions of time can also be used for functions of space variables or any other variables for that matter. If  $f(x)$  is a periodic function of the coordinate  $x$ , with period (repeat distance)  $\frac{2}{K}$ , then it can be represented in terms of Fourier functions as follows:

$$f(x) = \sum_{n=-\infty}^{\infty} f_n e^{inKx}$$

where the coefficients are

$$f_n = \frac{K}{2} \int_0^{\frac{2}{K}} f(x) e^{-inKx} dx .$$

If  $f(x)$  is a non-periodic function of the coordinate  $x$ , we write it as

$$f(x) = \int_{-\infty}^{\infty} f(k) e^{ikx} dk$$

and the Fourier transform is given as

$$f(k) = \int_{-\infty}^{\infty} f(x) e^{-ikx} dx$$

If  $f$  is a function of several spatial coordinates and/or time, one can Fourier transform (or express as Fourier series) simultaneously in as many variables as one wishes. You can even Fourier transform in some variables, expand in Fourier series in others, and not transform in another set of variables. It all depends on whether the functions are periodic or not, and whether you can solve the problem more easily after you have transformed it.

#### E. Comments

So far we have seen that a periodic function can be expanded in a discrete basis set of frequencies and a non-periodic function can be expanded in a continuous basis set of frequencies. The expansion process can be viewed as expressing a function in a different basis. These basis sets are the collections of solutions to a differential equation called the wave equation. These sets of solutions are useful because they are complete sets. Completeness means that any arbitrary function can be expressed exactly as a linear combination of these functions. Mathematically, completeness can be expressed as

$$1 = \int_{-\infty}^{\infty} \delta(x) dx$$

in the Fourier transform case, and

$$1 = \sum_n \delta_{n,n}$$

in the Fourier series case.

The only limitation on the function expressed is that it has to be a function that has the same boundary properties and depends on the same variables as the basis. You would not want to use Fourier series to express a function that is not periodic, nor would you want to express a three-dimensional vector using a two-dimensional or four-dimensional basis.

Besides the intrinsic usefulness of Fourier series and Fourier transforms for chemists (e.g., in FTIR spectroscopy), we have developed these ideas to illustrate a point that is important in quantum chemistry. Much of quantum chemistry is involved with basis sets and expansions. This has nothing in particular to do with quantum mechanics. Any time one is dealing with linear differential equations like those that govern light (e.g. spectroscopy) or matter (e.g. molecules), the solution can be written as linear combinations of complete sets of solutions.

## XII. Spherical Coordinates

### A. Definitions

The relationships among cartesian and spherical polar coordinates are given as follows:

$$z = r \cos \theta \quad r = \sqrt{x^2 + y^2 + z^2}$$

$$x = r \sin \theta \cos \phi \quad \theta = \cos^{-1} \frac{z}{\sqrt{x^2 + y^2 + z^2}}$$

$$y = r \sin \theta \sin \phi \quad \phi = \cos^{-1} \frac{x}{\sqrt{x^2 + y^2}}$$

The ranges of the polar variables are  $0 < r < \infty$ ,  $0 < \theta < \pi$ ,  $0 < \phi < 2\pi$ .

### B. The Jacobian in Integrals

In performing integration over all space, it is necessary to convert the multiple integral from cartesian to spherical coordinates:

$$dx \, dy \, dz \, f(x,y,z) = dr \, d\theta \, d\phi \, f(r, \theta, \phi) \, J,$$

where  $J$  is the so-called Jacobian of the transformation.  $J$  is computed by forming the determinant of the three-by-three matrix consisting of the partial derivatives relating  $x, y, z$  to  $r, \theta, \phi$ :

$$J = \frac{\partial(x,y,z)}{\partial(r, \theta, \phi)}$$

$$= \begin{vmatrix} \sin \theta \cos \phi & \sin \theta \sin \phi & \cos \theta \\ r \cos \theta \cos \phi & r \cos \theta \sin \phi & -r \sin \theta \\ -r \sin \theta \cos \phi & r \sin \theta \sin \phi & 0 \end{vmatrix}$$

The determinant,  $J$ , can be expanded, for example, by the method of diagonals, giving four nonzero terms:

$$J = r^2 \sin^3 \theta \cos^2 \phi + r^2 \cos^2 \theta \sin \theta \cos^2 \phi$$

$$+ r^2 \sin^3 \theta \sin^2 \phi + r^2 \cos^2 \theta \sin \theta \sin^2 \phi$$

$$= r^2 \sin \theta (\sin^2 \theta (\sin^2 \phi + \cos^2 \phi) + \cos^2 \theta (\cos^2 \phi + \sin^2 \phi))$$

$$= r^2 \sin \theta \, dr \, d\theta \, d\phi .$$

Hence in converting integrals from  $x, y, z$  to  $r, \theta, \phi$ , one writes as a short hand  $dx dy dz = r^2 \sin \theta \, dr \, d\theta \, d\phi$  .

### C. Transforming Operators

In many applications, derivative operators need to be expressed in spherical coordinates. In converting from cartesian to spherical coordinate derivatives, the chain rule is employed as follows:

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{\partial r}{\partial x} \frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x} \frac{\partial}{\partial \theta} + \frac{\partial \phi}{\partial x} \frac{\partial}{\partial \phi} \\ &= \sin \theta \cos \phi \frac{\partial}{\partial r} + \frac{\cos \theta \cos \phi}{r} \frac{\partial}{\partial \theta} - \frac{\sin \phi}{r \sin \theta} \frac{\partial}{\partial \phi} . \end{aligned}$$

Likewise

$$\frac{\partial}{\partial y} = \sin \theta \sin \phi \frac{\partial}{\partial r} + \frac{\cos \theta \sin \phi}{r} \frac{\partial}{\partial \theta} + \frac{\cos \phi}{r \sin \theta} \frac{\partial}{\partial \phi} ,$$

and

$$\frac{\partial}{\partial z} = \cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} + 0 \frac{\partial}{\partial \phi}$$

Now to obtain an expression for  $\frac{\partial^2}{\partial x^2}$  and other second derivatives, one needs to take the following derivatives:

$$\begin{aligned} \frac{\partial^2}{\partial x^2} &= \sin \theta \cos \phi \frac{\partial}{\partial r} \sin \theta \cos \phi \frac{\partial}{\partial r} + \frac{\cos \theta \cos \phi}{r} \frac{\partial}{\partial \theta} \frac{\partial}{\partial \theta} - \frac{\sin \phi}{r \sin \theta} \frac{\partial}{\partial \phi} \frac{\partial}{\partial \phi} \\ &\quad + \frac{\cos \theta \cos \phi}{r} \frac{\partial}{\partial r} \sin \theta \cos \phi \frac{\partial}{\partial r} + \frac{\cos \theta \cos \phi}{r} \frac{\partial}{\partial r} \frac{\partial}{\partial \theta} - \frac{\sin \phi}{r \sin \theta} \frac{\partial}{\partial r} \frac{\partial}{\partial \phi} \\ &\quad - \frac{\sin \phi}{r \sin \theta} \frac{\partial}{\partial \theta} \sin \theta \cos \phi \frac{\partial}{\partial r} + \frac{\cos \theta \cos \phi}{r} \frac{\partial}{\partial \theta} \frac{\partial}{\partial \theta} - \frac{\sin \phi}{r \sin \theta} \frac{\partial}{\partial \phi} \frac{\partial}{\partial \phi} \\ &= \sin^2 \theta \cos^2 \phi \frac{\partial^2}{\partial r^2} + \frac{\cos^2 \theta \cos^2 \phi}{r^2} \frac{\partial^2}{\partial \theta^2} + \frac{\sin^2 \phi}{r^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \end{aligned}$$

$$\begin{aligned}
& + \frac{2\sin\theta \cos\theta \cos^2\theta}{r} \frac{\partial^2}{\partial r^2} - \frac{2\cos\theta \sin\theta}{r} \frac{\partial^2}{\partial r \partial \theta} - \frac{2\cos\theta \cos\theta \sin\theta}{r^2 \sin\theta} \frac{\partial^2}{\partial \theta^2} \\
& + \frac{\cos^2\theta \cos^2\theta}{r} + \frac{\sin^2\theta}{r} \frac{\partial^2}{\partial r^2} + \frac{2\sin\theta \cos\theta \cos^2\theta}{r^2} + \frac{\sin^2\theta \cos\theta}{r^2 \sin\theta} \frac{\partial}{\partial \theta} \\
& + \frac{\cos\theta \sin\theta}{r^2} + \frac{\cos^2\theta \cos\theta \sin\theta}{r^2 \sin^2\theta} \frac{\partial}{\partial \theta}
\end{aligned}$$

Analogous steps can be performed for  $\frac{\partial^2}{\partial y^2}$  and  $\frac{\partial^2}{\partial z^2}$ . Adding up the three contributions, one obtains:

$$\Delta = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial}{\partial r} \right) + \frac{1}{r^2 \sin\theta} \frac{\partial}{\partial \theta} \left( \sin\theta \frac{\partial}{\partial \theta} \right) + \frac{1}{r^2 \sin^2\theta} \frac{\partial^2}{\partial \phi^2}.$$

As can be seen by reading Appendix G, the terms involving angular derivatives in  $\Delta$  is identical to  $-\frac{L^2}{r^2}$ , where  $L^2$  is the square of the rotational angular momentum operator.

Although in this appendix, we choose to treat it as a collection of differential operators that gives rise to differential equations in  $r, \theta, \phi$  to be solved, there are more general tools for treating all such angular momentum operators. These tools are developed in detail in Appendix G and are used substantially in the next two chapters.

### XIII. Separation of Variables

In solving differential equations such as the Schrödinger equation involving two or more variables (e.g., equations that depend on three spatial coordinates  $x, y,$  and  $z$  or  $r, \theta, \phi$ , and  $t$  or that depend on time  $t$  and several spatial variables denoted  $\mathbf{r}$ ), it is sometimes possible to reduce the solution of this one multi-variable equation to the solution of several equations each depending on fewer variables. A commonly used device for achieving this goal is the separation of variables technique.

This technique is not always successful, but can be used for the type of cases illustrated now. Consider a two-dimensional differential equation that is of second order and of the eigenvalue type:

$$A \frac{\partial^2}{\partial x^2} + B \frac{\partial^2}{\partial y^2} + C \frac{\partial}{\partial x} + D \frac{\partial}{\partial y} + E \psi = E \psi.$$

The solution  $\psi$  must be a function of  $x$  and  $y$  because the differential equation refers to  $\psi$ 's derivatives with respect to these variables.

The separations of variables device assumes that  $\psi(x,y)$  can be written as a product of a function of  $x$  and a function of  $y$ :

$$(x,y) = (x) (y).$$

Inserting this ansatz into the above differential equation and then dividing by  $(x) (y)$  produces:

$$A^{-1} \frac{2}{x^2} + B^{-1} \frac{2}{y^2} + C^{-1} \frac{1}{x} \frac{1}{y} + D = E.$$

The key observations to be made are:

A. If A is independent of y,  $A^{-1} \frac{2}{x^2}$  must be independent of y.

B. If B is independent of x,  $B^{-1} \frac{2}{y^2}$  must be independent of x.

C. If C vanishes and D does not depend on both x and y, then there are no "cross terms" in the above equations (i.e., terms that contain both x and y). For the sake of argument, let us take D to be dependent on x only for the remainder of this discussion; the case for which D depends on y only is handled in like manner.

Under circumstances for which all three of the above conditions are true, the left-hand side of the above second-order equation in two variables can be written as the sum of

$$A^{-1} \frac{2}{x^2} + D,$$

which is independent of y, and

$$B^{-1} \frac{2}{y^2}$$

which is independent of x.

The full equation states that the sum of these two pieces must equal the eigenvalue E, which is independent of both x and y.

Because E and  $B^{-1} \frac{2}{y^2}$  are both independent of x, the quantity  $A^{-1} \frac{2}{x^2} + D$  (as a whole) can not depend on x, since  $E - B^{-1} \frac{2}{y^2}$  must equal  $A^{-1} \frac{2}{x^2} + D$  for all values of x and y. That  $A^{-1} \frac{2}{x^2} + D$  is independent of x and (by assumption) independent of y allows us to write:

$$A^{-1} \frac{2}{x^2} + D = \text{constant},$$

a constant.

Likewise, because E and  $A^{-1} \frac{2}{x^2} + D$  are independent of y, the quantity  $B^{-1} \frac{2}{y^2}$  (as a whole) can not depend on y, since  $E - A^{-1} \frac{2}{x^2} - D$  must equal  $B^{-1} \frac{2}{y^2}$  for all values of x and y. That  $B^{-1} \frac{2}{y^2}$  is independent of y and (by assumption) independent of x allows us to write:

$$B^{-1} \frac{2}{y^2} = \text{constant},$$

another constant.

The net result is that we now have two first-order differential equations of the eigenvalue form:

$$A \frac{d^2}{dx^2} + D = \quad ,$$

and

$$B \frac{d^2}{dy^2} = \quad ' ,$$

and the solution of the original equation has been successfully subjected to separation of variables. The two eigenvalues  $\lambda$  and  $\mu$  of the separated x- and y- equations must obey:

$$\lambda + \mu = E,$$

which follows by using the two separate (x and y) eigenvalue equations in the full two-dimensional equation that contains E as its eigenvalue.

In summary, when separations of variables can be used, it:

- A. Reduces one multidimensional differential equation to two or more lower-dimensional differential equations.
- B. Expresses the eigenvalue of the original equation as a sum of eigenvalues (whose values are determined via boundary conditions as usual) of the lower-dimensional problems.
- C. Expresses the solutions to the original equation as a product of solutions to the lower-dimensional equations (i.e.,  $\psi(x,y) = \psi(x) \psi(y)$ , for the example considered above).

## The Hydrogen Atom Orbitals B

In Chapter 1 and Appendix A, the angular and radial parts of the Schrödinger equation for an electron moving in the potential of a nucleus of charge  $Z$  were obtained. These "hydrogen-like" atomic orbitals are proper eigenstates for H, He<sup>+</sup>, Li<sup>++</sup>, Be<sup>+++</sup>, etc. They also serve as useful starting points for treating many-electron atoms and molecules in that they have been found to be good basis functions in terms of which to expand the orbitals of most atoms and ions. In this Appendix, we will examine the sizes, energies, and shapes of these orbitals and consider, in particular, how these properties vary with principal quantum number  $n$ , charge  $Z$ , and angular quantum numbers  $l$  and  $m$ .

In Chapter 1 and Appendix A, it was shown that the total  $r, \theta, \phi$  dependence of these so-called hydrogenic orbitals is given as:

$$\psi_{n,l,m} = Y_{l,m}(\theta, \phi) R_{n,l}(r),$$

where the spherical harmonics  $Y_{l,m}$  are expressed in terms of the associated Legendre polynomials  $P_l^m(\cos \theta)$  as

$$Y_{l,m}(\theta, \phi) = P_l^{|m|}(\cos \theta) (2\pi)^{-1/2} \exp(im\phi),$$

and the radial functions  $R_{n,l}(r)$  are given in terms of Laguerre polynomials of order  $n-l-1$ ,  $L_{n-l-1}(\rho)$  as follows:

$$R_{n,l}(r) = N_{n,l}^{-1} e^{-\rho/2} L_{n-l-1}(\rho),$$

where  $N_{n,l}$  is a normalization constant. Here, the radial coordinate  $r$  of the electron is contained in the variable  $\rho$  defined as

$$\rho = Zr/a_0n,$$

where  $a_0$  is the Bohr radius

$$a_0 = \hbar^2 / \mu e^2 = 0.529 \text{ \AA}.$$

The energies of these hydrogenic orbitals, relative to an electron infinitely far from the nucleus and with zero kinetic energy, are

$$E = -\mu Z^2 e^4 / 2\hbar^2 n^2 = -27.21 Z^2 / 2n^2 \text{ eV}.$$

The  $n = 1, 2,$  and  $3$  wavefunctions are given explicitly as follows:

$$\begin{aligned}
 n=1, l=0, m=0, &= (Z/a_0)^{3/2} ( )^{-1/2} \exp(-Zr/a_0) \\
 n=2, l=0, m=0, &= (Z/a_0)^{3/2} (2 )^{-1/2} (1-Zr/2a_0)/2 \exp(-Zr/2a_0) \\
 n=2, l=1, m=0, &= (Z/a_0)^{3/2} (2 )^{-1/2} Zr/4a_0 \cos \exp(-Zr/2a_0) \\
 n=2, l=1, m=\pm 1, &= (Z/a_0)^{3/2} (4 )^{-1/2} Zr/4a_0 \sin \exp(\pm i ) \exp(-Zr/2a_0) \\
 n=3, l=0, m=0, &= (Z/a_0)^{3/2} (3 )^{-1/2} [27-18Zr/a_0+2(Zr/a_0)^2]/81 \\
 &\exp(-Zr/3a_0) \\
 n=3, l=1, m=0, &= (Z/a_0)^{3/2} (2 )^{-1/2} 2/81 [6Zr/a_0-(Zr/a_0)^2] \cos \\
 &\exp(-Zr/3a_0) \\
 n=3, l=1, m=\pm 1, &= (Z/a_0)^{3/2} ( )^{-1/2} 1/81 [6Zr/a_0-(Zr/a_0)^2] \sin \exp(\pm i ) \exp(-Zr/3a_0) \\
 n=3, l=2, m=0, &= (Z/a_0)^{3/2} (6 )^{-1/2} 1/81 (Zr/a_0)^2 [3\cos^2 -1] \\
 &\exp(-Zr/3a_0) \\
 n=3, l=2, m=\pm 1, &= (Z/a_0)^{3/2} ( )^{-1/2} 1/81 (Zr/a_0)^2 \sin \cos \exp(\pm i ) \exp(-Zr/3a_0) \\
 n=3, l=2, m=\pm 2, &= (Z/a_0)^{3/2} (4 )^{-1/2} 1/81 (Zr/a_0)^2 \sin^2 \exp(\pm 2i ) \exp(-Zr/3a_0).
 \end{aligned}$$

On pages 133-136 of Pauling and Wilson are tabulated the bound-state solutions of this problem for  $n$  ranging up to  $n=6$  and  $l$ -values up to  $l=5$ .

The above  $\pm m$  functions are appropriate whenever one wishes to describe orbitals that are eigenfunctions of  $L_z$ , the component of orbital angular momentum along the  $z$ -axis. The functions with  $\pm m$  quantum numbers can be combined in pairs to give the spatially oriented functions with which most chemists are familiar:

$$+ = 2^{-1/2} [ \psi_{m+} + \psi_{-m} ]$$

and

$$- = 2^{-1/2} [ \psi_{m-} - \psi_{-m} ].$$

For example, when applied to the  $2p_{\pm 1}$  functions, one forms the  $2p_x$  and  $2p_y$  functions; the  $3d_{\pm 1}$  functions combine to give the  $3d_{xz}$  and  $3d_{yz}$  functions; and the  $3d_{\pm 2}$  functions combine to give the  $3d_{xy}$  and  $3d_{x^2-y^2}$  functions. These spatially directed functions are more appropriate for use whenever one is dealing with the interaction of the atom on which these orbitals are located with surrounding "ligands" or bonding partners which are spatially arranged in a manner that destroys the spherical and axial symmetry of the system. One is permitted to combine the  $\pm m$  functions in this manner because these pairs of functions are energetically degenerate; any combination of them is therefore of the same energy.

There are several important points to stress about these hydrogenic atomic functions:

1. Their energies vary as  $-Z^2/n^2$ , so atoms with higher  $Z$  values will have more tightly bound electrons (for the same  $n$ ). In a many-electron atom, one often introduces the concept of an effective nuclear charge  $Z_{\text{eff}}$ , and takes this to be the full nuclear charge  $Z$  minus the number of electrons that occupy orbitals that reside radially "inside" the orbital in question. For example,  $Z_{\text{eff}} = 6-2=4$  for the  $n=2$  orbitals of Carbon in the  $1s^2 2s^2 2p^4$

configuration. This  $Z_{\text{eff}}$  is then used to qualitatively estimate the relative energetic stability by making use of the  $-Z_{\text{eff}}^2/n^2$  scaling of the energy.

2. The radial sizes of orbitals are governed by the product of  $\exp(-Z_{\text{eff}} r/n a_0)$ , an  $r^l$  factor which arises from the small- $r$  limit of  $\psi_{n,l,m}$ ,  $r^l$  which enters because the probability density is  $|\psi|^2 r^2$ , and the highest power of  $Zr/n a_0$  that appears in the Laguerre polynomial which is  $(Zr/n a_0)^{n-l-1}$ . This product's  $r$ -dependence reduces to  $\exp(-Z_{\text{eff}} r/n a_0) r^n$ , which has a maximum at  $r = n^2 a_0/Z_{\text{eff}}$ . So, orbitals with large  $Z_{\text{eff}}$  values are smaller and orbitals with larger  $n$ -values are larger.

3. The hydrogenic atom energy expression has no  $l$ -dependence; the 2s and 2p orbitals have exactly the same energy, as do the 3s, 3p, and 3d orbitals. This degree of degeneracy is only present in one-electron atoms and is the result of an additional symmetry (i.e., an additional operator that commutes with the Hamiltonian) that is not present once the atom contains two or more electrons. This additional symmetry is discussed on p. 77 of Atkins.

4. The radial part of  $\psi_{n,l,m}$  involves a polynomial in  $r$  of order  $n-l-1$

$$L_{n-l-1}(r) = \sum_k a_k r^k,$$
and the coefficients in these polynomials alternate in sign ( $a_k = c_k (-1)^k$ ). As a result, the radial functions possess nodes (i.e., values of  $r$  at which  $\psi_{n,l,m}$  vanishes) at points where the polynomial is equal to zero. There are  $n-l-1$  such radial nodes (excluding the nodes at  $r = \infty$  and at  $r = 0$ , the latter of which all but s-orbitals have). A 4s orbital has 3 such radial nodes, and a 3d orbital has none. Radial nodes provide a means by which an orbital acquires density closer to the nucleus than its  $\langle r^2 \rangle$  value would indicate; such enhanced density near the nucleus differentially stabilizes such orbitals relative to those with the same  $n$  but fewer nodes. In the language commonly used in inorganic chemistry, one says that nodes allow an orbital to "penetrate" the underlying inner-shell orbitals.

The presence of radial nodes also indicates that the electron has radial kinetic energy. The 3s orbital with 2 radial nodes has more radial kinetic energy than does the 3p which, in turn, has more than the 3d. On the other hand, the 3d orbital has the most angular energy

$l(l+1)\hbar^2/2mr^2$  (this is the analog of the rotational energy of a diatomic molecule), the 3p has an intermediate amount, and the 3s has the least.

## Quantum Mechanical Operators and Commutation C

### I. Bra-Ket Notation

It is conventional to represent integrals that occur in quantum mechanics in a notation that is independent of the number of coordinates involved. This is done because the fundamental structure of quantum chemistry applies to all atoms and molecules, regardless of how many electronic and atom-center coordinates arise. The most commonly used notation, which is referred to as 'Dirac' or 'bra-ket' notation, can be summarized as follows:

A. The wavefunction itself is represented as a so-called 'ket'  $|\psi\rangle$ .

B. The complex conjugate  $\psi^*$  of  $\psi$  is represented as a 'bra'  $\langle\psi|$ ; the complex conjugation is implied by writing  $\langle\psi|$ .

C. The integral, over all of the N coordinates ( $q_1 \dots q_N$ ) on which  $\psi$  depends, of the product of  $\psi^*$  and  $\psi$  is represented as a so-called 'bra-ket' or bracket:

$$\int \psi^* \psi dq_1 \dots dq_N = \langle\psi|\psi\rangle.$$

By convention, the two vertical lines that touch when  $\langle\psi|$  is placed against  $|\psi\rangle$  are merged into a single line in this notation.

D. Integrals involving one function ( $\psi^*$ ) and either another function ( $\psi$ ) or the result of an operator A acting on a function (e.g.,  $A\psi$  or  $A^2\psi$ ) are denoted as follows:

$$* \int dq_1 \dots dq_N = \langle \quad | \quad \rangle$$

$$* \int A \, dq_1 \dots dq_N = \langle \quad | A \quad \rangle = \langle \quad | A | \quad \rangle$$

$$* \int A \, dq_1 \dots dq_N = \langle \quad | A \quad \rangle = \langle \quad | A | \quad \rangle$$

$$(A)^* \int dq_1 \dots dq_N = \langle A \quad | \quad \rangle.$$

It is merely convention that an 'extra' vertical line (e.g., that appearing in  $\langle \quad | A \quad \rangle$ ) is inserted when an operator acting on the ket function appears in the integral.

## II. Hermitian Operators

In quantum mechanics, physically measurable quantities are represented by hermitian operators. Such operators  $\{R\}$  have matrix representations, in any basis spanning the space of functions on which the  $\{R\}$  act, that are hermitian:

$$\langle k | R | l \rangle = \langle l | R | k \rangle^* = \langle R | k | l \rangle.$$

The equality of the first and third terms expresses the so-called 'turn-over rule'; hermitian operators can act on the function to their right or, equivalently, on the function to their left.

Operators that do not obey the above identity are not hermitian. For such operators, it is useful to introduce the so-called adjoint operator as follows. If for the operator  $R$ , another operator  $R^+$  can be found that obeys

$$\langle k | R | l \rangle = \langle R^+ | k | l \rangle = \langle l | R^+ | k \rangle^*,$$

for all  $\{ | k \rangle \}$  within the class of functions on which  $R$  and  $R^+$  operate, then  $R^+$  is defined to be the adjoint of  $R$ . With this definition, it should be clear that hermitian operators are self-adjoint (i.e., they obey  $R^+ = R$ ).

The hermiticity property guarantees that the eigenvalues  $\{ \lambda_m \}$  of such operators are real numbers (i.e., not complex) and that the corresponding eigenfunctions  $\{ f_m \}$ , or their representations  $\{ V_{mk} \}$  within the  $\{ | k \rangle \}$  basis

$$f_m = \sum_k V_{mk} | k \rangle,$$

corresponding to different eigenvalues are orthonormal and that the eigenfunctions belonging to degenerate eigenvalues can be made orthonormal.

To prove these claims, start with  $R | k \rangle = \lambda_k | k \rangle$ . Multiplying on the left by the complex conjugate of  $\langle k |$  and integrating gives  $\langle k | R | k \rangle = \lambda_k \langle k | k \rangle$ . Taking the

complex conjugate of this equation and using the Hermiticity property  $\langle k | R | l \rangle = \langle l | R | k \rangle^*$  (applied with  $k=l$ ) gives  $k^* = k$ .

The orthogonality proof begins with  $R | k \rangle = k | k \rangle$ , and  $R | l \rangle = l | l \rangle$ . Multiplying the first of these on the left by  $\langle l |$  and the second by  $\langle k |$  gives  $\langle l | R | k \rangle = k \langle l | k \rangle$  and  $\langle k | R | l \rangle = l \langle k | l \rangle$ . The complex conjugate of the second reads  $\langle k | R | l \rangle^* = l^* \langle k | l \rangle^*$ ; using the Hermiticity property this reduces to  $\langle l | R | k \rangle = l^* \langle l | k \rangle$ . If  $k \neq l$ , this result can be consistent with  $\langle l | R | k \rangle = k \langle l | k \rangle$  only if  $\langle l | k \rangle$  vanishes.

### III. Meaning of the Eigenvalues and Eigenfunctions

In quantum mechanics, the eigenvalues of an operator represent the only numerical values that can be observed if the physical property corresponding to that operator is measured. Operators for which the eigenvalue spectrum (i.e., the list of eigenvalues) is discrete thus possess discrete spectra when probed experimentally.

For a system in a state that is an eigenfunction of  $R$

$$R \psi = R \psi,$$

measurement of the property corresponding to  $R$  will yield the value  $\lambda$ . For example, if an electron is in a  $2p_{-1}$  orbital and  $L^2$  is measured, the value  $L(L+1) \hbar^2 = 2\hbar^2$  (and only this value) will be observed; if  $L_z$  is measured, the value  $-\hbar$  (and only this value) will be observed. If the electron were in a  $2p_x$  orbital and  $L^2$  were measured, the value  $2\hbar^2$  will be found; however, if  $L_z$  is measured, we can not say that only one value will be observed because the  $2p_x$  orbital is not an eigenfunction of  $L_z$  (measurements in such non-eigenfunction situations are discussed below).

In general, if the property  $R$  is measured, any one of the eigenvalues  $\{ \lambda_m \}$  of the operator  $R$  may be observed. In a large number of such measurements (i.e., for an ensemble of systems all in states described by  $\psi$  that may or may not itself be an eigenfunction of  $R$ ), the probability or frequency of observing the particular eigenvalue  $\lambda_m$  is given by  $|C_m|^2$ , where  $C_m$  is the expansion coefficient of  $\psi$  in the eigenfunctions of  $R$ :

$$\psi = \sum_{\mu} C_{\mu} f_{\mu}.$$

In the special case treated earlier in which  $\psi$  is an eigenfunction of  $R$ , all but one of the  $C_m$  vanish; hence the probability of observing various  $\lambda_m$  values is zero except for the one state for which  $C_m$  is non-zero.

For a measurement that results in the observation of the particular value  $\lambda_m$ , a subsequent measurement of  $R$  on systems just found to have eigenvalue  $\lambda_m$  will result, with 100% certainty, in observation of this same value  $\lambda_m$ . The quantum mechanical interpretation of this experimental observation is to say that the act of measuring the property belonging to the operator  $R$  causes the wavefunction to be altered. Once this

measurement of R is made, the wavefunction is no longer  $\psi$ ; it is now  $f_m$  for those species for which the value  $\lambda_m$  is observed.

For example (this example and others included in this Appendix are also treated more briefly in Chapter 1), if the initial  $\psi$  discussed above were a so-called superposition state of the form

$$\psi = a(2p_0 + 2p_{-1} - 2p_1) + b(3p_0 - 3p_{-1}), \text{ then:}$$

A. If  $L^2$  were measured, the value  $2\hbar^2$  would be observed with probability  $3|a|^2 + 2|b|^2 = 1$ , since all of the non-zero  $C_m$  coefficients correspond to p-type orbitals for this  $\psi$ . After said measurement, the wavefunction would still be this same  $\psi$  because this entire  $\psi$  is an eigenfunction of  $L^2$ .

B. If  $L_z$  were measured for this

$$\psi = a(2p_0 + 2p_{-1} - 2p_1) + b(3p_0 - 3p_{-1}),$$

the values  $0\hbar$ ,  $1\hbar$ , and  $-1\hbar$  would be observed (because these are the only functions with non-zero  $C_m$  coefficients for the  $L_z$  operator) with respective probabilities  $|a|^2 + |b|^2$ ,  $|a|^2$ , and  $|a|^2 + |-b|^2$ .

C. After  $L_z$  were measured, if the sub-population for which  $-1\hbar$  had been detected were subjected to measurement of  $L^2$  the value  $2\hbar^2$  would certainly be found because the new wavefunction

$$\psi' = \{-a(2p_{-1}) - b(3p_{-1})\} (|a|^2 + |b|^2)^{-1/2}$$

is still an eigenfunction of  $L^2$  with this eigenvalue.

D. Again after  $L_z$  were measured, if the sub-population for which  $-1\hbar$  had been observed and for which the wavefunction is now

$$\psi' = \{-a(2p_{-1}) - b(3p_{-1})\} (|a|^2 + |b|^2)^{-1/2}$$

were subjected to measurement of the energy (through the Hamiltonian operator), two values would be found. With probability

$|a|^2 (|a|^2 + |b|^2)^{-1}$  the energy of the  $2p_{-1}$  orbital would be observed; with probability  $|b|^2 (|a|^2 + |b|^2)^{-1}$ , the energy of the  $3p_{-1}$  orbital would be observed.

The general observation to make is that, given an initial normalized  $\psi$  function, and a physical measurement (with operator R) to be made, one first must express  $\psi$  as a linear combination of the complete set of eigenfunctions of that R:

$$\psi = \sum_m C_m f_m.$$

The coefficients  $C_m$  tell, through  $|C_m|^2$ , the probabilities (since  $\psi$  is normalized to unity) of observing each of the  $R$  eigenvalues  $\lambda_m$  when the measurement is made. Once the measurement is made, that sub-population of the sample on which the experiment was run that gave the particular eigenvalue, say  $\lambda_p$ , now have a wavefunction that no longer is the above  $\psi$ ; their wavefunction now is  $\psi_p$ .

#### IV. Experiments Do Not Prepare Only Eigenstates

The above remarks should not be interpreted to mean that experiments are limited to preparing only eigenstates. Just as one can 'pluck' a violin string or 'pound' a drum head in any manner, experiments can prepare a system in states that are not pure eigenfunctions (i.e., states that do not contain just one eigenfunction in their expansion). However, no matter how the state is prepared, it can be interpreted, via expansion in the complete set of eigenfunctions of the operator(s) whose properties are to be measured, as a superposition of eigenfunctions. The superposition amplitudes provide the probabilities of observing each eigenfunction when the measurement is made.

For example, after the drum head has been hit, its shape will evolve spatially and in time in a manner that depends on how it was 'prepared' by the initial blow. However, if one carries out an experiment to detect and frequency-analyze the sound that emanates from this drum, thereby measuring differences in the eigen-energies of the system, one finds a set of discrete (quantized) frequencies  $\{\omega_k\}$  and corresponding amplitudes  $\{A_k\}$ . Repeating this experiment after a different 'blow' is used to prepare a different initial state of the drum head, one finds the same frequencies  $\{\omega_k\}$  but different amplitudes  $\{B_k\}$ .

The quantum mechanics interpretation of these observations is that the initial state of the drum head is a superposition of eigenstates:

$$\psi = \sum_n C_n \psi_n.$$

The  $\{C_n\}$  amplitudes are determined by how the drum head is 'prepared' in the blow that strikes it. As time evolves,  $\psi$  progresses according to the time-dependent Schrödinger equation:

$$i\hbar \frac{\partial \psi}{\partial t} = H \psi.$$

The frequencies emitted by this system will depend on the probability  $|C_n|^2$  of the system being in a particular eigenstate  $\psi_n$  and the energy  $E_n$  of each eigenstate.

The frequency spectrum measured for the drum head motion is therefore seen to have variable amplitudes for each observed 'peak' because the  $|C_n|^2$  vary depending on how the drum head was prepared. In contrast, the frequencies themselves are not characteristic of the preparation process but are properties of the drum head itself; they depend on the eigen-energies of the system, not on how the system is prepared.

This distinction between the characteristic eigenstates of the system with their intrinsic properties and the act of preparing the system in some state that may be a superposition of these eigenstates is essential to keep in mind when applying quantum mechanics to experimental observations.

## V. Operators That Commute and the Experimental Implications

Two hermitian operators that commute

$$[R, S] = RS - SR = 0$$

can be shown to possess complete sets of simultaneous eigenfunctions. That is, one can find complete sets of functions that are eigenfunctions of both R and of S.

The symbol  $[R, S]$  is used to denote what is called the commutator of the operators R and S. There are several useful identities that can be proven for commutators among operators A, B, C, and D, scalar numbers k, and integers n. Several of these are given as follows:

$$\begin{aligned} [A, A^n] &= 0 \\ [kA, B] &= [A, kB] = k[A, B] \\ [A, B+C] &= [A, B] + [A, C] \\ [A+B, C] &= [A, C] + [B, C] \\ [A+B, C+D] &= [A, C] + [A, D] + [B, C] + [B, D] \\ [A, BC] &= [A, B]C + B[A, C] \\ [AB, C] &= [A, C]B + A[B, C] \\ [AB, CD] &= [A, C]DB + C[A, D]B + A[B, C]D + AC[B, D]. \end{aligned}$$

The physical implications of the commutation of two operators are very important because they have to do with what pairs of measurements can be made without interfering with one another. For example, the fact that the x coordinate operator  $\mathbf{x} = x$  and its momentum operator  $\mathbf{p}_x = -i\hbar \frac{d}{dx}$  do not commute results in the well known Heisenberg uncertainty relationship  $\Delta x \Delta p_x \geq \hbar/2$  involving measurements of  $\mathbf{x}$  and  $\mathbf{p}_x$ .

There are two distinct cases that need to be addressed:

A. If the two operators act on different coordinates (or, more generally, on different sets of coordinates), then they obviously commute. Moreover, in this case, it is straightforward to find the complete set of eigenfunctions of both operators; one simply forms a product of any eigenfunction (say  $f_k$ ) of R and any eigenfunction (say  $g_n$ ) of S. The function  $f_k g_n$  is an eigenfunction of both R and S:

$$R f_k g_n = g_n (R f_k) = g_n (\lambda_k f_k) = \lambda_k f_k g_n ;$$

$$S f_k g_n = f_k (S g_n) = f_k (\mu g_n) = \mu f_k g_n .$$

In these equations use has been made of the fact that  $g_n$  and  $f_k$  are functions of different sets of coordinates that S and R, respectively, act upon.

Product functions such as  $f_k g_n$  yield predictable results when measurements are performed. If the property corresponding to R is measured, the value  $\lambda_k$  is observed, and the wavefunction remains

$f_k g_n$ . If S is measured, the value  $\mu$  is observed, and the wavefunction remains  $f_k g_n$ . For example, the two Hermitian operators  $-i\hbar \frac{d}{dx}$  and  $-i\hbar \frac{d}{dy}$  clearly commute. An eigenfunction of

$-i/\hbar$  is of the form  $\exp(ia)$  and an eigenfunction of  $-i/\hbar$  is of the form  $\exp(ibr)$ . The product  $\exp(ia) \exp(ibr)$  is an eigenfunction of both  $-i/\hbar$  and  $-i/\hbar$ . The corresponding eigenvalues are  $a$  and  $b$ , respectively; only these values will be observed if measurements of the properties corresponding to  $-i/\hbar$  and  $-i/\hbar$  are made for a system whose wavefunction is  $\exp(ia) \exp(ibr)$ .

B. If the operators  $R$  and  $S$  act on the same coordinates yet still commute, the implications of their commutation are somewhat more intricate to detail.

As a first step, consider the functions  $\{g_n\}$  that are eigenfunctions of  $S$  with eigenvalues  $\{\mu_n\}$ . Now, act on  $g_n$  with the  $SR$  operator and use the fact that  $SR = RS$  to obtain

$$S R g_n = R S g_n .$$

Because the  $\{g_n\}$  are eigenfunctions of  $S$  having eigenvalues  $\{\mu_n\}$ , this equation further reduces to:

$$S R g_n = R \mu_n g_n = \mu_n R g_n .$$

This is a key result. It shows that the function  $(R g_n)$  is itself either an eigenfunction of  $S$  having the same eigenvalue that  $g_n$  has or it vanishes.

If  $R g_n$  vanishes,  $g_n$  clearly is an eigenfunction of  $R$  (since  $R g_n = 0 g_n$ ) and of  $S$ . On the other hand, if  $R g_n$  is non-vanishing, it must be an eigenfunction of  $S$  having the same eigenvalue ( $\mu_n$ ) as  $g_n$ . If this eigenvalue is non-degenerate (i.e., if  $g_n$  is the only function with eigenvalue  $\mu_n$ ), then  $R g_n$  must be proportional to  $g_n$  itself:

$$R g_n = c_n g_n .$$

This also implies that  $g_n$  is an eigenfunction of both  $R$  and of  $S$ .

Thus far, we can say that functions which are eigenfunctions of  $S$  belonging to non-degenerate eigenvalues must also be eigenfunctions of  $R$ . On the other hand, if the  $\mu_n$  eigenvalue is degenerate (i.e., there are such functions  $g_n, g_n', g_n''$ , etc. that are  $S$ -eigenfunctions with the same  $\mu_n$  as their eigenvalue), all that can be said is that  $R g_n$  is some combination of this  $n$ -fold degenerate manifold of states:

$$R g_n = \sum_{n'} c_{n,n'} g_{n'} .$$

where the sum over  $n'$  runs only over the states with  $S$ -eigenvalues equal to  $\mu_n$ . This same conclusion can be reached no matter which particular state  $g_n$  among the degenerate manifold we begin with. Therefore, the above equation holds for all  $\{g_n\}$  that belong to this degenerate group of states.

The constants  $c_{n,n'}$  form a square (since we act on all states and produce combinations of states) Hermitian (since  $R$  is Hermitian) matrix; in fact,  $c_{n,n'}$  forms the matrix representation of the operator  $R$  within the  $n$ -dimensional space of orthonormal functions  $\{g_n\}$ . As with all Hermitian matrices, a unitary transformation can be employed

to bring it to diagonal form. That is, the orthonormal  $\{g_n\}$  functions can be unitarily combined:

$$G_p = \sum_n U_{p,n} g_n$$

to produce new orthonormal functions  $\{G_p\}$  for which the corresponding matrix elements  $c_{p,p'}$ , defined by

$$\begin{aligned} R G_p &= \sum_n U_{p,n} R g_n = \sum_{n,n'} U_{p,n} c_{n,n'} g_n \\ &= \sum_{p'} \sum_{n,n'} U_{p,n} c_{n,n'} U_{n',p'}^* G_{p'} = \sum_{p'} c_{p,p'} G_{p'} \end{aligned}$$

are diagonal

$$R G_p = c_{p,p} G_p.$$

This shows that the set of functions (the  $G_p$  in this degenerate case) that are eigenfunctions of  $S$  can also be eigenfunctions of  $R$ .

### C. Summary

In summary, we have shown that if  $R$  and  $S$  are operators that act on the same set of coordinates (e.g.,  $-i \hbar / x$  and  $x^2$  or  $\partial^2 / x^2$  and  $-i \hbar / x$ ), then an eigenfunction of  $R$  (denoted  $f_k$  and having eigenvalue  $\lambda_k$ ) must either (i) be eigenfunction of  $S$  (if its  $R$ -eigenvalue is non-degenerate) or (ii) be a member of a degenerate set of  $R$  eigenfunctions that can be combined among one another to produce eigenfunctions of  $S$ .

An example will help illustrate these points. The  $p_x$ ,  $p_y$  and  $p_z$  orbitals are eigenfunctions of the  $L^2$  angular momentum operator with eigenvalues equal to  $L(L+1) \hbar^2 = 2 \hbar^2$ . Since  $L^2$  and  $L_z$  commute and act on the same (angle) coordinates, they possess a complete set of simultaneous eigenfunctions. Although the  $p_x$ ,  $p_y$  and  $p_z$  orbitals are not eigenfunctions of  $L_z$ , they can be combined (as above) to form the  $G_p$  functions) to form three new orbitals:  $p_0 = p_z$ ,  $p_1 = 2^{-1/2} [p_x + i p_y]$ , and  $p_{-1} = 2^{-1/2} [p_x - i p_y]$  that are still eigenfunctions of  $L^2$  but are now eigenfunctions of  $L_z$  also (with eigenvalues  $0\hbar$ ,  $1\hbar$ , and  $-1\hbar$ , respectively).

It should be mentioned that if two operators do not commute, they may still have some eigenfunctions in common, but they will not have a complete set of simultaneous eigenfunctions. For example, the  $L_z$  and  $L_x$  components of the angular momentum operator do not commute; however, a wavefunction with  $L=0$  (i.e., an S-state) is an eigenfunction of both operators.

### D. Experimental Significance

We use an example to illustrate the importance of two operators commuting to quantum mechanics' interpretation of experiments. Assume that an experiment has been carried out on an atom to measure its total angular momentum  $L^2$ . According to quantum mechanics, only values equal to  $L(L+1) \hbar^2$  will be observed. Further assume, for the

particular experimental sample subjected to observation, that values of  $L^2$  equal to  $2\hbar^2$  and  $0\hbar^2$  were detected in relative amounts of 64 % and 36 % , respectively. This means that the atom's original wavefunction could be represented as:

$$\psi = 0.8 P + 0.6 S,$$

where P and S represent the P-state and S-state components of  $\psi$ . The squares of the amplitudes 0.8 and 0.6 give the 64 % and 36 % probabilities mentioned above.

Now assume that a subsequent measurement of the component of angular momentum along the lab-fixed z-axis is to be measured for that sub-population of the original sample found to be in the P-state. For that population, the wavefunction is now a pure P-function:

$$\psi' = P.$$

However, at this stage we have no information about how much of this  $\psi'$  is of  $m = 1, 0,$  or  $-1$ , nor do we know how much  $2p, 3p, 4p, \dots np$  component this state contains.

Because the property corresponding to the operator  $L_z$  is about to be measured, we express the above  $\psi'$  in terms of the eigenfunctions of  $L_z$ :

$$\psi' = P = \sum_{m=1,0,-1} C'_m P_m.$$

When the measurement of  $L_z$  is made, the values  $1\hbar, 0\hbar,$  and  $-1\hbar$  will be observed with probabilities given by  $|C'_1|^2, |C'_0|^2,$  and  $|C'_{-1}|^2$ , respectively. For that sub-population found to have, for example,  $L_z$  equal to  $-1\hbar$ , the wavefunction then becomes

$$\psi'' = P_{-1}.$$

At this stage, we do not know how much of  $2p_{-1}, 3p_{-1}, 4p_{-1}, \dots np_{-1}$  this wavefunction contains. To probe this question another subsequent measurement of the energy (corresponding to the  $H$  operator) could be made. Doing so would allow the amplitudes in the expansion of the above  $\psi'' = P_{-1}$

$$\psi'' = P_{-1} = \sum_n C''_n nP_{-1}$$

to be found.

The kind of experiment outlined above allows one to find the content of each particular component of an initial sample's wavefunction. For example, the original wavefunction has

$0.64 |C''_n|^2 |C'_m|^2$  fractional content of the various  $nP_m$  functions.

Let us consider another experiment in which an initial sample (with wavefunction  $\psi$ ) is first subjected to measurement of  $L_z$  and then subjected to measurement of  $L^2$  and then of the energy. In this order, one would first find specific values (integer multiples of  $\hbar$ ) of  $L_z$  and one would express  $\psi$  as

$$\psi = \sum_m D_m \psi_m.$$

At this stage, the nature of each  $\psi_m$  is unknown (e.g., the  $\psi_1$  function can contain  $n p_1$ ,  $n' d_1$ ,  $n'' f_1$ , etc. components); all that is known is that  $\psi_m$  has  $m \hbar$  as its  $L_z$  value.

Taking that sub-population ( $|D_m|^2$  fraction) with a particular  $m \hbar$  value for  $L_z$  and subjecting it to subsequent measurement of  $L^2$  requires the current wavefunction  $\psi_m$  to be expressed as

$$\psi_m = \sum_L D_{L,m} \psi_{L,m}.$$

When  $L^2$  is measured the value  $L(L+1) \hbar^2$  will be observed with probability  $|D_{m,L}|^2$ , and the wavefunction for that particular sub-population will become

$$\psi'' = \psi_{L,m}.$$

At this stage, we know the value of  $L$  and of  $m$ , but we do not know the energy of the state. For example, we may know that the present sub-population has  $L=1$ ,  $m=-1$ , but we have no knowledge (yet) of how much  $2p_{-1}$ ,  $3p_{-1}$ , ...  $np_{-1}$  the system contains.

To further probe the sample, the above sub-population with  $L=1$  and  $m=-1$  can be subjected to measurement of the energy. In this case, the function  $\psi_{1,-1}$  must be expressed as

$$\psi_{1,-1} = \sum_n D_n'' \psi_{n,-1}.$$

When the energy measurement is made, the state  $\psi_{n,-1}$  will be found  $|D_n''|^2$  fraction of the time.

We now need to explain how the fact that  $L_z$ ,  $L^2$ , and  $H$  all commute with one another (i.e., are mutually commutative) makes the series of measurements described above more straightforward than if these operators did not commute. In the first experiment, the fact that they are mutually commutative allowed us to expand the 64 % probable  $L^2$  eigenstate with  $L=1$  in terms of functions that were eigenfunctions of the operator for which measurement was about to be made without destroying our knowledge of the value of  $L^2$ . That is, because  $L^2$  and  $L_z$  can have simultaneous eigenfunctions, the  $L=1$  function can be expanded in terms of functions that are eigenfunctions of both  $L^2$  and  $L_z$ . This in turn, allowed us to find experimentally the sub-population that had, for example a  $-1 \hbar$  value of  $L_z$  while retaining knowledge that the state remains an eigenstate of  $L^2$  (the state at this time had  $L=1$  and  $m=-1$  and was denoted  $\psi_{1,-1}$ ) Then, when this  $\psi_{1,-1}$  state was subjected to energy measurement, knowledge of the energy of the sub-population could be gained without giving up knowledge of the  $L^2$  and  $L_z$  information; upon carrying out said measurement, the state became  $\psi_{n,-1}$ .

In contrast, if (hypothetically)  $L^2$  and  $L_z$  did not commute, the  $L=1$  function originally detected with 64 % probability would be altered by the subsequent  $L_z$  measurement in a manner that destroys our knowledge of  $L^2$ . The  $\psi$  function could still have been expanded in terms of the eigenfunctions of the property about to be probed ( $L_z$ )

$$\psi = \sum_{m=1,0,-1} C'_m \psi_m.$$

However, because  $L^2$  and  $L_z$  do not commute in this hypothetical example, the states  $\psi_m$  that are eigenfunctions of  $L_z$  will not, in general, also be eigenfunctions of  $L^2$ . Hence, when  $L_z$  is measured and a particular value (say  $-1\hbar$ ) is detected, the wavefunction becomes

$$\psi' = \psi_{-1},$$

which is no longer an eigenfunction of  $L^2$ .

The essential observations to be made are:

1. After the first measurement is made (say for operator R), the wavefunction becomes an eigenfunction of R with a well defined R-eigenvalue (say  $r$ ):  $\psi = \psi(r)$ .

2. The eigenfunctions of the second operator S (i.e., the operator corresponding to the measurement about to be made) can be taken to also be eigenfunctions of R if R and S commute. This then allows  $\psi(r)$  to be expanded in terms of functions that are both R-eigenfunctions (with eigenvalue  $r$ ) and S-eigenfunctions (with various eigenvalues  $\mu_n$ ):

$\psi(r) = \sum_n C_n \psi(r, \mu_n)$ . Upon measurement of S, the wavefunction becomes one of these  $\psi(r, \mu_n)$  functions. When the system is in this state, both R- and S- eigenvalues are known precisely; they are  $r$  and  $\mu_n$ .

3. The eigenfunctions of the second operator S (i.e., the operator corresponding to the measurement about to be made) can not be taken to also be eigenfunctions of R if R and S do not commute. The function  $\psi(r)$  can still be expanded in terms of functions that are both S-eigenfunctions (with various eigenvalues  $\mu_n$ ):  $\psi(r) = \sum_n C_n \psi(\mu_n)$ . However, because R and S do not commute, these  $\psi(\mu_n)$  functions are not, in general, also R-eigenfunctions; they are only S-eigenfunctions. Then, upon measurement of S, the wavefunction becomes one of these  $\psi(\mu_n)$  functions. When the system is in this state, the S- eigenvalue is known precisely; it is  $\mu_n$ . The R-eigenvalue is no longer specified. In fact, the new state  $\psi(\mu_n)$  may contain components of all different R-eigenvalues, which can be represented by expanding  $\psi(\mu_n)$  in terms of the R-eigenfunctions:  $\psi(\mu_n) = \sum_k D_k \psi(r_k)$ . If R were measured again, after the state has been prepared in  $\psi(\mu_n)$ , the R-eigenvalues  $\{r_k\}$  would be observed with probabilities  $\{|D_k|^2\}$ , and the wavefunction would, for these respective sub-populations, become  $\psi(r_k)$ .

It should now be clear that the act of carrying out an experimental measurement disturbs the system in that it causes the system's wavefunction to become an eigenfunction of the operator whose property is measured. If two properties whose corresponding operators commute are measured, the measurement of the second property does not destroy knowledge of the first property's value gained in the first measurement. If the two properties do not commute, the second measurement does destroy knowledge of the first property's value. It is thus often said that 'measurements for operators that do not commute interfere with one another'.

## Time Independent Perturbation Theory D

Perturbation theory is used in two qualitatively different contexts in quantum chemistry. It allows one to estimate (because perturbation theory is usually employed through some finite order and may not even converge if extended to infinite order) the splittings and shifts in energy levels and changes in wavefunctions that occur when an external field (e.g., electric or magnetic or that due to a surrounding set of 'ligands'- a crystal field) is applied to an atom, molecule, ion, or solid whose 'unperturbed' states are known. These 'perturbations' in energies and wavefunctions are expressed in terms of the (complete) set of unperturbed states. For example, the distortion of the 2s orbital of a Li atom due to the application of an external electric field along the y-axis is described by adding to the (unperturbed) 2s orbital components of p<sub>y</sub>-type orbitals (2p, 3p, etc.):

$$\psi = a 2s + \sum_n C_n n p_y.$$

The effect of adding in the p<sub>y</sub> orbitals is to polarize the 2s orbital along the y-axis. The amplitudes C<sub>n</sub> are determined via the equations of perturbation theory developed below; the change in the energy of the 2s orbital caused by the application of the field is expressed in terms of the C<sub>n</sub> coefficients and the (unperturbed) energies of the 2s and n p<sub>y</sub> orbitals.

There is another manner in which perturbation theory is used in quantum chemistry that does not involve an externally applied perturbation. Quite often one is faced with solving a Schrödinger equation to which no exact solution has been (yet) or can be found. In such cases, one often develops a 'model' Schrödinger equation which in some sense is designed to represent the system whose full Schrödinger equation can not be solved. The difference between the Hamiltonia of the full and model problems, H and H<sup>0</sup>, respectively is used to define a perturbation V=H-H<sup>0</sup>. Perturbation theory is then employed to approximate the energy levels and wavefunctions of the full H in terms of the energy levels and wavefunctions of the model system (which, by assumption, can be found). The 'imperfection' in the model problem is therefore used as the perturbation. The success of such an approach depends strongly on how well the model H<sup>0</sup> represents the true problem (i.e., on how 'small' V is). For this reason, much effort is often needed to develop approximate Hamiltonia for which V is small and for which the eigenfunctions and energy levels can be found.

### I. Structure of Time-Independent Perturbation Theory

#### A. The Power Series Expansions of the Wavefunction and Energy

Assuming that all wavefunctions  $\psi_k$  and energies  $E_k^0$  belonging to the unperturbed Hamiltonian  $H^0$  are known

$$H^0 \psi_k = E_k^0 \psi_k,$$

and given that one wishes to find eigenstates ( $\psi_k$  and  $E_k$ ) of the perturbed Hamiltonian

$$H=H^0+ V,$$

perturbation theory begins by expressing  $\psi_k$  and  $E_k$  as power series in the perturbation strength  $\lambda$ :

$$\psi_k = \sum_{n=0}^{\infty} \lambda^n \psi_k^{(n)}$$

$$E_k = \sum_{n=0}^{\infty} \lambda^n E_k^{(n)}$$

Moreover, it is assumed that, as the strength of the perturbation is reduced to zero,  $\psi_k$  reduces to one of the unperturbed states  $\psi_k^0$  and that the full content of  $\psi_k$  in  $\psi_k^0$  is contained in the first term  $\psi_k^{(0)}$ .

This means that  $\psi_k^{(0)} = \psi_k^0$  and  $E_k^{(0)} = E_k^0$ , and so

$$\psi_k = \psi_k^0 + \sum_{n=1}^{\infty} \lambda^n \psi_k^{(n)} = \psi_k^0 + \psi_k'$$

$$E_k = E_k^0 + \sum_{n=1}^{\infty} \lambda^n E_k^{(n)} = E_k^0 + E_k'$$

In the above expressions,  $\lambda$  would be proportional to the strength of the electric or magnetic field if one is dealing with an external-field case. When dealing with the situation for which  $V$  is the imperfection in the model  $H^0$ ,  $\lambda$  is equal to unity; in this case, one thinks of formulating and solving for the perturbation expansion for variable  $\lambda$  after which  $\lambda$  is set equal to unity.

## B. The Order-by-Order Energy Equations

Equations for the order-by-order corrections to the wavefunctions and energies are obtained by using these power series expressions in the full Schrödinger equation:

$$(H - E_k) \psi_k = 0.$$

Multiplying through by  $\psi_k^*$  and integrating gives the expression in terms of which the total energy is obtained:

$$\langle \psi_k | H | \psi_k \rangle = E_k \langle \psi_k | \psi_k \rangle = E_k.$$

Using the fact that  $\psi_k^0$  is an eigenfunction of  $H^0$  and employing the power series expansion of  $\psi_k$  allows one to generate the fundamental relationships among the energies  $E_k^{(n)}$  and the wavefunctions  $\psi_k^{(n)}$ :

$$E_k = \langle \psi_k | H^0 | \psi_k \rangle + \langle \psi_k | V | \psi_k \rangle = E_k^0 + \langle \psi_k | V | \sum_{n=0}^{\infty} \lambda^n \psi_k^{(n)} \rangle.$$

The lowest few orders in this expansion read as follows:

$$E_k = E_k^0 + \langle \psi_k^0 | V | \psi_k^0 \rangle + 2 \langle \psi_k^0 | V | \psi_k^{(1)} \rangle + 3 \langle \psi_k^0 | V | \psi_k^{(2)} \rangle + \dots$$

If the various  $\psi_k^{(n)}$  can be found, then this equation can be used to compute the order-by-order energy expansion.

Notice that the first-order energy correction is given in terms of the zeroth-order (i.e., unperturbed) wavefunction as:

$$E_k^{(1)} = \langle \psi_k | V | \psi_k \rangle,$$

the average value of the perturbation taken over  $\psi_k$ .

### C. The Order-by-Order Wavefunction Equations

To obtain workable expressions for the perturbative corrections to the wavefunction  $\psi_k$ , the full Schrödinger equation is first projected against all of the unperturbed eigenstates  $\{ | \psi_j \rangle \}$  other than the state  $\psi_k$  whose perturbative corrections are sought:

$$\langle \psi_j | H | \psi_k \rangle = E_k \langle \psi_j | \psi_k \rangle, \text{ or}$$

$$\langle \psi_j | H^0 | \psi_k \rangle + \langle \psi_j | V | \psi_k \rangle = E_k \langle \psi_j | \psi_k \rangle, \text{ or}$$

$$\langle \psi_j | \psi_k \rangle E_j^0 + \langle \psi_j | V | \psi_k \rangle = E_k \langle \psi_j | \psi_k \rangle, \text{ or finally}$$

$$\langle \psi_j | V | \psi_k \rangle = [ E_k - E_j^0 ] \langle \psi_j | \psi_k \rangle.$$

Next, each component  $\psi_k^{(n)}$  of the eigenstate  $\psi_k$  is expanded in terms of the unperturbed eigenstates (as they can be because the  $\{ \psi_k \}$  form a complete set of functions):

$$\psi_k = \psi_k + \sum_{j \neq k, n=1}^{\infty} \langle \psi_j | \psi_k^{(n)} \rangle | \psi_j \rangle.$$

Substituting this expansion for  $\psi_k$  into the preceding equation gives

$$\langle \psi_j | V | \psi_k \rangle + \sum_{l \neq k, n=1}^{\infty} \langle \psi_l | \psi_k^{(n)} \rangle \langle \psi_j | V | \psi_l \rangle$$

$$= [ E_k - E_j^0 ] \sum_{n=1}^{\infty} \langle \psi_j | \psi_k^{(n)} \rangle.$$

To extract from this set of coupled equations relations that can be solved for the coefficients  $\langle \psi_j | \psi_k^{(n)} \rangle$ , which embodies the desired wavefunction perturbations  $\psi_k^{(n)}$ , one collects together all terms with like power of  $\lambda$  in the above general equation (in doing so, it is important to keep in mind that  $E_k$  itself is given as a power series in  $\lambda$ ).

The  $\lambda^0$  terms vanish, and the first-order terms reduce to:

$$\langle \psi_j | V | \psi_k \rangle = [ E_k^0 - E_j^0 ] \langle \psi_j | \psi_k^{(1)} \rangle,$$

which can be solved for the expansion coefficients of the so-called first-order wavefunction  $\psi_k^{(1)}$ :

$$\psi_k^{(1)} = \sum_j \langle j | V | k \rangle / [E_k^0 - E_j^0] | j \rangle .$$

When this result is used in the earlier expression for the second-order energy correction, one obtains:

$$E_k^{(2)} = \sum_j \langle j | V | k \rangle^2 / [E_k^0 - E_j^0] .$$

The terms proportional to  $\psi_k^{(2)}$  are as follows:

$$\begin{aligned} \langle 1 | k \rangle \langle 1 | \psi_k^{(1)} \rangle &= \langle j | V | 1 \rangle \\ &= [E_k^0 - E_j^0] \langle j | \psi_k^{(2)} \rangle + E_k^{(1)} \langle j | \psi_k^{(1)} \rangle . \end{aligned}$$

The solution to this equation can be written as:

$$\langle j | \psi_k^{(2)} \rangle = [E_k^0 - E_j^0]^{-1} \langle 1 | k \rangle \langle 1 | \psi_k^{(1)} \rangle \{ \langle j | V | 1 \rangle - \sum_l E_k^{(1)} \} .$$

Because the expansion coefficients  $\langle 1 | \psi_k^{(1)} \rangle$  of  $\psi_k^{(1)}$  are already known, they can be used to finally express the expansion coefficients of  $\psi_k^{(2)}$  totally in terms of zeroth-order quantities:

$$\begin{aligned} \langle j | \psi_k^{(2)} \rangle &= [E_k^0 - E_j^0]^{-1} \langle 1 | k \rangle \{ \langle j | V | 1 \rangle - \sum_l E_k^{(1)} \} \\ &\langle 1 | V | k \rangle [E_k^0 - E_l^0]^{-1} , \end{aligned}$$

which then gives

$$\begin{aligned} \psi_k^{(2)} &= \sum_j \langle j | \psi_k^{(2)} \rangle | j \rangle \\ &= \sum_j \langle j | V | k \rangle [E_k^0 - E_l^0]^{-1} | j \rangle . \end{aligned}$$

#### D. Summary

An essential thing to stress concerning the above development of so-called Rayleigh-Schrödinger perturbation theory (RSPT) is that each of the energy corrections  $E_k^{(n)}$  and wavefunction corrections  $\psi_k^{(n)}$  are expressed in terms of integrals over the unperturbed wavefunctions  $\psi_k$  involving the perturbation (i.e.,  $\langle j | V | 1 \rangle$ ) and the unperturbed energies  $E_j^0$ . As such, these corrections can be symmetry-analyzed to determine, for example, whether perturbations of a given symmetry will or will not affect particular states. For example, if the state under study belongs to a non-degenerate representation in the absence of the perturbation  $V$ , then its first-order energy correction  $\langle k | V | k \rangle$  will be non-zero only if  $V$  contains a totally symmetric component (because the direct product of the symmetry of  $\psi_k$  with itself is the totally symmetric representation). Such an analysis predicts, for example, that the energy of an s orbital of an atom will be unchanged, in first-order, by the application of an external electric field because the perturbation

$$V = e \mathbf{E} \cdot \mathbf{r}$$

is odd under the inversion operation (and hence can not be totally symmetric). This same analysis, when applied to  $E_k^{(2)}$  shows that contributions to the second-order energy of an orbital arise only from unperturbed orbitals  $\psi_j$  that are odd under inversion because only in such cases will the integrals  $\langle \psi_s | e \mathbf{E} \cdot \mathbf{r} | \psi_j \rangle$  be non-zero.

## II. The Møller-Plesset Perturbation Series

### A. The Choice of $H^0$

Let us assume that an SCF calculation has been carried out using the set of  $N$  spin-orbitals  $\{\psi_a\}$  that are occupied in the reference configuration  $\psi_k$  to define the corresponding Fock operator:

$$F = h + \sum_{a(\text{occupied})} [J_a - K_a] .$$

Further, we assume that all of the occupied  $\{\psi_a\}$  and virtual  $\{\psi_m\}$  spin-orbitals and orbital energies have been determined and are available.

This Fock operator is used to define the unperturbed Hamiltonian of Møller-Plesset perturbation theory (MPPT):

$$H^0 = \sum_i F(\mathbf{r}_i).$$

This particular Hamiltonian, when acting on any Slater determinant formed by placing  $N$  electrons into the SCF spin-orbitals, yields a zeroth order eigenvalue equal to the sum of the orbital energies of the spin-orbitals appearing in that determinant:

$$H^0 | \psi_{j_1} \psi_{j_2} \psi_{j_3} \psi_{j_4} \dots \psi_{j_N} \rangle = (\epsilon_{j_1} + \epsilon_{j_2} + \epsilon_{j_3} + \epsilon_{j_4} + \dots + \epsilon_{j_N}) | \psi_{j_1} \psi_{j_2} \psi_{j_3} \psi_{j_4} \dots \psi_{j_N} \rangle$$

because the spin-orbitals obey

$$F \psi_j = \epsilon_j \psi_j,$$

where  $j$  runs over all (occupied (a, b, ...) and virtual (m, n, ...)) spin-orbitals. This result is the MPPT embodiment of  $H^0 \psi_k = E_k^0 \psi_k$ .

### B. The Perturbation $V$

The perturbation  $V$  appropriate to this MPPT case is the difference between the full  $N$ -electronic Hamiltonian and this  $H^0$ :

$$V = H - H^0.$$

Matrix elements of  $V$  among determinantal wavefunctions constructed from the SCF spin-orbitals  $\langle \psi_l | V | \psi_k \rangle$  can be expressed, using the Slater-Condon rules, in terms of matrix elements over the full Hamiltonian  $H$

$$\langle \phi_k | V | \phi_k \rangle = \langle \phi_k | H | \phi_k \rangle - \sum_{i,j} E_k^0,$$

because each such determinant is an eigenfunction of  $H^0$ .

### C. The MPPT Energy Corrections

Given this particular choice of  $H^0$ , it is possible to apply the general RSPT energy and wavefunction correction formulas developed above to generate explicit results in terms of spin-orbital energies and one- and two-electron integrals,  $\langle \phi_i | h | \phi_j \rangle$  and  $\langle \phi_i \phi_j | g | \phi_k \phi_l \rangle = \langle ij | kl \rangle$ , over these spin-orbitals. In particular, the first-order energy correction is given as follows:

$$\begin{aligned} E_k^{(1)} &= \langle \phi_k | V | \phi_k \rangle = \langle \phi_k | H | \phi_k \rangle - \sum_a \sum_a \\ &= \sum_a \sum_a - \sum_{a,b} [J_{a,b} - K_{a,b}] - \sum_a \sum_a \\ &= - \sum_{a,b} [J_{a,b} - K_{a,b}] = - \sum_{a,b} [\langle ab | ab \rangle - \langle ab | ba \rangle]. \end{aligned}$$

Thus  $E_k^0$  (the sum of orbital energies) and  $E_k^{(1)}$  (the correction for double counting) add up to produce the proper expectation value energy.

The second-order energy correction can be evaluated in like fashion by noting that  $\langle \phi_k | H | \phi_l \rangle = 0$  according to the Brillouin theorem for all singly excited  $\phi_l$ , and that  $\langle \phi_k | H | \phi_l \rangle = \langle ab | mn \rangle - \langle ab | nm \rangle$  for doubly excited  $\phi_l$  in which excitations from  $\phi_a$  and  $\phi_b$  into  $\phi_m$  and  $\phi_n$  are involved:

$$\begin{aligned} E_k^{(2)} &= \sum_j \langle \phi_j | V | \phi_k \rangle^2 / [E_k^0 - E_j^0] \\ &= \sum_{a,b;m<n} |\langle ab | mn \rangle - \langle ab | nm \rangle|^2 / (\epsilon_a + \epsilon_b - \epsilon_m - \epsilon_n). \end{aligned}$$

### D. The Wavefunction Corrections

The first-order MPPT wavefunction can be evaluated in terms of Slater determinants that are excited relative to the SCF reference function  $\phi_k$ . Realizing again that the perturbation coupling matrix elements  $\langle \phi_k | H | \phi_l \rangle$  are non-zero only for doubly excited CSF's, and denoting such doubly excited  $\phi_l$  by  $\phi_{a,b;m,n}$ , the first-order wavefunction can be written as:

$$\begin{aligned} \phi_k^{(1)} &= \sum_j \langle \phi_j | V | \phi_k \rangle / [E_k^0 - E_j^0] | \phi_j \rangle \\ &= \sum_{a,b;m<n} \phi_{a,b;m,n} [\langle ab | mn \rangle - \langle ab | nm \rangle] / (\epsilon_a + \epsilon_b - \epsilon_m - \epsilon_n). \end{aligned}$$

## III. Conceptual Use of Perturbation Theory

The first- and second- order RSPT energy and first-order RSPT wavefunction correction expressions form not only a useful computational tool but are also of great use in understanding how strongly a perturbation will affect a particular state of the system. By

examining the symmetries of the state of interest  $\psi_k$  (this can be an orbital of an atom or molecule, an electronic state of same, or a vibrational/rotational wavefunction of a molecule) and of the perturbation  $V$ , one can say whether  $V$  will have a significant effect on the energy  $E_k$  of  $\psi_k$ ; if  $\langle \psi_k | V | \psi_k \rangle$  is non-zero, the effect can be expected to be significant.

Sometimes the perturbation is of the wrong symmetry to directly (i.e., in a first-order manner) affect  $E_k$ . In such cases, one considers whether nearby states  $\{ \psi_j, E_j \}$  exist which could couple through  $V$  with  $\psi_k$ ; the second-order energy expression, which contains  $\sum_j |\langle \psi_j | V | \psi_k \rangle|^2 / [E_k^0 - E_j^0]$  directs one to seek states whose symmetries are contained in the direct product of the symmetries of  $V$  and of  $\psi_k$  and which are close to  $E_k$  in energy.

It is through such symmetry and 'coupling matrix element' considerations that one can often 'guess' whether a given perturbation will have an appreciable effect on the state of interest.

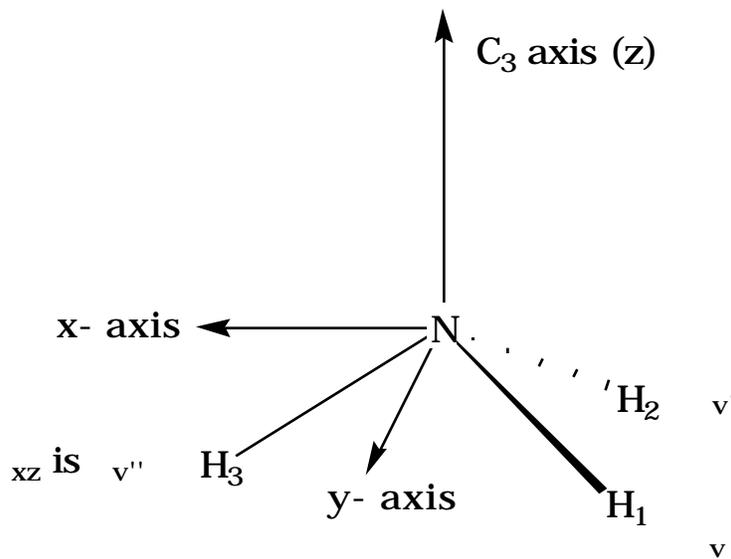
The nature of the perturbation is not important to such considerations. It could be the physical interaction that arises as two previously non-interacting atoms are brought together (in which case  $V$  would have axial point group symmetry) or it could describe the presence of surrounding ligands on a central transition metal ion (in which case  $V$  would carry the symmetry of the 'ligand field'). Alternatively, the perturbation might describe the electric dipole interaction of the electrons and nuclei of the atom or molecule with an externally applied electric field  $\mathbf{E}$ , in which case  $V = - \sum_j e \mathbf{r}_j \cdot \mathbf{E} + \sum_a Z_a e \mathbf{R}_a \cdot \mathbf{E}$  contains components that transform as  $x$ ,  $y$ , and  $z$  in the point group appropriate to the system (because the electronic  $\mathbf{r}_j$  and nuclear  $\mathbf{R}_a$  coordinate vectors so transform).

## Point Group Symmetry E

It is assumed that the reader has previously learned, in undergraduate inorganic or physical chemistry classes, how symmetry arises in molecular shapes and structures and what symmetry elements are (e.g., planes, axes of rotation, centers of inversion, etc.). For the reader who feels, after reading this appendix, that additional background is needed, the texts by Cotton and EWK, as well as most physical chemistry texts can be consulted. We review and teach here only that material that is of direct application to symmetry analysis of molecular orbitals and vibrations and rotations of molecules. We use a specific example, the ammonia molecule, to introduce and illustrate the important aspects of point group symmetry.

### I. The $C_{3v}$ Symmetry Group of Ammonia - An Example

The ammonia molecule  $\text{NH}_3$  belongs, in its ground-state equilibrium geometry, to the  $C_{3v}$  point group. Its symmetry operations consist of two  $C_3$  rotations,  $C_3$ ,  $C_3^2$  (rotations by  $120^\circ$  and  $240^\circ$ , respectively about an axis passing through the nitrogen atom and lying perpendicular to the plane formed by the three hydrogen atoms), three vertical reflections,  $\sigma_v$ ,  $\sigma_v'$ ,  $\sigma_v''$ , and the identity operation. Corresponding to these six operations are symmetry elements: the three-fold rotation axis,  $C_3$  and the three symmetry planes  $\sigma_v$ ,  $\sigma_v'$  and  $\sigma_v''$  that contain the three NH bonds and the z-axis (see figure below).



These six symmetry operations form a mathematical group. A group is defined as a set of objects satisfying four properties.

1. A combination rule is defined through which two group elements are combined to give a result which we call the product. The product of two elements in the group must also be a member of the group (i.e., the group is closed under the combination rule).
2. One special member of the group, when combined with any other member of the group, must leave the group member unchanged (i.e., the group contains an identity element).

- Every group member must have a reciprocal in the group. When any group member is combined with its reciprocal, the product is the identity element.
- The associative law must hold when combining three group members (i.e.,  $(AB)C$  must equal  $A(BC)$ ).

The members of symmetry groups are symmetry operations; the combination rule is successive operation. The identity element is the operation of doing nothing at all. The group properties can be demonstrated by forming a multiplication table. Let us label the rows of the table by the first operation and the columns by the second operation. Note that this order is important because most groups are not commutative. The  $C_{3v}$  group multiplication table is as follows:

	E	$C_3$	$C_3^2$	$\nu$	$\nu'$	$\nu''$	second operation
E	E	$C_3$	$C_3^2$	$\nu$	$\nu'$	$\nu''$	
$C_3$	$C_3$	$C_3^2$	E	$\nu'$	$\nu''$	$\nu$	
$C_3^2$	$C_3^2$	E	$C_3$	$\nu''$	$\nu$	$\nu'$	
$\nu$	$\nu$	$\nu''$	$\nu'$	E	$C_3^2$	$C_3$	
$\nu'$	$\nu'$	$\nu$	$\nu''$	$C_3$	E	$C_3^2$	
$\nu''$	$\nu''$	$\nu'$	$\nu$	$C_3^2$	$C_3$	E	
First operation							

Note the reflection plane labels do not move. That is, although we start with  $H_1$  in the  $\nu$  plane,  $H_2$  in  $\nu''$ , and  $H_3$  in  $\nu'$ , if  $H_1$  moves due to the first symmetry operation,  $\nu$  remains fixed and a different H atom lies in the  $\nu$  plane.

## II. Matrices as Group Representations

In using symmetry to help simplify molecular orbital or vibration/rotation energy level calculations, the following strategy is followed:

- A set of  $M$  objects belonging to the constituent atoms (or molecular fragments, in a more general case) is introduced. These objects are the orbitals of the individual atoms (or of the fragments) in the m.o. case; they are unit vectors along the  $x$ ,  $y$ , and  $z$  directions located on each of the atoms, and representing displacements along each of these directions, in the vibration/rotation case.
- Symmetry tools are used to combine these  $M$  objects into  $M$  new objects each of which belongs to a specific symmetry of the point group. Because the hamiltonian (electronic in the m.o. case and vibration/rotation in the latter case) commutes with the symmetry operations of the point group, the matrix representation of  $\mathbf{H}$  within the symmetry adapted basis will be "block diagonal". That is, objects of different symmetry will not interact; only interactions among those of the same symmetry need be considered.

To illustrate such symmetry adaptation, consider symmetry adapting the  $2s$  orbital of N and the three  $1s$  orbitals of H. We begin by determining how these orbitals transform under the symmetry operations of the  $C_{3v}$  point group. The act of each of the six symmetry operations on the four atomic orbitals can be denoted as follows:

$(S_N, S_1, S_2, S_3)$	E	$(S_N, S_1, S_2, S_3)$
	$C_3$	$(S_N, S_3, S_1, S_2)$
	$C_3^2$	$(S_N, S_2, S_3, S_1)$
	$\nu$	$(S_N, S_1, S_3, S_2)$
	$\nu''$	$(S_N, S_3, S_2, S_1)$
	$\nu'$	$(S_N, S_2, S_1, S_3)$

Here we are using the active view that a  $C_3$  rotation rotates the molecule by  $120^\circ$ . The equivalent passive view is that the 1s basis functions are rotated  $-120^\circ$ . In the  $C_3$  rotation,  $S_3$  ends up where  $S_1$  began,  $S_1$ , ends up where  $S_2$  began and  $S_2$  ends up where  $S_3$  began.

These transformations can be thought of in terms of a matrix multiplying a vector with elements  $(S_N, S_1, S_2, S_3)$ . For example, if  $D^{(4)}(C_3)$  is the representation matrix giving the  $C_3$  transformation, then the above action of  $C_3$  on the four basis orbitals can be expressed as:

$$D^{(4)}(C_3) \begin{pmatrix} S_N \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} S_N \\ S_1 \\ S_2 \\ S_3 \end{pmatrix} = \begin{pmatrix} S_N \\ S_3 \\ S_1 \\ S_2 \end{pmatrix}$$

We can likewise write matrix representations for each of the symmetry operations of the  $C_{3v}$  point group:

$$D^{(4)}(C_3^2) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad D^{(4)}(E) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$D^{(4)}(\nu) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad D^{(4)}(\nu') = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$D^{(4)}(\sigma_v'') = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

It is easy to verify that a  $C_3$  rotation followed by a  $\sigma_v$  reflection is equivalent to a  $\sigma_v'$  reflection alone. In other words

$$\sigma_v C_3 = \sigma_v', \text{ or, } \begin{matrix} & S_1 & & S_3 & & S_3 \\ & & C_3 & & & \sigma_v \\ \sigma_v C_3 = \sigma_v' & S_2 & S_3 & S_1 & S_2 & S_2 & S_1 \end{matrix}$$

Note that this same relationship is carried by the matrices:

$$D^{(4)}(\sigma_v) D^{(4)}(C_3) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} = D^{(4)}(\sigma_v')$$

Likewise we can verify that  $C_3 \sigma_v = \sigma_v''$  directly and we can notice that the matrices also show the same identity:

$$D^{(4)}(C_3) D^{(4)}(\sigma_v) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = D^{(4)}(\sigma_v'')$$

In fact, one finds that the six matrices,  $D^{(4)}(R)$ , when multiplied together in all 36 possible ways obey the same multiplication table as did the six symmetry operations. We say the matrices form a representation of the group because the matrices have all the properties of the group.

### A. Characters of Representations

One important property of a matrix is the sum of its diagonal elements

$$\text{Tr}(D) = \sum_i D_{ii} = \chi$$

is called the trace or character of the matrix. In the above example

$$\chi(E) = 4$$

$$\chi(C_3) = \chi(C_3^2) = 1$$

$$\chi(\sigma_v) = \chi(\sigma_v') = \chi(\sigma_v'') = 2.$$

The importance of the characters of the symmetry operations lies in the fact that they do not depend on the specific basis used to form them. That is, they are invariant to a unitary or orthogonal transformation of the objects used to define the matrices. As a result, they contain information about the symmetry operation itself and about the space spanned by the set of objects. The significance of this observation for our symmetry adaptation process will become clear later.

Note that the characters of both rotations are the same as are those of all three reflections. Collections of operations having identical characters are called classes. Each operation in a class of operations has the same character as other members of the class. The character of a class depends on the space spanned by the basis of functions on which the symmetry operations act. Above we used  $(S_N, S_1, S_2, S_3)$  as a basis.

## B. Another Basis and Another Representation

If, alternatively, we use the one-dimensional basis consisting of the 1s orbital on the N-atom, we obtain different characters, as we now demonstrate.

The act of the six symmetry operations on this  $S_N$  can be represented as follows:

$$\begin{array}{ccc} S_N \begin{array}{c} E \\ S_N \end{array} S_N; & S_N \begin{array}{c} C_3 \\ S_N \end{array} S_N; & S_N \begin{array}{c} C_3^2 \\ S_N \end{array} S_N; \\ S_N \begin{array}{c} \nu \\ S_N \end{array} S_N; & S_N \begin{array}{c} \nu' \\ S_N \end{array} S_N; & S_N \begin{array}{c} \nu'' \\ S_N \end{array} S_N. \end{array}$$

We can represent this group of operations in this basis by the one-dimensional set of matrices:

$$\begin{array}{ccc} D^{(1)}(E) = 1; & D^{(1)}(C_3) = 1; & D^{(1)}(C_3^2) = 1, \\ D^{(1)}(\nu) = 1; & D^{(1)}(\nu'') = 1; & D^{(1)}(\nu') = 1. \end{array}$$

Again we have

$$D^{(1)}(\nu) D^{(1)}(C_3) = 1 \quad 1 = D^{(1)}(\nu''), \text{ and}$$

$$D^{(1)}(C_3) D^{(1)}(\nu) = 1 \quad 1 = D^{(1)}(\nu').$$

These six matrices form another representation of the group. In this basis, each character is equal to unity. The representation formed by allowing the six symmetry operations to act on the 1s N-atom orbital is clearly not the same as that formed when the same six operations acted on the  $(S_N, S_1, S_2, S_3)$  basis. We now need to learn how to further analyze the information content of a specific representation of the group formed when the symmetry operations act on any specific set of objects.

## III. Reducible and Irreducible Representations

### A. A Reducible Representation

Note that every matrix in the four dimensional group representation labeled  $D^{(4)}$  has the so-called block diagonal form

$$\begin{array}{|c|ccc|} \hline 1 & 0 & 0 & 0 \\ \hline 0 & \multicolumn{3}{|c|}{3 \times 3 \text{ matrix}} \\ \hline 0 & & & \\ \hline 0 & & & \\ \hline \end{array}$$

This means that these  $D^{(4)}$  matrices are really a combination of two separate group representations (mathematically, it is called a direct sum representation). We say that  $D^{(4)}$  is reducible into a one-dimensional representation  $D^{(1)}$  and a three-dimensional representation formed by the  $3 \times 3$  submatrices which we will call  $D^{(3)}$ .

$$\begin{array}{l}
 \begin{array}{ccc}
 1 & 0 & 0 \\
 0 & 1 & 0 \\
 0 & 0 & 1
 \end{array} ; \quad
 \begin{array}{ccc}
 0 & 0 & 1 \\
 1 & 0 & 0 \\
 0 & 1 & 0
 \end{array} ; \quad
 \begin{array}{ccc}
 0 & 1 & 0 \\
 0 & 0 & 1 \\
 1 & 0 & 0
 \end{array} \\
 \\
 \begin{array}{ccc}
 1 & 0 & 0 \\
 0 & 0 & 1 \\
 0 & 1 & 0
 \end{array} ; \quad
 \begin{array}{ccc}
 0 & 0 & 1 \\
 0 & 1 & 0 \\
 1 & 0 & 0
 \end{array} ; \quad
 \begin{array}{ccc}
 0 & 1 & 0 \\
 1 & 0 & 0 \\
 0 & 0 & 1
 \end{array}
 \end{array}$$

The characters of  $D^{(3)}$  are  $\chi(E) = 3$ ,  $\chi(2C_3) = 0$ ,  $\chi(3C_2) = 1$ . Note that we would have obtained this  $D^{(3)}$  representation directly if we had originally chosen to examine the basis  $(S_1, S_2, S_3)$ ; also note that these characters are equal to those of  $D^{(4)}$  minus those of  $D^{(1)}$ .

### B. A Change in Basis

Now let us convert to a new basis that is a linear combination of the original  $S_1, S_2, S_3$  basis:

$$T_1 = S_1 + S_2 + S_3$$

$$T_2 = 2S_1 - S_2 - S_3$$

$$T_3 = S_2 - S_3$$

(Don't worry about how we construct  $T_1$ ,  $T_2$ , and  $T_3$  yet. As will be demonstrated later, we form them by using symmetry projection operators defined below) We determine how the "T" basis functions behave under the group operations by allowing the operations to act on the  $S_j$  and interpreting the results in terms of the  $T_i$ . In particular,

$$\begin{array}{l}
 (T_1, T_2, T_3) \xrightarrow{v} (T_1, T_2, -T_3); \quad (T_1, T_2, T_3) \xrightarrow{E} (T_1, T_2, T_3); \\
 (T_1, T_2, T_3) \xrightarrow{v'} (S_3 + S_2 + S_1, 2S_3 - S_2 - S_1, S_2 - S_1) = (T_1, -\frac{1}{2}T_2 - \frac{3}{2}T_3, -\frac{1}{2}T_2 + \frac{1}{2}T_3); \\
 (T_1, T_2, T_3) \xrightarrow{v''} (S_2 + S_1 + S_3, 2S_2 - S_1 - S_3, S_1 - S_3) = (T_1, -\frac{1}{2}T_2 + \frac{3}{2}T_3, \frac{1}{2}T_2 + \frac{1}{2}T_3); \\
 (T_1, T_2, T_3) \xrightarrow{C_3} (S_3 + S_1 + S_2, 2S_3 - S_1 - S_2, S_1 - S_2) = (T_1, -\frac{1}{2}T_2 - \frac{3}{2}T_3, \frac{1}{2}T_2 - \frac{1}{2}T_3); \\
 (T_1, T_2, T_3) \xrightarrow{C_3^2} (S_2 + S_3 + S_1, 2S_2 - S_3 - S_1, S_3 - S_1) = (T_1, -\frac{1}{2}T_2 + \frac{3}{2}T_3, -\frac{1}{2}T_2 - \frac{1}{2}T_3).
 \end{array}$$

So the matrix representations in the new  $T_i$  basis are:

$$\begin{aligned}
D^{(3)}(E) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} ; D^{(3)}(C_3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{3}{2} \\ 0 & +\frac{1}{2} & -\frac{1}{2} \end{pmatrix} ; \\
D^{(3)}(C_3^2) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & +\frac{3}{2} \\ 0 & +\frac{1}{2} & -\frac{1}{2} \end{pmatrix} ; D^{(3)}(v) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} ; \\
D^{(3)}(v') &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & -\frac{3}{2} \\ 0 & +\frac{1}{2} & -\frac{1}{2} \end{pmatrix} ; D^{(3)}(v'') = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{2} & +\frac{3}{2} \\ 0 & +\frac{1}{2} & +\frac{1}{2} \end{pmatrix} .
\end{aligned}$$

### C. Reduction of the Reducible Representation

These six matrices can be verified to multiply just as the symmetry operations do; thus they form another three-dimensional representation of the group. We see that in the  $T_i$  basis the matrices are block diagonal. This means that the space spanned by the  $T_i$  functions, which is the same space as the  $S_j$  span, forms a reducible representation that can be decomposed into a one dimensional space and a two dimensional space (via formation of the  $T_i$  functions). Note that the characters (traces) of the matrices are not changed by the change in bases.

The one-dimensional part of the above reducible three-dimensional representation is seen to be the same as the totally symmetric representation we arrived at before,  $D^{(1)}$ . The two-dimensional representation that is left can be shown to be irreducible; it has the following matrix representations:

$$\begin{aligned}
D^{(2)}(E) &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} ; D^{(2)}(C_3) = \begin{pmatrix} -\frac{1}{2} & -\frac{3}{2} \\ +\frac{1}{2} & -\frac{1}{2} \end{pmatrix} ; D^{(2)}(C_3^2) = \begin{pmatrix} -\frac{1}{2} & +\frac{3}{2} \\ -\frac{1}{2} & -\frac{1}{2} \end{pmatrix} ; \\
D^{(2)}(v) &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} ; D^{(2)}(v') = \begin{pmatrix} -\frac{1}{2} & -\frac{3}{2} \\ -\frac{1}{2} & +\frac{1}{2} \end{pmatrix} ; D^{(2)}(v'') = \begin{pmatrix} -\frac{1}{2} & +\frac{3}{2} \\ +\frac{1}{2} & +\frac{1}{2} \end{pmatrix} .
\end{aligned}$$

The characters can be obtained by summing diagonal elements:

$$(E) = 2, \quad (2C_3) = -1, \quad (3v) = 0.$$

### D. Rotations as a Basis

Another one-dimensional representation of the group can be obtained by taking rotation about the Z-axis (the  $C_3$  axis) as the object on which the symmetry operations act:

$$\begin{array}{lll} \text{E} & & \\ \text{R}_z & \text{R}_z; & \text{R}_z \text{C}_3 \text{R}_z; \text{R}_z \text{C}_3^2 \text{R}_z'; \\ \text{R}_z \text{v} & -\text{R}_z; & \text{R}_z \text{v}'' -\text{R}_z; \quad \text{R}_z \text{v}' -\text{R}_z. \end{array}$$

In writing these relations, we use the fact that reflection reverses the sense of a rotation. The matrix representations corresponding to this one-dimensional basis are:

$$\begin{array}{lll} \text{D}^{(1)}(\text{E}) = 1; & \text{D}^{(1)}(\text{C}_3) = 1; & \text{D}^{(1)}(\text{C}_3^2) = 1; \\ \text{D}^{(1)}(\text{v}) = -1; & \text{D}^{(1)}(\text{v}'') = -1; & \text{D}^{(1)}(\text{v}') = -1. \end{array}$$

These one-dimensional matrices can be shown to multiply together just like the symmetry operations of the  $C_{3v}$  group. They form an irreducible representation of the group (because it is one-dimensional, it can not be further reduced). Note that this one-dimensional representation is not identical to that found above for the 1s N-atom orbital, or the  $T_1$  function.

#### E. Overview

We have found three distinct irreducible representations for the  $C_{3v}$  symmetry group; two different one-dimensional and one two dimensional representations. Are there any more? An important theorem of group theory shows that the number of irreducible representations of a group is equal to the number of classes. Since there are three classes of operation, we have found all the irreducible representations of the  $C_{3v}$  point group. There are no more.

The irreducible representations have standard names the first  $D^{(1)}$  (that arising from the  $T_1$  and  $1s_N$  orbitals) is called  $A_1$ , the  $D^{(1)}$  arising from  $R_z$  is called  $A_2$  and  $D^{(2)}$  is called E (not to be confused with the identity operation E).

Thus, our original  $D^{(4)}$  representation was a combination of two  $A_1$  representations and one E representation. We say that  $D^{(4)}$  is a direct sum representation:  $D^{(4)} = 2A_1 + E$ . A consequence is that the characters of the combination representation  $D^{(4)}$  can be obtained by adding the characters of its constituent irreducible representations.

		E	$2C_3$	$3\sigma_v$
$A_1$		1	1	1
$A_1$		1	1	1
E		2	-1	0
$2A_1 + E$		4	1	2

#### F. How to Decompose Reducible Representations in General

Suppose you were given only the characters (4,1,2). How can you find out how many times  $A_1$ , E, and  $A_2$  appear when you reduce  $D^{(4)}$  to its irreducible parts? You want to find a linear combination of the characters of  $A_1$ ,  $A_2$  and E that add up (4,1,2). You can treat the characters of matrices as vectors and take the dot product of  $A_1$  with  $D^{(4)}$

$$\begin{array}{cccccc}
 & & & & & 4 & E \\
 & & & & & 1 & C_3 \\
 1 & 1 & 1 & 1 & 1 & 1 & \\
 E & C_3 & & & & 2 & \nu \\
 & & & & & 2 & \\
 & & & & & 2 & 
 \end{array} = 4 + 1 + 1 + 2 + 2 + 2 = 12.$$

The vector (1,1,1,1,1,1) is not normalized; hence to obtain the component of (4,1,1,2,2,2) along a unit vector in the (1,1,1,1,1,1) direction, one must divide by the norm of

(1,1,1,1,1,1); this norm is 6. The result is that the reducible representation contains  $\frac{12}{6} = 2$

$A_1$  components. Analogous projections in the  $E$  and  $A_2$  directions give components of 1 and 0, respectively. In general, to determine the number  $n$  of times irreducible

representation appears in the reducible representation with characters  $\chi_{red}$ , one calculates

$$n = \frac{1}{g} \sum_R \chi(R) \chi_{red}(R),$$

where  $g$  is the order of the group and  $\chi(R)$  are the characters of the  $i$ th irreducible representation.

### G. Commonly Used Bases

We could take any set of functions as a basis for a group representation. Commonly used sets include: coordinates (x,y,z) located on the atoms of a polyatomic molecule (their symmetry treatment is equivalent to that involved in treating a set of p orbitals on the same atoms), quadratic functions such as d orbitals - xy,yz,xz,x<sup>2</sup>-y<sup>2</sup>,z<sup>2</sup>, as well as rotations about the x, y and z axes. The transformation properties of these very commonly used bases are listed in the character tables shown at the end of this appendix.

### H. Summary

The basic idea of symmetry analysis is that any basis of orbitals, displacements, rotations, etc. transforms either as one of the irreducible representations or as a direct sum (reducible) representation. Symmetry tools are used to first determine how the basis transforms under action of the symmetry operations. They are then used to decompose the resultant representations into their irreducible components.

## III. Another Example

### A. The 2p Orbitals of Nitrogen

For a function to transform according to a specific irreducible representation means that the function, when operated upon by a point-group symmetry operator, yields a linear combination of the functions that transform according to that irreducible representation. For example, a 2p<sub>z</sub> orbital (z is the C<sub>3</sub> axis of NH<sub>3</sub>) on the nitrogen atom belongs to the A<sub>1</sub> representation because it yields unity times itself when C<sub>3</sub>, C<sub>3</sub><sup>2</sup>,  $\nu$ ,  $\nu'$ ,  $\nu''$  or the identity operation act on it. The factor of 1 means that 2p<sub>z</sub> has A<sub>1</sub> symmetry



are the projections along each orbital of the effect of the symmetry operation acting on that orbital. For example, a diagonal element of the  $C_3$  matrix is the component of  $C_3 2p_y$  along the  $2p_y$  direction. More rigorously, it

is  $\int 2p_y^* C_3 2p_y d\tau$ . Thus, the character of the  $C_3$  matrix is the sum of  $\int 2p_x^* C_3 2p_x d\tau$  and  $\int 2p_y^* C_3 2p_y d\tau$ . In general, the character of any symmetry operation  $S$  can be computed by allowing  $S$  to operate on each orbital  $\psi_i$ , then projecting  $S \psi_i$  along  $\psi_i$  (i.e., forming  $\int \psi_i^* S \psi_i d\tau$ ), and summing these terms,

$$\chi(S) = \sum_i \int \psi_i^* S \psi_i d\tau = \text{Tr}(S).$$

If these rules are applied to the  $2p_x$  and  $2p_y$  orbitals of nitrogen within the  $C_{3v}$  point group, one obtains

$$\chi(E) = 2, \quad \chi(C_3) = \chi(C_3^2) = -1, \quad \chi(\sigma_v) = \chi(\sigma_v'') = \chi(\sigma_v') = 0.$$

This set of characters is the same as  $D^{(2)}$  above and agrees with those of the  $E$  representation for the  $C_{3v}$  point group. Hence,  $2p_x$  and  $2p_y$  belong to or transform as the  $E$  representation. This is why  $(x,y)$  is to the right of the row of characters for the  $E$  representation in the  $C_{3v}$  character table. In similar fashion, the  $C_{3v}$  character table states that  $d_{x^2-y^2}$  and  $d_{xy}$  orbitals on nitrogen transform as  $E$ , as do  $d_{xy}$  and  $d_{yz}$ , but  $d_z^2$  transforms as  $A_1$ .

Earlier, we considered in some detail how the three  $1s_H$  orbitals on the hydrogen atoms transform. Repeating this analysis using the short-cut rule just described, the traces (characters) of the  $3 \times 3$  representation matrices are computed by allowing  $E$ ,  $2C_3$ , and  $3\sigma_v$  to operate on  $1s_{H1}$ ,  $1s_{H2}$ , and  $1s_{H3}$  and then computing the component of the resulting function along the original function. The resulting characters are  $\chi(E) = 3$ ,  $\chi(C_3) = \chi(C_3^2) = 0$ , and  $\chi(\sigma_v) = \chi(\sigma_v') = \chi(\sigma_v'') = 1$ , in agreement with what we calculated before.

Using the orthogonality of characters taken as vectors we can reduce the above set of characters to  $A_1 + E$ . Hence, we say that our orbital set of three  $1s_H$  orbitals forms a reducible representation consisting of the sum of  $A_1$  and  $E$  IR's. This means that the three  $1s_H$  orbitals can be combined to yield one orbital of  $A_1$  symmetry and a pair that transform according to the  $E$  representation.

#### IV. Projector Operators: Symmetry Adapted Linear Combinations of Atomic Orbitals

To generate the above  $A_1$  and  $E$  symmetry-adapted orbitals, we make use of so-called symmetry projection operators  $P_E$  and  $P_{A_1}$ . These operators are given in terms of linear combinations of products of characters times elementary symmetry operations as follows:

$$P_{A_1} = \frac{1}{h} \sum_S \chi_A(S) S$$

$$P_E = \frac{1}{S} \sum_E \chi_E(S) S$$

where S ranges over  $C_3, C_3^2, \sigma_v, \sigma_v'$  and  $\sigma_v''$  and the identity operation. The result of applying  $P_{A_1}$  to say  $1s_{H_1}$  is

$$\begin{aligned} P_{A_1} 1s_{H_1} &= 1s_{H_1} + 1s_{H_2} + 1s_{H_3} + 1s_{H_2} + 1s_{H_3} + 1s_{H_1} \\ &= 2(1s_{H_1} + 1s_{H_2} + 1s_{H_3}) = A_1, \end{aligned}$$

which is an (unnormalized) orbital having  $A_1$  symmetry. Clearly, this same  $A_1$  would be generated by  $P_{A_1}$  acting on  $1s_{H_2}$  or  $1s_{H_3}$ . Hence, only one  $A_1$  orbital exists.

Likewise,

$$P_E 1s_{H_1} = \frac{2}{3} (1s_{H_1} - 1s_{H_2} - 1s_{H_3}) = E_{1,1}$$

which is one of the symmetry adapted orbitals having E symmetry. The other E orbital can be obtained by allowing  $P_E$  to act on  $1s_{H_2}$  or  $1s_{H_3}$ :

$$P_E 1s_{H_2} = \frac{2}{3} (1s_{H_2} - 1s_{H_1} - 1s_{H_3}) = E_{2,2}$$

$$P_E 1s_{H_3} = \frac{2}{3} (1s_{H_3} - 1s_{H_1} - 1s_{H_2}) = E_{3,3}$$

It might seem as though three orbitals having E symmetry were generated, but only two of these are really independent functions. For example,  $E_{3,3}$  is related to  $E_{1,1}$  and  $E_{2,2}$  as follows:

$$E_{3,3} = -(E_{1,1} + E_{2,2})$$

Thus, only  $E_{1,1}$  and  $E_{2,2}$  are needed to span the two-dimensional space of the E representation. If we include  $E_{1,1}$  in our set of orbitals and require our orbitals to be orthogonal, then we must find numbers a and b such that  $\psi_E = a \psi_{E,2} + b \psi_{E,3}$  is

orthogonal to  $E_{1,1}$ :  $\int \psi_E \psi_{E,1} d\tau = 0$ . A straightforward calculation gives  $a = -b$  or  $\psi_E = a(1s_{H_2} - 1s_{H_3})$  which agrees with what we used earlier to construct the  $T_i$  functions in terms of the  $S_j$  functions.

## V. Summary

Let us now summarize what we have learned. Any given set of atomic orbitals  $\{\psi_i\}$ , atom-centered displacements or rotations can be used as a basis for the symmetry operations of the point group of the molecule. The characters  $\chi(S)$  belonging to the operations S of this point group within any such space can be found by summing the integrals  $\int \psi_i^* S \psi_i d\tau$  over all the atomic orbitals (or corresponding unit vector atomic displacements). The resultant characters will, in general, be reducible to a combination of the characters of the irreducible representations  $\chi_i(S)$ . To decompose the characters  $\chi(S)$  of the reducible representation to a sum of characters  $\chi_i(S)$  of the irreducible representation

$$\chi(S) = \sum_i n_i \chi_i(S),$$

it is necessary to determine how many times,  $n_i$ , the  $i$ -th irreducible representation occurs in the reducible representation. The expression for  $n_i$  is (see the text by Cotton)

$$n_i = \frac{1}{g} \sum_S \chi(S) \chi_i(S)$$

in which  $g$  is the order of the point group; the total number of symmetry operations in the group (e.g.,  $g = 6$  for  $C_{3v}$ ).

For example, the reducible representation  $\chi(E) = 3$ ,  $\chi(C_3) = 0$ , and  $\chi(\sigma_v) = 1$  formed by the three  $1s_H$  orbitals discussed above can be decomposed as follows:

$$n_{A_1} = \frac{1}{6} (3 + 1 + 2 + 0 + 1 + 3 + 1 + 1) = 1,$$

$$n_{A_2} = \frac{1}{6} (3 + 1 + 2 + 0 + 1 + 3 + 1 + (-1)) = 0,$$

$$n_E = \frac{1}{6} (3 + 2 + 2 + 0 + (-1) + 3 + 1 + 0) = 1.$$

These equations state that the three  $1s_H$  orbitals can be combined to give one  $A_1$  orbital and, since  $E$  is degenerate, one pair of  $E$  orbitals, as established above. With knowledge of the  $n_i$ , the symmetry-adapted orbitals can be formed by allowing the projectors

$$P_i = \sum_i \chi_i(S) S$$

to operate on each of the primitive atomic orbitals. How this is carried out was illustrated for the  $1s_H$  orbitals in our earlier discussion. These tools allow a symmetry decomposition of any set of atomic orbitals into appropriate symmetry-adapted orbitals.

Before considering other concepts and group-theoretical machinery, it should once again be stressed that these same tools can be used in symmetry analysis of the translational, vibrational and rotational motions of a molecule. The twelve motions of  $NH_3$  (three translations, three rotations, six vibrations) can be described in terms of combinations of displacements of each of the four atoms in each of three ( $x, y, z$ ) directions. Hence, unit vectors placed on each atom directed in the  $x$ ,  $y$ , and  $z$  directions form a basis for action by the operations  $\{S\}$  of the point group. In the case of  $NH_3$ , the characters of the resultant  $12 \times 12$  representation matrices form a reducible representation in the  $C_{2v}$  point group:  $\chi(E) = 12$ ,  $\chi(C_3) = \chi(C_3^2) = 0$ ,  $\chi(\sigma_v) = \chi(\sigma_v') = \chi(\sigma_v'') = 2$ .

(You should try to prove this. For example under  $\sigma_v$ , the  $H_2$  and  $H_3$  atoms are interchanged, so unit vectors on either one will not contribute to the trace. Unit  $z$ -vectors on  $N$  and  $H_1$  remain unchanged as well as the corresponding  $y$ -vectors. However, the  $x$ -vectors on  $N$  and  $H_1$  are reversed in sign. The total character for  $\sigma_v'$  the  $H_2$  and  $H_3$  atoms are interchanged, so unit vectors on either one will not contribute to the trace. Unit  $z$ -vectors on  $N$  and  $H_1$  remain unchanged as well as the corresponding  $y$ -vectors.

However, the  $x$ -vectors on  $N$  and  $H_1$  are reversed in sign. The total character for  $\sigma_v''$  is thus  $4 - 2 = 2$ . This representation can be decomposed as follows:

$$n_{A_1} = \frac{1}{6} [1 + 1 + 12 + 2 + 1 + 0 + 3 + 1 + 2] = 3,$$

$$n_{A_2} = \frac{1}{6} [1 + 1 + 12 + 2 + 1 + 0 + 3 + (-1) + 2] = 1,$$

$$n_E = \frac{1}{6} [1 + 2 + 12 + 2 + (-1) + 0 + 3 + 0 + 2] = 4.$$

From the information on the right side of the  $C_{3v}$  character table, translations of all four atoms in the z, x and y directions transform as  $A_1(z)$  and  $E(x,y)$ , respectively, whereas rotations about the z( $R_z$ ), x( $R_x$ ), and y( $R_y$ ) axes transform as  $A_2$  and E. Hence, of the twelve motions, three translations have  $A_1$  and E symmetry and three rotations have  $A_2$  and E symmetry. This leaves six vibrations, of which two have  $A_1$  symmetry, none have  $A_2$  symmetry, and two (pairs) have E symmetry. We could obtain symmetry-adapted vibrational and rotational bases by allowing symmetry projection operators of the irreducible representation symmetries to operate on various elementary cartesian (x,y,z) atomic displacement vectors. Both Cotton and Wilson, Decius and Cross show in detail how this is accomplished.

## VI. Direct Product Representations

### A. Direct Products in N-Electron Wavefunctions

We now return to the symmetry analysis of orbital products. Such knowledge is important because one is routinely faced with constructing symmetry-adapted N-electron configurations that consist of products of N individual orbitals. A point-group symmetry operator S, when acting on such a product of orbitals, gives the product of S acting on each of the individual orbitals

$$S(\psi_1 \psi_2 \psi_3 \dots \psi_N) = (S \psi_1) (S \psi_2) (S \psi_3) \dots (S \psi_N).$$

For example, reflection of an N-orbital product through the  $\sigma_v$  plane in  $NH_3$  applies the reflection operation to all N electrons.

Just as the individual orbitals formed a basis for action of the point-group operators, the configurations (N-orbital products) form a basis for the action of these same point-group operators. Hence, the various electronic configurations can be treated as functions on which S operates, and the machinery illustrated earlier for decomposing orbital symmetry can then be used to carry out a symmetry analysis of configurations.

Another shortcut makes this task easier. Since the symmetry adapted individual orbitals  $\{\psi_i, i = 1, \dots, M\}$  transform according to irreducible representations, the representation matrices for the N-term products shown above consist of products of the matrices belonging to each  $\psi_i$ . This matrix product is not a simple product but a direct product. To compute the characters of the direct product matrices, one multiplies the characters of the individual matrices of the irreducible representations of the N orbitals that appear in the electron configuration. The direct-product representation formed by the orbital products can therefore be symmetry-analyzed (reduced) using the same tools as we used earlier.

For example, if one is interested in knowing the symmetry of an orbital product of the form  $a_1^2 a_2^2 e^2$  (note: lower case letters are used to denote the symmetry of electronic orbitals) in  $C_{3v}$  symmetry, the following procedure is used. For each of the six symmetry operations in the  $C_{2v}$  point group, the product of the characters associated with each of the six spin orbitals (orbital multiplied by  $\uparrow$  or  $\downarrow$  spin) is formed

$$\chi_i(S) = (\chi_{A_1}(S))^2 (\chi_{A_2}(S))^2 (\chi_E(S))^2.$$

In the specific case considered here,  $\chi(E) = 4$ ,  $\chi(2C_3) = 1$ , and  $\chi(3C_2) = 0$  (You should try this.). Notice that the contributions of any doubly occupied nondegenerate orbitals (e.g.,  $a_1^2$ , and  $a_2^2$ ) to these direct product characters  $\chi(S)$  are unity because for all operators  $(\chi_k(S))^2 = 1$  for any one-dimensional irreducible representation. As a result, only the singly occupied or degenerate orbitals need to be considered when forming the characters of the reducible direct-product representation  $\chi(S)$ . For this example this means that the direct-product characters can be determined from the characters  $\chi_E(S)$  of the two active (i.e., nonclosed-shell) orbitals - the  $e^2$  orbitals. That is,  $\chi(S) = \chi_E(S) \chi_E(S)$ .

From the direct-product characters  $\chi(S)$  belonging to a particular electronic configuration (e.g.,  $a_1^2 a_2^2 e^2$ ), one must still decompose this list of characters into a sum of irreducible characters. For the example at hand, the direct-product characters  $\chi(S)$  decompose into one  $A_1$ , one  $A_2$ , and one  $E$  representation. This means that the  $e^2$  configuration contains  $A_1$ ,  $A_2$ , and  $E$  symmetry elements. Projection operators analogous to those introduced earlier for orbitals can be used to form symmetry-adapted orbital products from the individual basis orbital products of the form  $a_1^2 a_2^2 e_x^m e_y^{m'}$ , where  $m$  and  $m'$  denote the occupation (1 or 0) of the two degenerate orbitals  $e_x$  and  $e_y$ . When dealing with indistinguishable particles such as electrons, it is also necessary to further project the resulting orbital products to make them antisymmetric (for Fermions) or symmetric (for Bosons) with respect to interchange of any pair of particles. This step reduces the set of  $N$ -electron states that can arise. For example, in the above  $e^2$  configuration case, only  $^3A_2$ ,  $^1A_1$ , and  $^1E$  states arise; the  $^3E$ ,  $^3A_1$ , and  $^1A_2$  possibilities disappear when the antisymmetry projector is applied. In contrast, for an  $e^1 e^1$  configuration, all states arise even after the wavefunction has been made antisymmetric. The steps involved in combining the point group symmetry with permutational antisymmetry are illustrated in Chapter 10. In Appendix III of Electronic Spectra and Electronic Structure of Polyatomic Molecules, G. Herzberg, Van Nostrand Reinhold Co., New York, N.Y. (1966) the resolution of direct products among various representations within many point groups are tabulated.

## B. Direct Products in Selection Rules

Two states  $\psi_a$  and  $\psi_b$  that are eigenfunctions of a Hamiltonian  $H_0$  in the absence of some external perturbation (e.g., electromagnetic field or static electric field or potential due to surrounding ligands) can be "coupled" by the perturbation  $V$  only if the symmetries of  $V$  and of the two wavefunctions obey a so-called selection rule. In particular, only if the coupling integral (see Appendix D which deals with time independent perturbation theory)

$$\int \psi_a^* V \psi_b d\tau = V_{a,b}$$

is non-vanishing will the two states be coupled by  $V$ .

The role of symmetry in determining whether such integrals are non-zero can be demonstrated by noting that the integrand, considered as a whole, must contain a component that is invariant under all of the group operations (i.e., belongs to the totally

symmetric representation of the group). In terms of the projectors introduced above in Sec. IV, of this Appendix we must have

$$A(S) S_a^* V_b$$

not vanish. Here the subscript A denotes the totally symmetric representation of the group. The symmetry of the product  $S_a^* V_b$  is, according to what was covered earlier in this Section, given by the direct product of the symmetries of  $S_a^*$  of  $V$  and of  $b$ . So, the conclusion is that the integral will vanish unless this triple direct product contains, when it is reduced to its irreducible components, a component of the totally symmetric representation.

To see how this result is used, consider the integral that arises in formulating the interaction of electromagnetic radiation with a molecule within the electric-dipole approximation:

$$S_a^* \mathbf{r}_b d .$$

Here  $\mathbf{r}$  is the vector giving, together with  $e$ , the unit charge, the quantum mechanical dipole moment operator

$$\mathbf{r} = e \sum_n Z_n \mathbf{R}_n - e \sum_j \mathbf{r}_j ,$$

where  $Z_n$  and  $\mathbf{R}_n$  are the charge and position of the  $n^{\text{th}}$  nucleus and  $\mathbf{r}_j$  is the position of the  $j^{\text{th}}$  electron. Now, consider evaluating this integral for the singlet  $n \rightarrow n^*$  transition in formaldehyde. Here, the closed-shell ground state is of  $^1A_1$  symmetry and the excited state, which involves promoting an electron from the non-bonding  $b_2$  lone pair orbital on the Oxygen into the  $a_2$  orbital on the CO moiety, is of  $^1A_2$  symmetry ( $b_1 \times b_2 = a_2$ ). The direct product of the two wavefunction symmetries thus contains only  $a_2$  symmetry. The three components (x, y, and z) of the dipole operator have, respectively,  $b_1$ ,  $b_2$ , and  $a_1$  symmetry. Thus, the triple direct products give rise to the following possibilities:

$$a_2 \times b_1 = b_2,$$

$$a_2 \times b_2 = b_1,$$

$$a_2 \times a_1 = a_2 .$$

There is no component of  $a_1$  symmetry in the triple direct product, so the integral vanishes. This allows us to conclude that the  $n \rightarrow n^*$  excitation in formaldehyde is electric dipole forbidden.

## VII. Overview

This appendix has reviewed how to make a symmetry decomposition of a basis of atomic orbitals (or cartesian displacements or orbital products) into irreducible representation components. This tool is most helpful when constructing the orbital correlation diagrams that form the basis of the Woodward-Hoffmann rules. We also learned how to form the direct-product symmetries that arise when considering configurations consisting of products of symmetry-adapted spin orbitals. This step is essential for the construction of configuration and state correlation diagrams upon which one ultimately bases a prediction about whether a reaction is allowed or forbidden. Finally, we learned how the direct product analysis allows one to determine whether or not integrals of products of wave functions with operators between them vanish. This tool is of utmost importance in determining selection rules in spectroscopy and for

determining the effects of external perturbations on the states of the species under investigation.

## Character Tables

$C_1$	E
A	1

$C_s$	E	h		
A'	1	1	x,y,R <sub>z</sub>	x <sup>2</sup> ,y <sup>2</sup> ,z <sup>2</sup> ,xy
A''	1	-1	z,R <sub>x</sub> ,R <sub>y</sub>	yz,xz

$C_i$	E	i		
A <sub>g</sub>	1	1	R <sub>x</sub> ,R <sub>y</sub> ,R <sub>z</sub>	x <sup>2</sup> ,y <sup>2</sup> ,z <sup>2</sup> ,xy,xz,yz
A <sub>u</sub>	1	-1	x,y,z	

$C_2$	E	$C_2$		
A	1	1	z,R <sub>z</sub>	x <sup>2</sup> ,y <sup>2</sup> ,z <sup>2</sup> ,xy
B	1	-1	x,y,R <sub>x</sub> ,R <sub>y</sub>	yz,xz

$D_2$	E	$C_2(z)$	$C_2(y)$	$C_2(x)$		
A	1	1	1	1		x <sup>2</sup> ,y <sup>2</sup> ,z <sup>2</sup>
B <sub>1</sub>	1	1	-1	-1	z,R <sub>z</sub>	xy
B <sub>2</sub>	1	-1	1	-1	y,R <sub>y</sub>	xz
B <sub>3</sub>	1	-1	-1	1	x,R <sub>x</sub>	yz

$D_3$	E	$2C_3$	$3C_2$		
A <sub>1</sub>	1	1	1		x <sup>2</sup> +y <sup>2</sup> ,z <sup>2</sup>
A <sub>2</sub>	1	1	-1	z,R <sub>z</sub>	
E	2	-1	0	(x,y) (R <sub>x</sub> ,R <sub>y</sub> )	(x <sup>2</sup> -y <sup>2</sup> ,xy) (xz,yz)

D <sub>4</sub>	E	2C <sub>4</sub>	C <sub>2</sub> (=C <sub>4</sub> <sup>2</sup> )	2C <sub>2</sub> '	2C <sub>2</sub> "		
A <sub>1</sub>	1	1	1	1	1	z, R <sub>Z</sub>	x <sup>2</sup> +y <sup>2</sup> , z <sup>2</sup>
A <sub>2</sub>	1	1	1	-1	-1		x <sup>2</sup> -y <sup>2</sup>
B <sub>1</sub>	1	-1	1	1	-1		xy
B <sub>2</sub>	1	-1	1	-1	1		(xz, yz)
E	2	0	-2	0	0		(x, y) (R <sub>x</sub> , R <sub>y</sub> )

C <sub>2v</sub>	E	C <sub>2</sub>	v(xz)	v'(yz)		
A <sub>1</sub>	1	1	1	1	z	x <sup>2</sup> , y <sup>2</sup> , z <sup>2</sup>
A <sub>2</sub>	1	1	-1	-1	R <sub>Z</sub>	xy
B <sub>1</sub>	1	-1	1	-1	x, R <sub>y</sub>	xz
B <sub>2</sub>	1	-1	-1	1	y, R <sub>x</sub>	yz

C <sub>3v</sub>	E	2C <sub>3</sub>	3 v		
A <sub>1</sub>	1	1	1	z	x <sup>2</sup> +y <sup>2</sup> , z <sup>2</sup>
A <sub>2</sub>	1	1	-1	R <sub>Z</sub>	
E	2	-1	0	(x, y) (R <sub>x</sub> , R <sub>y</sub> )	(x <sup>2</sup> -y <sup>2</sup> , xy) (xz, yz)

C <sub>4v</sub>	E	2C <sub>4</sub>	C <sub>2</sub>	2 v	2 d		
A <sub>1</sub>	1	1	1	1	1	z	x <sup>2</sup> +y <sup>2</sup> , z <sup>2</sup>
A <sub>2</sub>	1	1	1	-1	-1	R <sub>Z</sub>	
B <sub>1</sub>	1	-1	1	1	-1		x <sup>2</sup> -y <sup>2</sup>
B <sub>2</sub>	1	-1	1	-1	1		xy
E	2	0	-2	0	0	(x, y) (R <sub>x</sub> , R <sub>y</sub> )	(xz, yz)

C <sub>2h</sub>	E	C <sub>2</sub>	i	h		
A <sub>g</sub>	1	1	1	1	R <sub>Z</sub>	x <sup>2</sup> , y <sup>2</sup> , z <sup>2</sup> , xy
B <sub>g</sub>	1	-1	1	-1	R <sub>x</sub> , R <sub>y</sub>	xz, yz
A <sub>u</sub>	1	1	-1	-1	z	
B <sub>u</sub>	1	-1	-1	1	x, y	

D <sub>2h</sub>	E	C <sub>2</sub> (z)	C <sub>2</sub> (y)	C <sub>2</sub> (x)	i	(xy)	(xz)	(yz)		
A <sub>g</sub>	1	1	1	1	1	1	1	1		x <sup>2</sup> ,y <sup>2</sup> ,z <sup>2</sup>
B <sub>1g</sub>	1	1	-1	-1	1	1	-1	-1	R <sub>z</sub>	xy
B <sub>2g</sub>	1	-1	1	-1	1	-1	1	-1	R <sub>y</sub>	xz
B <sub>3g</sub>	1	-1	-1	1	1	-1	-1	1	R <sub>x</sub>	yz
A <sub>u</sub>	1	1	1	1	-1	-1	-1	-1		
B <sub>1u</sub>	1	1	-1	-1	-1	-1	1	1	z	
B <sub>2u</sub>	1	-1	1	-1	-1	1	-1	1	y	
B <sub>3u</sub>	1	-1	-1	1	-1	1	1	-1	x	

D <sub>3h</sub>	E	2C <sub>3</sub>	3C <sub>2</sub>	h	2S <sub>3</sub>	3 v		
A <sub>1</sub> '	1	1	1	1	1	1		x <sup>2</sup> +y <sup>2</sup> ,z <sup>2</sup>
A <sub>2</sub> '	1	1	-1	1	1	-1	R <sub>z</sub>	
E'	2	-1	0	2	-1	0	(x,y)	(x <sup>2</sup> -y <sup>2</sup> ,xy)
A <sub>1</sub> "	1	1	1	-1	-1	-1		
A <sub>2</sub> "	1	1	-1	-1	-1	1	z	
E"	2	-1	0	-2	1	0	(R <sub>x</sub> ,R <sub>y</sub> )	(xz,yz)

D <sub>4h</sub>	E	2C <sub>4</sub>	C <sub>2</sub>	2C <sub>2</sub> '	2C <sub>2</sub> "	i	2S <sub>4</sub>	h	2 v	2 d		
A <sub>1g</sub>	1	1	1	1	1	1	1	1	1	1		x <sup>2</sup> +y <sup>2</sup> ,z <sup>2</sup>
A <sub>2g</sub>	1	1	1	-1	-1	1	1	1	-1	-1	R <sub>z</sub>	
B <sub>1g</sub>	1	-1	1	1	-1	1	-1	1	1	-1		x <sup>2</sup> -y <sup>2</sup>
B <sub>2g</sub>	1	-1	1	-1	1	1	-1	1	-1	1		xy
E <sub>g</sub>	2	0	-2	0	0	2	0	-2	0	0	(R <sub>x</sub> ,R <sub>y</sub> )	(xz,yz)
A <sub>1u</sub>	1	1	1	1	1	-1	-1	-1	-1	-1		
A <sub>2u</sub>	1	1	1	-1	-1	-1	-1	-1	1	1	z	
B <sub>1u</sub>	1	-1	1	1	-1	-1	1	-1	-1	1		
B <sub>2u</sub>	1	-1	1	-1	1	-1	1	-1	1	-1		
E <sub>u</sub>	2	0	-2	0	0	-2	0	2	0	0	(x,y)	

D <sub>6h</sub>	E	2C <sub>6</sub>	2C <sub>3</sub>	C <sub>2</sub>	3C <sub>2</sub> '	3C <sub>2</sub> "	i	2S <sub>3</sub>	2S <sub>6</sub>	h	3 d	3 v			
A <sub>1g</sub>	1	1	1	1	1	1	1	1	1	1	1	1		R <sub>Z</sub>	x <sup>2</sup> +y <sup>2</sup> ,z <sup>2</sup>
A <sub>2g</sub>	1	1	1	1	-1	-1	1	1	1	1	-1	-1			
B <sub>1g</sub>	1	-1	1	-1	1	-1	1	-1	1	-1	1	-1			
B <sub>2g</sub>	1	-1	1	-1	-1	1	1	-1	1	-1	-1	1			
E <sub>1g</sub>	2	1	-1	-2	0	0	2	1	-1	-2	0	0	(R <sub>x</sub> ,R <sub>y</sub> )	(xz,yz)	
E <sub>2g</sub>	2	-1	-1	2	0	0	2	-1	-1	2	0	0		(x <sup>2</sup> -y <sup>2</sup> ,xy)	
A <sub>1u</sub>	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1			
A <sub>2u</sub>	1	1	1	1	-1	-1	-1	-1	-1	-1	1	1	z		
B <sub>1u</sub>	1	-1	1	-1	1	-1	-1	1	-1	1	-1	1			
B <sub>2u</sub>	1	-1	1	-1	-1	1	-1	1	-1	1	1	-1			
E <sub>1u</sub>	2	1	-1	-2	0	0	-2	-1	1	2	0	0	(x,y)		
E <sub>2u</sub>	2	-1	-1	2	0	0	-2	1	1	-2	0	0			

D <sub>2d</sub>	E	2S <sub>4</sub>	C <sub>2</sub>	2C <sub>2</sub> '	2 d		
A <sub>1</sub>	1	1	1	1	1		x <sup>2</sup> +y <sup>2</sup> ,z <sup>2</sup>
A <sub>2</sub>	1	1	1	-1	-1	R <sub>Z</sub>	
B <sub>1</sub>	1	-1	1	1	-1		x <sup>2</sup> -y <sup>2</sup>
B <sub>2</sub>	1	-1	1	-1	1	z	xy
E	2	0	-2	0	0	(x,y) (R <sub>x</sub> ,R <sub>y</sub> )	(xz,yz)

D <sub>3d</sub>	E	2C <sub>3</sub>	3C <sub>2</sub>	i	2S <sub>6</sub>	3 d		
A <sub>1g</sub>	1	1	1	1	1	1		x <sup>2</sup> +y <sup>2</sup> ,z <sup>2</sup>
A <sub>2g</sub>	1	1	-1	1	1	-1	R <sub>Z</sub>	
E <sub>g</sub>	2	-1	0	2	-1	0	(R <sub>x</sub> ,R <sub>y</sub> )	(x <sup>2</sup> -y <sup>2</sup> ,xy) (xz,yz)
A <sub>1u</sub>	1	1	1	-1	-1	-1		
A <sub>2u</sub>	1	1	-1	-1	-1	1	z	
E <sub>u</sub>	2	-1	0	-2	1	0	(x,y)	

S <sub>4</sub>	E	S <sub>4</sub>	C <sub>2</sub>	S <sub>4</sub> <sup>3</sup>		
A	1	1	1	1	R <sub>Z</sub>	x <sup>2</sup> +y <sup>2</sup> ,z <sup>2</sup>
B	1	-1	1	-1	z	x <sup>2</sup> -y <sup>2</sup> ,xy
E	{ 1 -1 }	{ i -i }	{ -1 -1 }	{ -i i }	(x,y) (R <sub>x</sub> ,R <sub>y</sub> )	(xz,yz)

T	E	4C <sub>3</sub>	4C <sub>3</sub> <sup>2</sup>	3C <sub>2</sub>		=exp(2 i/3)
A	1	1	1	1		x <sup>2</sup> +y <sup>2</sup> +z <sup>2</sup>
E	{1 1	*	*	{1 -1}		(2z <sup>2</sup> -x <sup>2</sup> -y <sup>2</sup> , x <sup>2</sup> -y <sup>2</sup> )
T	3	0	0	-1	(R <sub>x</sub> , R <sub>y</sub> , R <sub>z</sub> ) (x, y, z)	(xy, xz, yz)

T <sub>h</sub>	E	4C <sub>3</sub>	4C <sub>3</sub> <sup>2</sup>	3C <sub>2</sub>	i	4S <sub>6</sub>	4S <sub>6</sub> <sup>5</sup>	3 h		=exp(2 i/3)
A <sub>g</sub>	1	1	1	1	1	1	1	1		x <sup>2</sup> +y <sup>2</sup> +z <sup>2</sup>
A <sub>u</sub>	1	1	1	1	-1	-1	-1	-1		
E <sub>g</sub>	{1 1	*	*	{1 1	1	*	*	1		(2z <sup>2</sup> -x <sup>2</sup> -y <sup>2</sup> , x <sup>2</sup> -y <sup>2</sup> )
E <sub>u</sub>	{1 1	*	*	{1 1	-1	-	-	{-1 -1}		
T <sub>g</sub>	3	0	0	-1	1	0	0	-1	(R <sub>x</sub> , R <sub>y</sub> , R <sub>z</sub> )	
T <sub>u</sub>	3	0	0	-1	-1	0	0	1	(x, y, z)	(xy, xz, yz)

T <sub>d</sub>	E	8C <sub>3</sub>	3C <sub>2</sub>	6S <sub>4</sub>	6 d		
A <sub>1</sub>	1	1	1	1	1		x <sup>2</sup> +y <sup>2</sup> +z <sup>2</sup>
A <sub>2</sub>	1	1	1	-1	-1		
E	2	-1	2	0	0		(2z <sup>2</sup> -x <sup>2</sup> -y <sup>2</sup> , x <sup>2</sup> -y <sup>2</sup> )
T <sub>1</sub>	3	0	-1	1	-1	(R <sub>x</sub> , R <sub>y</sub> , R <sub>z</sub> )	
T <sub>2</sub>	3	0	-1	-1	1	(x, y, z)	(xy, xz, yz)

O	E	6C <sub>4</sub>	3C <sub>2</sub> (=C <sub>4</sub> <sup>2</sup> )	8C <sub>3</sub>	6C <sub>2</sub>		
A <sub>1</sub>	1	1	1	1	1		x <sup>2</sup> +y <sup>2</sup> +z <sup>2</sup>
A <sub>2</sub>	1	-1	1	1	-1		
E	2	0	2	-1	0		(2z <sup>2</sup> -x <sup>2</sup> -y <sup>2</sup> , x <sup>2</sup> -y <sup>2</sup> )
T <sub>1</sub>	3	1	-1	0	-1	(R <sub>x</sub> , R <sub>y</sub> , R <sub>z</sub> ) (x, y, z)	
T <sub>2</sub>	3	-1	-1	0	1		(xy, xz, yz)



## Qualitative Orbital Picture and Semi-Empirical Methods F

Some of the material contained in the early parts of this Appendix appears, in condensed form, near the end of Chapter 7. For the sake of completeness and clarity of presentation, this material is repeated and enhanced here.

In the simplest picture of chemical bonding, the valence molecular orbitals  $\psi_i$  are constructed as linear combinations of valence atomic orbitals  $\chi_\mu$  according to the LCAO-MO formula:

$$\psi_i = \sum_{\mu} C_{i\mu} \chi_{\mu}$$

The core electrons are not explicitly included in such a treatment, although their effects are felt through an electrostatic potential  $V$  as detailed below. The electrons that reside in the occupied MO's are postulated to experience an effective potential  $V$  that has the following properties:

1.  $V$  contains contributions from all of the nuclei in the molecule exerting coulombic attractions on the electron, as well as coulombic repulsions and exchange interactions exerted by the other electrons on this electron;
2. As a result of the (assumed) cancellation of attractions from distant nuclei and repulsions from the electron clouds (i.e., the core, lone-pair, and valence orbitals) that surround these distant nuclei, the effect of  $V$  on any particular MO  $\psi_i$  depends primarily on the atomic charges and local bond polarities of the atoms over which  $\psi_i$  is delocalized.

As a result of these assumptions, qualitative molecular orbital models can be developed in which one assumes that each MO  $\psi_i$  obeys a one-electron Schrödinger equation

$$\hat{h} \psi_i = \epsilon_i \psi_i$$

Here the orbital-level hamiltonian  $\hat{h}$  contains the kinetic energy of motion of the electron and the potential  $V$  mentioned above:

$$\left[ -\frac{\hbar^2}{2m_e} \nabla^2 + V \right] \psi_i = \epsilon_i \psi_i$$

Expanding the MO  $\psi_i$  in the LCAO-MO manner, substituting this expansion into the above Schrödinger equation, multiplying on the left by  $\chi_{\mu}^*$ , and integrating over the coordinates of the electron generates the following orbital-level eigenvalue problem:

$$\sum_{\mu} \left\langle \chi_{\mu} \left| -\frac{\hbar^2}{2m_e} \nabla^2 + V \right| \chi_{\mu} \right\rangle C_{i\mu} = \epsilon_i \sum_{\mu} \langle \chi_{\mu} | \chi_{\mu} \rangle C_{i\mu}$$

If the constituent atomic orbitals  $\{\chi_{\mu}\}$  have been orthonormalized as discussed in Chapter 7, the overlap integrals  $\langle \chi_{\mu} | \chi_{\mu} \rangle$  reduce to  $\delta_{\mu\mu}$ .

In all semi-empirical models, the quantities that explicitly define the potential  $V$  are not computed from first principles as they are in so-called ab initio methods. Rather, either

experimental data or results of ab initio calculations are used to determine the parameters in terms of which  $V$  is expressed. The various semi-empirical methods discussed below differ in the sophistication used to include electron-electron interactions as well as in the manner experimental data or ab initio computational results are used to specify  $V$ .

If experimental data is used to parameterize a semi-empirical model, then the model should not be extended beyond the level at which it has been parameterized. For example, orbitals obtained from a semi-empirical theory parameterized such that bond energies, excitation energies, and ionization energies predicted by theory agree with experimental data should not subsequently be used in a configuration interaction (CI) calculation. To do so would be inconsistent because the CI treatment, which is designed to treat dynamical correlations among the electrons, would duplicate what using the experimental data (which already contains mother nature's electronic correlations) to determine the parameters had accomplished.

Alternatively, if results of ab initio theory at the single-configuration orbital level are used to define the parameters of a semi-empirical model, it would be proper to use the semi-empirical orbitals in a subsequent higher-level treatment of electron correlations.

### A. The Hückel Model

In the most simplified embodiment of the above orbital-level model, the following additional approximations are introduced:

1. The diagonal values  $\langle \mu | -\hbar^2/2m_e \nabla^2 + V | \mu \rangle$ , which are usually denoted  $\epsilon_\mu$ , are taken to be equal to the energy of an electron in the atomic orbital  $\mu$  and, as such, are evaluated in terms of atomic ionization energies (IP's) and electron affinities (EA's):

$$\langle \mu | -\hbar^2/2m_e \nabla^2 + V | \mu \rangle = -IP_\mu,$$

for atomic orbitals that are occupied in the atom, and

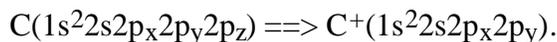
$$\langle \mu | -\hbar^2/2m_e \nabla^2 + V | \mu \rangle = -EA_\mu,$$

for atomic orbitals that are not occupied in the atom.

These approximations assume that contributions in  $V$  arising from coulombic attraction to nuclei other than the one on which  $\mu$  is located, and repulsions from the core, lone-pair, and valence electron clouds surrounding these other nuclei cancel to an extent that

$\langle \mu | V | \mu \rangle$  contains only potentials from the atom on which  $\mu$  sits.

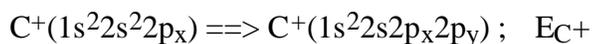
It should be noted that the IP's and EA's of valence-state orbitals are not identical to the experimentally measured IP's and EA's of the corresponding atom, but can be obtained from such information. For example, the 2p valence-state IP (VSIP) for a Carbon atom is the energy difference associated with the hypothetical process



If the energy differences for the "promotion" of C



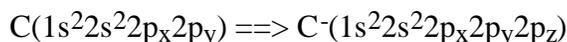
and for the promotion of  $C^+$



are known, the desired VSIP is given by:

$$IP_{2p_z} = IP_C + E_{C^+} - E_C.$$

The EA of the 2p orbital is obtained from the



energy gap, which means that  $EA_{2p_z} = EA_C$ . Some common IP's of valence 2p orbitals in eV are as follows: C (11.16), N (14.12), N<sup>+</sup> (28.71), O (17.70), O<sup>+</sup> (31.42), F<sup>+</sup> (37.28).

2. The off-diagonal elements  $\langle \mu | -\hbar^2/2m_e \nabla^2 + V | \nu \rangle$  are taken as zero if  $\mu$  and  $\nu$  belong to the same atom because the atomic orbitals are assumed to have been constructed to diagonalize the one-electron hamiltonian appropriate to an electron moving in that atom. They are set equal to a parameter denoted  $\mu_{\nu}$ , if  $\mu$  and  $\nu$  reside on neighboring atoms that are chemically bonded. If  $\mu$  and  $\nu$  reside on atoms that are not bonded neighbors, then the off-diagonal matrix element is set equal to zero.

3. The geometry dependence of the  $\mu_{\nu}$  parameters is often approximated by assuming that  $\mu_{\nu}$  is proportional to the overlap  $S_{\mu\nu}$  between the corresponding atomic orbitals:

$$\mu_{\nu} = \mu_0 S_{\mu\nu}.$$

Here  $\mu_0$  is a constant (having energy units) characteristic of the bonding interaction between  $\mu$  and  $\nu$ ; its value is usually determined by forcing the molecular orbital energies obtained from such a qualitative orbital treatment to yield experimentally correct ionization potentials, bond dissociation energies, or electronic transition energies.

The particular approach described thus far forms the basis of the so-called Hückel model. Its implementation requires knowledge of the atomic  $\mu$  and  $\mu_0$  values, which are eventually expressed in terms of experimental data, as well as a means of calculating the geometry dependence of the  $\mu_{\nu}$ 's (e.g., some method for computing overlap matrices  $S_{\mu\nu}$ ).

## B. The Extended Hückel Method

It is well known that bonding and antibonding orbitals are formed when a pair of atomic orbitals from neighboring atoms interact. The energy splitting between the bonding and antibonding orbitals depends on the overlap between the pair of atomic orbitals. Also, the energy of the antibonding orbital lies higher above the arithmetic mean  $E_{ave} = (E_A + E_B)/2$  of the energies of the constituent atomic orbitals ( $E_A$  and  $E_B$ ) than the bonding orbital lies below  $E_{ave}$ . If overlap is ignored, as in conventional Hückel theory (except in

parameterizing the geometry dependence of  $\langle \mu, \nu | H | \mu, \nu \rangle$ , the differential destabilization of antibonding orbitals compared to stabilization of bonding orbitals can not be accounted for.

By parameterizing the off-diagonal Hamiltonian matrix elements in the following overlap-dependent manner:

$$H_{\nu, \mu} = \langle \nu | -\frac{\hbar^2}{2m_e} \nabla^2 + V | \mu \rangle = 0.5 K (h_{\mu, \mu} + h_{\nu, \nu}) S_{\mu, \nu},$$

and explicitly treating the overlaps among the constituent atomic orbitals  $\{ \mu \}$  in solving the orbital-level Schrödinger equation

$$\langle \mu | -\frac{\hbar^2}{2m_e} \nabla^2 + V | \mu \rangle C_{i\mu} = \epsilon_i \langle \mu | \mu \rangle C_{i\mu},$$

Hoffmann introduced the so-called extended Hückel method. He found that a value for  $K=1.75$  gave optimal results when using Slater-type orbitals as a basis (and for calculating the  $S_{\mu, \nu}$ ). The diagonal  $h_{\mu, \mu}$  elements are given, as in the conventional Hückel method, in terms of valence-state IP's and EA's. Cusachs later proposed a variant of this parameterization of the off-diagonal elements:

$$H_{\nu, \mu} = 0.5 K (h_{\mu, \mu} + h_{\nu, \nu}) S_{\mu, \nu} (2 - |S_{\mu, \nu}|).$$

For first- and second-row atoms, the 1s or (2s, 2p) or (3s, 3p, 3d) valence-state ionization energies ( $\epsilon_{\mu}$ 's), the number of valence electrons (#Elec.) as well as the orbital exponents ( $e_s$ ,  $e_p$  and  $e_d$ ) of Slater-type orbitals used to calculate the overlap matrix elements  $S_{\mu, \nu}$ , corresponding are given below.

Atom	# Elec.	$e_s=e_p$	$e_d$	$s(\text{eV})$	$p(\text{eV})$	$d(\text{eV})$
H	1	1.3		-13.6		
Li	1	0.650		-5.4	-3.5	
Be	2	0.975		-10.0	-6.0	
B	3	1.300		-15.2	-8.5	
C	4	1.625		-21.4	-11.4	
N	5	1.950		-26.0	-13.4	
O	6	2.275		-32.3	-14.8	
F	7	2.425		-40.0	-18.1	
Na	1	0.733		-5.1	-3.0	
Mg	2	0.950		-9.0	-4.5	
Al	3	1.167		-12.3	-6.5	
Si	4	1.383	1.383	-17.3	-9.2	-6.0
P	5	1.600	1.400	-18.6	-14.0	-7.0
S	6	1.817	1.500	-20.0	-13.3	-8.0
Cl	7	2.033	2.033	-30.0	-15.0	-9.0

In the Hückel or extended Hückel methods no explicit reference is made to electron-electron interactions although such contributions are absorbed into the V potential, and hence into the  $\mu$  and  $\mu$ , parameters of Hückel theory or the  $h_{\mu,\mu}$  and  $h_{\mu}$ , parameters of extended Hückel theory. As electron density flows from one atom to another (due to electronegativity differences), the electron-electron repulsions in various atomic orbitals changes. To account for such charge-density-dependent coulombic energies, one must use an approach that includes explicit reference to inter-orbital coulomb and exchange interactions. There exists a large family of semi-empirical methods that permit explicit treatment of electronic interactions; some of the more commonly used approaches are discussed below.

### C. Semi-Empirical Models that Treat Electron-Electron Interactions

#### 1. The ZDO Approximation

Most methods of this type are based on the so-called zero-differential overlap (ZDO) approximation. Their development begins by using an approximation to the atomic-orbital-based two-electron integrals introduced by Mulliken:

$$\langle a | b | g | c | d \rangle = S_{a,c} S_{b,d} \{ \langle a, b | + \langle a, d | + \langle c, b | + \langle c, d | \} / 4,$$

where  $S_{a,c}$  is the overlap integral between  $a$  and  $c$ , and

$$\langle a, b | = \langle a | b | g | a | b \rangle$$

is the coulomb integral between the charge densities  $|a|^2$  and  $|b|^2$ . Then, when the so-called zero-overlap approximation

$$S_{a,c} = \delta_{a,c}$$

is made, the general four-orbital two-electron integral given above reduces to its zero-differential overlap value:

$$\langle a | b | g | c | d \rangle = a_{,c} b_{,d} a_{,b}$$

This fundamental approximation allows the two-electron integrals that enter into the expression for the Fock matrix elements to be expressed in terms of the set of two-orbital coulomb interaction integrals  $a_{,b}$  as well as experimental or ab initio values for valence-state IP's and EA's, as is now illustrated.

## 2. Resulting Fock Matrices

Using the ZDO approximation, the Fock matrix elements over the valence atomic orbitals (the cores are still treated through an effective electrostatic potential as above)

$$F_{\mu, \nu} = \langle \mu | h | \nu \rangle + \sum_b [ \delta_{\mu, \nu} \langle \mu | g | \nu \rangle - \delta_{\mu, \nu}^{\text{ex}} \langle \mu | g | \nu \rangle ],$$

reduce, to:

$$F_{\mu, \mu} = \langle \mu | h | \mu \rangle + \sum_b \delta_{\mu, \mu} - \sum_b \delta_{\mu, \mu}^{\text{ex}} \mu_{, \mu},$$

for the diagonal elements and

$$F_{\mu, \nu} = \langle \mu | h | \nu \rangle - \sum_b \delta_{\mu, \nu}^{\text{ex}} \mu_{, \nu},$$

for the off-diagonal elements. Here,  $h$  represents the kinetic energy  $-\hbar^2 \nabla^2 / 2m$  operator plus the sum of the attractive coulombic potential energies to each of the nuclei  $-\sum_a Z_a e^2 / |r - R_a|$  and the electrostatic repulsions of the core electrons (i.e., all those not explicitly treated as valence in this calculation) around each of the nuclei.

Further reduction of the diagonal  $F_{\mu, \mu}$  expression is achieved by:

a. Combining terms in the sum involving orbitals on atomic centers other than where  $\mu$  sits (atom  $a$ ) together with the sum of coulomb attractions (which appear in  $h$ ) over these same centers:

$$\begin{aligned} & \sum_b (\text{not on atom } a) \delta_{\mu, \nu} \mu_{, \nu} \\ & - \sum_b (\text{not on atom } a) \langle \mu | Z_b e^2 / |r - R_b|^{-1} | \mu \rangle \\ & = \sum_b (\text{not center } a) \left( \delta_{\mu, \nu} \mu_{, \nu} - V_{\mu, b} \right). \end{aligned}$$

This combination represents the net coulombic interaction of  $|\mu|^2$  with the total electron density (first sum) and the total attractive positive density (second sum) on atoms other than the atom on which  $\mu$  sits.

## b. Recognizing

$$\langle \mu | -\hbar^2 \nabla^2 / 2m | \mu \rangle + \langle \mu | Z_a e^2 / |r - R_a| | \mu \rangle = U_{\mu, \mu}$$

as the average value of the atomic Fock operator (i.e., kinetic energy plus attractive colomb potential to that atom's nucleus plus coulomb and exchange interactions with other electrons on that atom) for an electron in  $\mu$  on the nucleus a. As in Hückel theory, the values of these parameters  $U_{\mu, \mu}$ , which play the role of the Hückel  $H_{\mu, \mu}$ , can be determined from atomic valence-state ionization potentials and electron affinities. These quantities, in turn, may be obtained either from experimental data or from the results of ab initio calculations.

As a result, the diagonal F matrix elements are given by

$$F_{\mu, \mu} = U_{\mu, \mu} + \langle \mu | -\hbar^2 \nabla^2 / 2m | \mu \rangle + \langle \mu | Z_a e^2 / |r - R_a| | \mu \rangle + \langle \mu | \sum_{b \neq a} (J_{\mu, b} - K_{\mu, b}) | \mu \rangle - V_{\mu, b}$$

The evaluation of the quantities entering into this expression and that for the off-diagonal  $F_{\mu, \nu}$  elements differs from one semi-empirical method to another; this topic is covered later in this Appendix.

Reduction of the off-diagonal elements involving orbitals  $\mu$  and  $\nu$  on the same atom (a) is achieved by assuming that the atomic orbitals have been formed in a manner that makes the contributions to  $F_{\mu, \nu}$  from atom a vanish

$$\langle \mu | -\hbar^2 \nabla^2 / 2m | \nu \rangle - \langle \mu | Z_a e^2 / |r - R_a| | \nu \rangle - \langle \mu | \sum_{b \neq a} (J_{\mu, b} - K_{\mu, b}) | \nu \rangle + \langle \mu | \sum_{b \neq a} (J_{\nu, b} - K_{\nu, b}) | \nu \rangle = 0,$$

and then neglecting, to be consistent with the ZDO assumption, the contributions from atoms other than atom a

$$- \langle \mu | Z_a e^2 / |r - R_a| | \nu \rangle + \langle \nu | Z_a e^2 / |r - R_a| | \mu \rangle = 0,$$

Hence, the off-diagonal F matrix elements vanish,  $F_{\mu, \nu} = 0$  for  $\mu$  and  $\nu$  both on the same atom (a).

The off-diagonal F matrix elements coupling orbitals from different atoms (a and b) are expressed as

$$F_{\mu, \nu} = \langle \mu | -\hbar^2 \nabla^2 / 2m | \nu \rangle - \langle \mu | Z_a e^2 / |r - R_a| | \nu \rangle - \langle \mu | Z_b e^2 / |r - R_b| | \nu \rangle - \langle \mu | \sum_{c \neq a, b} (J_{\mu, c} - K_{\mu, c}) | \nu \rangle + \langle \nu | \sum_{c \neq a, b} (J_{\nu, c} - K_{\nu, c}) | \mu \rangle$$

Contributions to these elements from atoms other than a and b are neglected, again to be consistent with the ZDO approximation.

Unlike the Hückel and extended Hückel methods, the semi-empirical approaches that explicitly treat electron-electron interactions give rise to Fock matrix element

expressions that depend on the atomic-orbital-based density matrix  $\rho_{\mu,\nu}$ . This quantity is computed using the LCAO-MO coefficients  $\{C_{i,\mu}\}$  of the occupied molecular orbitals from the previous iteration of the

$$F_{\mu,\nu} = \sum_i \rho_{i,\mu} \langle i | \nu \rangle C_{i,\mu}$$

equations. In particular,

$$F_{\mu,\mu} = \sum_{i(\text{occupied})} n_i C_{i,\mu}^2$$

$$F_{\mu,\mu}^{\text{ex}} = \sum_{i(\text{occupied and of spin } \sigma)} C_{i,\mu}^2$$

Here,  $n_i$  is the number (0, 1, or 2) of electrons that occupy the  $i^{\text{th}}$  molecular orbital, and  $\sigma$  denotes the spin ( $\uparrow$  or  $\downarrow$ ) of the orbital whose Fock matrix is being formed. For example, when studying doublet radicals having  $K$  doubly occupied orbitals and one half-filled orbital ( $K+1$ ) in which an  $\uparrow$  electron resides,  $\sigma = \uparrow$ . In this case,  $n_i = 2$  for the first  $K$  orbitals and  $n_i = 1$  for the last occupied orbital. Moreover, the Fock matrix elements  $F_{\mu,\nu}$  for  $\nu$  orbitals contains contributions from  $F_{\mu,\nu}^{\text{ex}}$  that are of the form

$$F_{\mu,\nu}^{\text{ex}} = \sum_{i=1,K} C_{i,\mu} C_{i,\nu}$$

while the Fock matrix elements  $F_{\mu,\nu}$  for  $\nu$  orbitals contains

$$F_{\mu,\nu}^{\text{ex}} = \sum_{i=1,K} C_{i,\mu} C_{i,\nu} + C_{K+1,\mu} C_{K+1,\nu}$$

For both  $F_{\mu,\mu}$  and  $F_{\mu,\mu}^{\text{ex}}$ , coulomb contributions arise as

$$F_{\mu,\mu} = 2 \sum_{i=1,K} C_{i,\mu}^2 + C_{K+1,\mu}^2$$

### 3. Various Semi-Empirical Methods

#### a. The Pariser-Parr-Pople (PPP) Method for $\pi$ -Orbitals

In the PPP method, only the  $\pi$ -orbitals and the corresponding  $\pi$ -electrons are considered. The parameters included in the F matrix

$$F_{\mu,\nu} = U_{\mu,\nu} + \left( \rho_{\mu,\mu} - \rho_{\mu,\mu}^{\text{ex}} \right) \rho_{\mu,\nu} + \rho_{\nu,\nu} \left( \delta_{\mu,\nu} - V_{\mu,\nu} \right)$$

$$F_{\mu,\mu}^{\text{ex}} = \rho_{\mu,\mu} - \rho_{\mu,\mu}^{\text{ex}}$$

are obtained as follows:

i). The diagonal integrals  $U_{a,a}$ , which represent the mutual coulomb repulsions between a pair of electrons in the valence-state orbital labeled  $a$ , can be estimated, as suggested by Pariser, in terms of the valence-state IP and EA of that orbital:

$$a,a = IP_a - EA_a .$$

Alternatively, these one-center coulomb integrals can be computed from first principles using Slater or Gaussian type orbitals.

ii). The off-diagonal coulomb integrals  $a,b$  are commonly approximated either by the Mataga-Nishimoto formula:

$$a,b = e^2 / (R_{a,b} + x_{a,b}),$$

where

$$x_{a,b} = 2e^2 / (a,a + b,b),$$

or by the Dewar-Ohno-Klopman expression:

$$a,b = \frac{e^2}{\sqrt{(R_{a,b}^2 + 0.25 e^4 (1/a,a + 1/b,b)^2)}} .$$

iii). The valence-state IP's and EA's,  $U_{\mu,\mu}$  are evaluated from experimental data or from the results of ab initio calculations of the atomic IP's and EA's.

iv). The  $\mu, \mu$  integrals are usually chosen to make bond lengths, bond energies, or electronic excitation energies in the molecule agree with experimental data. The geometry dependence of  $\mu, \mu$  is often parameterized as in Hückel theory  $\mu, \mu = \rho_{\mu}, S_{\mu, \mu}$ , and the overlap is then computed from first principles.

v). The  $b$  (not center a)  $V_{\mu,b}$  term, which represents the coulombic attraction of an electron in  $\mu$  to the nucleus at center b, is often approximated as  $Z_b \mu, \mu$ , where  $\mu, \mu$  labels the one orbital on center b, and  $Z_b$  is the number of electrons contributed by center b. This parameterization then permits the attractive interaction for center b to be combined with the repulsive interaction to give

$$b, \mu \text{ (not center a)} = \mu, \mu - Z_b \mu, \mu .$$

## b. All Valence Electron Methods

The CNDO, INDO, NDDO, MNDO, and MINDO methods all are defined in terms of an orbital-level Fock matrix with elements

$$F_{\mu,\mu} = U_{\mu,\mu} + (\mu, \mu - \mu, \mu^{ex}) \mu, \mu + b, \mu \text{ (not center a)} = \mu, \mu - V_{\mu,b} .$$

$$F_{\mu, \mu} = \mu, \mu - \mu, \mu^{ex} \mu, \mu .$$

They differ among one another in two ways: (i) in the degree to which they employ the ZDO approximation to eliminate two-electron integrals, and (ii) in whether they employ

experimental data (MINDO, MNDO, CNDO/S) or results of ab initio one-electron calculations (CNDO, INDO, NDDO) to define their parameters.

The CNDO and CNDO/S methods apply the ZDO approximation to all integrals, regardless of whether the orbitals are located on the same atom or not. In the INDO method, which was designed to improve the treatment of spin densities at nuclear centers and to handle singlet-triplet energy differences for open-shell species, exchange integrals  $\langle a | g | b \rangle$  involving orbitals  $a$  and  $b$  on the same atom are retained. In the NDDO approach, the ZDO approximation is applied only to integrals involving orbitals on two or more different atoms; that is, all one center integrals are retained. The text Approximate Molecular Orbital Theory by J. A. Pople and D. L. Beveridge, McGraw-Hill, New York (1970) gives a treatment of several of these semi-empirical methods beyond the introduction provided here.

To illustrate the differences among the various approaches and to clarify how their parameters are obtained, let us consider two specific and popular choices- CNDO/2 and MINDO.

#### i. The CNDO/2 and CNDO/S Models

In the CNDO/2 approach as originally implemented, ab initio (orbital-level) calculated values of the energies mentioned below are used in determining the requisite parameters. In the later CNDO/S method, experimental values of these energies are employed. Briefly, in any CNDO method:

1). The diagonal integrals  $U_{a,a}$ , which represent the mutual coulomb repulsions between a pair of electrons in the valence-state orbital labeled  $a$ , are calculated in terms of the valence-state IP and EA of that orbital:

$$U_{a,a} = IP_a - EA_a .$$

2). The valence-state IP's and EA's, and hence the  $U_{\mu,\mu}$  are evaluated from the results of ab initio calculations (CNDO/2) or experimental measurement (CNDO/S) of the atomic IP's and EA's. The expressions used are:

$$-IP_A = U_{a,a} + (Z_A - 1) U_{a,a},$$

for orbitals  $a$ , and

$$-EA_A = U_{b,b} + Z_A U_{b,b},$$

for orbitals  $b$ . Here  $Z_A$  is the effective core charge of atom A (the nuclear charge minus the number of "core" electrons not explicitly treated). For first row atoms, several  $U_{a,a}$  and  $EA_{a,a}$  values are tabulated below (all quantities are in eV).

	H	Li	Be	B	C	N	O	F
$U_{s,s}$	-13.6	-5.00	-15.4	-30.37	-50.69	-70.09	-101.3	-129.5
$U_{p,p}$		-3.67	-12.28	-24.7	-41.53	-57.85	-84.28	-108.9
$EA_{a,a}$	12.85	3.46	5.95	8.05	10.33	11.31	13.91	15.23

3). The off-diagonal coulomb integrals  $J_{a,b}$  are commonly approximated either by the Mataga-Nishimoto formula:

$$J_{a,b} = e^2 / (R_{a,b} + x_{a,b}),$$

where

$$x_{a,b} = 2e^2 / (I_{a,a} + I_{b,b}),$$

or by the Dewar-Ohno-Klopman expression:

$$J_{a,b} = \frac{e^2}{\sqrt{(R_{a,b}^2 + 0.25 e^4 (1/I_{a,a} + 1/I_{b,b})^2)}}.$$

4). As in PPP theory, the term  $J_{b, \mu}$  (not center a) ( $\mu, -V_{\mu,b}$ ) is approximated by  $J_{b, \mu} = (I_{\mu, \mu} - Z_b) V_{\mu,b}$ , where  $Z_b$  is the number of valence electrons contributed by atom b, and  $\mu$  is one of the valence electrons on atom b.

5. The  $S_{\mu, \nu}$  parameters are approximated as  $S_{\mu, \nu} = S_{\mu, \nu} (I_{a,a} + I_{b,b})$ , where  $S_{\mu, \nu}$  is the overlap between the orbitals  $\mu$  and  $\nu$ , and  $I_{a,a}$  and  $I_{b,b}$  are atom-dependent parameters given below for first row atoms:

	H	Li	Be	B	C	N	O	F
$I_{a,a}$ (eV)	-9	-9	-13	-17	-21	-25	-31	-39

## ii. The INDO (and MINDO-Type) Methods

In these methods, the specification of the parameters entering into  $F_{\mu, \nu}$  is carried out in the same fashion as in the CNDO/2 approach, except that:

1. The ZDO approximation is made only for two-center integrals; one-center coulomb  $J_{a,b} = \langle a | g | a \rangle$  and exchange

$K_{a,b} = \langle a | g | b \rangle$  integrals are retained. In the INDO approach, the values of these single-atom integrals are determined by requiring the results of the calculation, performed at the Fock-like orbital level, to agree with results of ab initio Fock-level calculations. In the MINDO approach, experimental electronic spectra of the particular atom are used to determine these parameters. The "diagonal" values  $J_{a,a}$  are determined, as indicated earlier, from valence-state energies (ab initio for INDO and experimental for MINDO) of the atom A on which  $a$  resides.

2. The values of the  $U_{a,a}$  parameters are determined according to the following equations:

$$U_{s,s} = -0.5 (IP_H + EA_H) - 0.5 I_{s,s}$$

for Hydrogen's 1s orbital;

$$U_{s,s} = -0.5(IP_s + EA_s) - (Z_A - 0.5) \int_{s,s} + 1/6 (Z_A - 1.5) G^1(s,s)$$

for Boron through Fluorine's 2s orbitals; and

$$U_{p,p} = -0.5(IP_p + EA_p) - (Z_A - 0.5) \int_{p,p} + 2/25 (Z_A - 1.5) F^2(p,p) + \\ 1/3 G^1(p,p).$$

Here,  $F^2$  and  $G^1$  represent the well known Slater-Condon integrals in terms of which the coulomb and exchange integrals can be expressed:

$$F^k(nl, n'l') = \int_0^\infty \int_0^\infty |R_{nl}(r)|^2 |R_{n'l'}(r')|^2 2r_{<}^k / r_{>}^{k+1} r^2 r'^2 dr dr'$$

$$G^k(nl, n'l') = \int_0^\infty \int_0^\infty |R_{nl}(r)R_{n'l'}(r)|^2 2r_{<}^k / r_{>}^{k+1} r^2 r'^2 dr dr'$$

and  $Z_A$  is the effective core charge (the nuclear charge minus the number of "core" electrons not explicitly treated in the calculation) of the atom A on which the orbitals in question reside. In the definitions of the integrals,  $r_{<}$  and  $r_{>}$  represent, respectively, the smaller and larger of  $r$  and  $r'$ . Again, ab initio calculational data is used in the INDO method, and experimental data in the MINDO method to fix the parameters entering these expressions.

#### D. Summary

As presented, semi-empirical methods are based on a single-configuration picture of electronic structure. Extensions of such approaches to permit consideration of more than a single important configuration have been made (for excellent overviews, see Approximate Molecular Orbital Theory by J. A. Pople and D. L. Beveridge, McGraw-Hill, New York (1970) and Valence Theory, 2<sup>nd</sup> Ed., by J. N. Murrell, S. F. A. Kettle, and J. M. Tedder, John Wiley, London (1965)). Pople and co-workers preferred to use data from ab initio calculations in developing sets of parameters to use in such methods because they viewed semi-empirical methods as approximations to ab initio methods. Others use experimental data to determine parameters because they view semi-empirical methods as models of mother nature.

## Angular Momentum Operator Identities G

### I. Orbital Angular Momentum

A particle moving with momentum  $\mathbf{p}$  at a position  $\mathbf{r}$  relative to some coordinate origin has so-called orbital angular momentum equal to  $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ . The three components of this angular momentum vector in a cartesian coordinate system located at the origin mentioned above are given in terms of the cartesian coordinates of  $\mathbf{r}$  and  $\mathbf{p}$  as follows:

$$L_z = x p_y - y p_x ,$$

$$L_x = y p_z - z p_y ,$$

$$L_y = z p_x - x p_z .$$

Using the fact that the quantum mechanical coordinate operators  $\{q_k\} = x, y, z$  as well as the conjugate momentum operators  $\{p_j\} = p_x, p_y, p_z$  are Hermitian, it is possible to show that  $L_x, L_y,$  and  $L_z$  are also Hermitian, as they must be if they are to correspond to experimentally measurable quantities.

Using the fundamental commutation relations among the cartesian coordinates and the cartesian momenta:

$$[q_k, p_j] = q_k p_j - p_j q_k = i\hbar \delta_{j,k} \quad (j, k = x, y, z) ,$$

it can be shown that the above angular momentum operators obey the following set of commutation relations:

$$[L_x, L_y] = i\hbar L_z ,$$

$$[L_y, L_z] = i\hbar L_x ,$$

$$[L_z, L_x] = i\hbar L_y .$$

Although the components of  $\mathbf{L}$  do not commute with one another, they can be shown to commute with the operator  $L^2$  defined by

$$L^2 = L_x^2 + L_y^2 + L_z^2 .$$

This new operator is referred to as the square of the total angular momentum operator.

The commutation properties of the components of  $\mathbf{L}$  allow us to conclude that complete sets of functions can be found that are eigenfunctions of  $L^2$  and of one, but not more than one, component of  $\mathbf{L}$ . It is convention to select this one component as  $L_z$ , and to label the resulting simultaneous eigenstates of  $L^2$  and  $L_z$  as  $|l, m\rangle$  according to the corresponding eigenvalues:

$$L^2 |l, m\rangle = \hbar^2 l(l+1) |l, m\rangle, \quad l = 0, 1, 2, 3, \dots$$

$$L_z |l, m\rangle = \hbar m |l, m\rangle, \quad m = \pm 1, \pm(1-1), \pm(1-2), \dots \pm(1-(l-1)), 0.$$

That these eigenvalues assume the values specified in these identities is proven in considerable detail below. These eigenfunctions of  $L^2$  and of  $L_z$  will not, in general, be eigenfunctions of either  $L_x$  or of  $L_y$ . This means that any measurement of  $L_x$  or  $L_y$  will necessarily change the wavefunction if it begins as an eigenfunction of  $L_z$ .

The above expressions for  $L_x$ ,  $L_y$ , and  $L_z$  can be mapped into quantum mechanical operators by substituting  $x$ ,  $y$ , and  $z$  as the corresponding coordinate operators and  $-i\hbar / x$ ,  $-i\hbar / y$ , and  $-i\hbar / z$  for  $p_x$ ,  $p_y$ , and  $p_z$ , respectively. The resulting operators can then be transformed into spherical coordinates by using the techniques provided in Appendix (A), the results of which are:

$$L_z = -i\hbar \frac{\partial}{\partial \phi} ,$$

$$L_x = i\hbar \left\{ \sin \theta \frac{\partial}{\partial \phi} + \cot \theta \cos \theta \frac{\partial}{\partial \theta} \right\} ,$$

$$L_y = -i\hbar \left\{ \cos \theta \frac{\partial}{\partial \phi} - \cot \theta \sin \theta \frac{\partial}{\partial \theta} \right\} ,$$

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

At this point, it should be again stressed that the above form for  $L^2$  appears explicitly when the kinetic energy operator  $-\hbar^2/2m \nabla^2$  is expressed in spherical coordinates; in particular, the term  $L^2/2mr^2$  is what enters. This means that our study of the properties of angular momenta will also help us to understand the angular-motion components of the Hamiltonian for spherically symmetric systems (i.e., those for which the potential  $V$  contains no angle dependence, and hence for which the total angle dependence is contained in the kinetic energy term  $L^2/2mr^2$ ).

## II. Properties of General Angular Momenta

There are many types of angular momenta that one encounters in chemistry. Orbital angular momenta, such as that introduced above, arise in electronic motion in atoms, in atom-atom and electron-atom collisions, and in rotational motion in molecules. Intrinsic spin angular momentum is present in electrons,  $H^1$ ,  $H^2$ ,  $C^{13}$ , and many other nuclei. In this section, we will deal with the behavior of any and all angular momenta and their corresponding eigenfunctions.

At times, an atom or molecule contains more than one type of angular momentum. The Hamiltonian's interaction potentials present in a particular species may or may not cause these individual angular momenta to be coupled to an appreciable extent (i.e., the Hamiltonian may or may not contain terms that refer simultaneously to two or more of these angular momenta). For example, the  $NH^-$  ion, which has a  $^2 \Sigma^-$  ground electronic state (its electronic configuration is  $1s_N^2 2s^2 2p_x^2 2p_y^1$ ) has electronic spin, electronic orbital, and molecular rotational angular momenta. The full Hamiltonian  $H$  contains spin-orbit coupling terms that couple the electronic spin and orbital angular momenta, thereby causing them individually to not commute with  $H$ .  $H$  also contains terms that couple the ion's rotational and electronic angular momenta, thereby making these quantities no longer "good" quantum numbers (i.e., making the corresponding operators no longer commute with  $H$ ).

In such cases, the eigenstates of the system can be labeled rigorously only by angular momentum quantum numbers  $j$  and  $m$  belonging to the total angular momentum  $\mathbf{J}$ . The total angular momentum of a collection of individual angular momenta is defined, component-by-component, as follows:

$$\mathbf{J}_k = \sum_i \mathbf{J}_k(i),$$

where  $k$  labels  $x$ ,  $y$ , and  $z$ , and  $i$  labels the constituents whose angular momenta couple to produce  $\mathbf{J}$ .

For the remainder of this Appendix, we will study eigenfunction-eigenvalue relationships that are characteristic of all angular momenta and which are consequences of the commutation relations among the angular momentum vector's three components. We will also study how one combines eigenfunctions of two or more angular momenta  $\{\mathbf{J}(i)\}$  to produce eigenfunctions of the the total  $\mathbf{J}$ .

### A. Consequences of the Commutation Relations

Any set of three Hermitian operators that obey

$$[J_x, J_y] = i\hbar J_z,$$

$$[J_y, J_z] = i\hbar J_x,$$

$$[J_z, J_x] = i\hbar J_y,$$

will be taken to define an angular momentum  $\mathbf{J}$ , whose square  $J^2 = J_x^2 + J_y^2 + J_z^2$  commutes with all three of its components. It is useful to also introduce two combinations of the three fundamental operators:

$$J_{\pm} = J_x \pm i J_y,$$

and to refer to them as raising and lowering operators for reasons that will be made clear below. These new operators can be shown to obey the following commutation relations:

$$[J^2, J_{\pm}] = 0,$$

$$[J_z, J_{\pm}] = \pm \hbar J_{\pm}.$$

These two operators are not Hermitian operators (although  $J_x$  and  $J_y$  are), but they are adjoints of one another:

$$J_+^\dagger = J_-,$$

$$J_-^\dagger = J_+.$$

as can be shown using the self-adjoint nature of  $J_x$  and  $J_y$ .

Using only the above commutation properties, it is possible to prove important properties of the eigenfunctions and eigenvalues of  $J^2$  and  $J_z$ . Let us assume that we have found a set of simultaneous eigenfunctions of  $J^2$  and  $J_z$ ; the fact that these two operators

commute tells us that this is possible. Let us label the eigenvalues belonging to these functions:

$$J^2 |j,m\rangle = \hbar^2 f(j,m) |j,m\rangle,$$

$$J_z |j,m\rangle = \hbar m |j,m\rangle,$$

in terms of the quantities  $m$  and  $f(j,m)$ . Although we certainly "hint" that these quantities must be related to certain  $j$  and  $m$  quantum numbers, we have not yet proven this, although we will soon do so. For now, we view  $f(j,m)$  and  $m$  simply as symbols that represent the respective eigenvalues. Because both  $J^2$  and  $J_z$  are Hermitian, eigenfunctions belonging to different  $f(j,m)$  or  $m$  quantum numbers must be orthogonal:

$$\langle j,m | j',m' \rangle = \delta_{m,m'} \delta_{j,j'}.$$

We now prove several identities that are needed to discover the information about the eigenvalues and eigenfunctions of general angular momenta that we are after. Later in this Appendix, the essential results are summarized.

#### 1. There is a Maximum and a Minimum Eigenvalue for $J_z$

Because all of the components of  $\mathbf{J}$  are Hermitian, and because the scalar product of any function with itself is positive semi-definite, the following identity holds:

$$\langle j,m | J_x^2 + J_y^2 | j,m \rangle = \langle J_x | j,m \rangle \langle J_x | j,m \rangle + \langle J_y | j,m \rangle \langle J_y | j,m \rangle \geq 0.$$

However,  $J_x^2 + J_y^2$  is equal to  $J^2 - J_z^2$ , so this inequality implies that

$$\langle j,m | J^2 - J_z^2 | j,m \rangle = \hbar^2 \{f(j,m) - m^2\} \geq 0,$$

which, in turn, implies that  $m^2$  must be less than or equal to  $f(j,m)$ . Hence, for any value of the total angular momentum eigenvalue  $f$ , the  $z$ -projection eigenvalue ( $m$ ) must have a maximum and a minimum value and both of these must be less than or equal to the total angular momentum squared eigenvalue  $f$ .

#### 2. The Raising and Lowering Operators Change the $J_z$ Eigenvalue but not the $J^2$ Eigenvalue When Acting on $|j,m\rangle$

Applying the commutation relations obeyed by  $J_{\pm}$  to  $|j,m\rangle$  yields another useful result:

$$J_z J_{\pm} |j,m\rangle - J_{\pm} J_z |j,m\rangle = \pm \hbar J_{\pm} |j,m\rangle,$$

$$J^2 J_{\pm} |j,m\rangle - J_{\pm} J^2 |j,m\rangle = 0.$$

Now, using the fact that  $|j,m\rangle$  is an eigenstate of  $J^2$  and of  $J_z$ , these identities give

$$J_z J_{\pm} |j,m\rangle = (m\hbar \pm \hbar) J_{\pm} |j,m\rangle = \hbar (m \pm 1) |j,m\rangle,$$

$$J^2 J_{\pm} |j,m\rangle = \hbar^2 f(j,m) J_{\pm} |j,m\rangle.$$

These equations prove that the functions  $J_{\pm} |j,m\rangle$  must either themselves be eigenfunctions of  $J^2$  and  $J_z$ , with eigenvalues  $\hbar^2 f(j,m)$  and  $\hbar (m\pm 1)$  or  $J_{\pm} |j,m\rangle$  must equal zero. In the former case, we see that  $J_{\pm}$  acting on  $|j,m\rangle$  generates a new eigenstate with the same  $J^2$  eigenvalue as  $|j,m\rangle$  but with one unit of  $\hbar$  higher in  $J_z$  eigenvalue. It is for this reason that we call  $J_{\pm}$  raising and lowering operators. Notice that, although  $J_{\pm} |j,m\rangle$  is indeed an eigenfunction of  $J_z$  with eigenvalue  $(m\pm 1)\hbar$ ,  $J_{\pm} |j,m\rangle$  is not identical to  $|j,m\pm 1\rangle$ ; it is only proportional to  $|j,m\pm 1\rangle$ :

$$J_{\pm} |j,m\rangle = C_{\pm,j,m} |j,m\pm 1\rangle.$$

Explicit expressions for these  $C_{\pm,j,m}$  coefficients will be obtained below. Notice also that because the  $J_{\pm} |j,m\rangle$ , and hence  $|j,m\pm 1\rangle$ , have the same  $J^2$  eigenvalue as  $|j,m\rangle$  (in fact, sequential application of  $J_{\pm}$  can be used to show that all  $|j,m'\rangle$ , for all  $m'$ , have this same  $J^2$  eigenvalue), the  $J^2$  eigenvalue  $f(j,m)$  must be independent of  $m$ . For this reason,  $f$  can be labeled by one quantum number  $j$ .

### 3. The $J^2$ Eigenvalues are Related to the Maximum and Minimum $J_z$ Eigenvalues Which are Related to One Another

Earlier, we showed that there exists a maximum and a minimum value for  $m$ , for any given total angular momentum. It is when one reaches these limiting cases that  $J_{\pm} |j,m\rangle = 0$  applies. In particular,

$$J_+ |j,m_{\max}\rangle = 0,$$

$$J_- |j,m_{\min}\rangle = 0.$$

Applying the following identities:

$$J_- J_+ = J^2 - J_z^2 - \hbar J_z,$$

$$J_+ J_- = J^2 - J_z^2 + \hbar J_z,$$

respectively, to  $|j,m_{\max}\rangle$  and  $|j,m_{\min}\rangle$  gives

$$\hbar^2 \{ f(j,m_{\max}) - m_{\max}^2 - m_{\max} \} = 0,$$

$$\hbar^2 \{ f(j,m_{\min}) - m_{\min}^2 + m_{\min} \} = 0,$$

which immediately gives the  $J^2$  eigenvalue  $f(j,m_{\max})$  and  $f(j,m_{\min})$  in terms of  $m_{\max}$  or  $m_{\min}$ :

$$f(j,m_{\max}) = m_{\max} (m_{\max} + 1),$$

$$f(j,m_{\min}) = m_{\min} (m_{\min} - 1).$$

So, we now know the  $J^2$  eigenvalues for  $|j, m_{\max}\rangle$  and  $|j, m_{\min}\rangle$ . However, we earlier showed that  $|j, m\rangle$  and  $|j, m-1\rangle$  have the same  $J^2$  eigenvalue (when we treated the effect of  $J_{\pm}$  on  $|j, m\rangle$ ) and that the  $J^2$  eigenvalue is independent of  $m$ . If we therefore define the quantum number  $j$  to be  $m_{\max}$ , we see that the  $J^2$  eigenvalues are given by

$$J^2 |j, m\rangle = \hbar^2 j(j+1) |j, m\rangle.$$

We also see that

$$f(j, m) = j(j+1) = m_{\max}(m_{\max}+1) = m_{\min}(m_{\min}-1),$$

from which it follows that

$$m_{\min} = -m_{\max}.$$

#### 4. The $j$ Quantum Number Can Be Integer or Half-Integer

The fact that the  $m$ -values run from  $j$  to  $-j$  in unit steps (because of the property of the  $J_{\pm}$  operators), there clearly can be only integer or half-integer values for  $j$ . In the former case, the  $m$  quantum number runs over  $-j, -j+1, -j+2, \dots, -j+(j-1), 0, 1, 2, \dots, j$ ; in the latter,  $m$  runs over  $-j, -j+1, -j+2, \dots, -j+(j-1/2), 1/2, 3/2, \dots, j$ . Only integer and half-integer values can range from  $j$  to  $-j$  in steps of unity. Species with integer spin are known as Bosons and those with half-integer spin are called Fermions.

#### 5. More on $J_{\pm} |j, m\rangle$

Using the above results for the effect of  $J_{\pm}$  acting on  $|j, m\rangle$  and the fact that  $J_+$  and  $J_-$  are adjoints of one another, allows us to write:

$$\begin{aligned} \langle j, m | J_- J_+ |j, m\rangle &= \langle j, m | (J^2 - J_z^2 - \hbar J_z) |j, m\rangle \\ &= \hbar^2 \{j(j+1) - m(m+1)\} = \langle J_+ \langle j, m | J_+ |j, m\rangle = (C^+_{j, m})^2, \end{aligned}$$

where  $C^+_{j, m}$  is the proportionality constant between  $J_+ |j, m\rangle$  and the normalized function  $|j, m+1\rangle$ . Likewise, the effect of  $J_-$  can be

$$\begin{aligned} \langle j, m | J_+ J_- |j, m\rangle &= \langle j, m | J_+ (J^2 - J_z^2 + \hbar J_z) |j, m\rangle \\ &= \hbar^2 \{j(j+1) - m(m-1)\} = \langle J_- \langle j, m | J_- |j, m\rangle = (C^-_{j, m})^2, \end{aligned}$$

where  $C^-_{j, m}$  is the proportionality constant between  $J_- |j, m\rangle$  and the normalized  $|j, m-1\rangle$ . Thus, we can solve for  $C^{\pm}_{j, m}$  after which the effect of  $J_{\pm}$  on  $|j, m\rangle$  is given by:

$$J_{\pm} |j, m\rangle = \hbar \{j(j+1) - m(m\pm 1)\}^{1/2} |j, m\pm 1\rangle.$$

## B. Summary

The above results apply to any angular momentum operators. The essential findings can be summarized as follows:

(i)  $J^2$  and  $J_z$  have complete sets of simultaneous eigenfunctions. We label these eigenfunctions  $|j,m\rangle$ ; they are orthonormal in both their  $m$ - and  $j$ -type indices:  $\langle j,m|j',m'\rangle = \delta_{m,m'} \delta_{j,j'}$ .

(ii) These  $|j,m\rangle$  eigenfunctions obey:

$$J^2 |j,m\rangle = \hbar^2 j(j+1) |j,m\rangle, \quad \{ j = \text{integer or half-integer} \},$$

$$J_z |j,m\rangle = \hbar m |j,m\rangle, \quad \{ m = -j, \text{ in steps of } 1 \text{ to } +j \}.$$

(iii) The raising and lowering operators  $J_{\pm}$  act on  $|j,m\rangle$  to yield functions that are eigenfunctions of  $J^2$  with the same eigenvalue as  $|j,m\rangle$  and eigenfunctions of  $J_z$  with eigenvalue of  $(m\pm 1)\hbar$ :

$$J_{\pm} |j,m\rangle = \hbar \{ j(j+1) - m(m\pm 1) \}^{1/2} |j,m\pm 1\rangle.$$

(iv) When  $J_{\pm}$  acts on the "extremal" states  $|j,j\rangle$  or  $|j,-j\rangle$ , respectively, the result is zero.

The results given above are, as stated, general. Any and all angular momenta have quantum mechanical operators that obey these equations. It is convention to designate specific kinds of angular momenta by specific letters; however, it should be kept in mind that no matter what letters are used, there are operators corresponding to  $J^2$ ,  $J_z$ , and  $J_{\pm}$  that obey relations as specified above, and there are eigenfunctions and eigenvalues that have all of the properties obtained above. For electronic or collisional orbital angular momenta, it is common to use  $L^2$  and  $L_z$ ; for electron spin,  $S^2$  and  $S_z$  are used; for nuclear spin  $I^2$  and  $I_z$  are most common; and for molecular rotational angular momentum,  $N^2$  and  $N_z$  are most common (although sometimes  $J^2$  and  $J_z$  may be used). Whenever two or more angular momenta are combined or coupled to produce a "total" angular momentum, the latter is designated by  $J^2$  and  $J_z$ .

### III. Coupling of Angular Momenta

If the Hamiltonian under study contains terms that couple two or more angular momenta  $\mathbf{J}(i)$ , then only the components of the total angular momentum  $\mathbf{J} = \sum_i \mathbf{J}(i)$  and  $J^2$  will commute with  $H$ . It is therefore essential to label the quantum states of the system by the eigenvalues of  $J_z$  and  $J^2$  and to construct variational trial or model wavefunctions that are eigenfunctions of these total angular momentum operators. The problem of angular momentum coupling has to do with how to combine eigenfunctions of the uncoupled angular momentum operators, which are given as simple products of the eigenfunctions of the individual angular momenta  $\sum_i |j_i, m_i\rangle$ , to form eigenfunctions of  $J^2$  and  $J_z$ .

#### A. Eigenfunctions of $J_z$

Because the individual elements of  $\mathbf{J}$  are formed additively, but  $J^2$  is not, it is straightforward to form eigenstates of

$$J_z = \sum_i J_z(i);$$

simple products of the form  $\prod_i |j_i, m_i\rangle$  are eigenfunctions of  $J_z$ :

$$J_z \prod_i |j_i, m_i\rangle = \sum_k J_z(k) \prod_i |j_i, m_i\rangle = \sum_k \hbar m_k \prod_i |j_i, m_i\rangle,$$

and have  $J_z$  eigenvalues equal to the sum of the individual  $m_k \hbar$  eigenvalues. Hence, to form an eigenfunction with specified  $J$  and  $M$  eigenvalues, one must combine only those product states  $\prod_i |j_i, m_i\rangle$  whose  $m_i \hbar$  sum is equal to the specified  $M$  value.

## B. Eigenfunctions of $J^2$ ; the Clebsch-Gordon Series

The task is then reduced to forming eigenfunctions  $|J, M\rangle$ , given particular values for the  $\{j_i\}$  quantum numbers (e.g., to couple the  $^3P$  states of the Si atom, which are eigenfunctions of  $L^2$  and of  $S^2$ , to produce a  $^3P_1$  state which is an eigenfunction of  $J^2$ , where  $J=L+S$ ). When coupling pairs of angular momenta  $\{|j, m\rangle$  and  $|j', m'\rangle\}$ , the total angular momentum states can be written, according to what we determined above, as

$$|J, M\rangle = \sum_{m, m'} C^{J, M}_{j, m; j', m'} |j, m\rangle |j', m'\rangle,$$

where the coefficients  $C^{J, M}_{j, m; j', m'}$  are called vector coupling coefficients (because angular momentum coupling is viewed much like adding two vectors  $\mathbf{j}$  and  $\mathbf{j}'$  to produce another vector  $\mathbf{J}$ ), and where the sum over  $m$  and  $m'$  is restricted to those terms for which  $m+m' = M$ . It is more common to express the vector coupling or so-called **Clebsch-Gordon (CG) coefficients** as  $\langle j, m; j' m' | J, M \rangle$  and to view them as elements of a "matrix" whose columns are labeled by the coupled-state  $J, M$  quantum numbers and whose rows are labeled by the quantum numbers characterizing the uncoupled "product basis"  $j, m; j', m'$ . It turns out (see Chapt. 2 of Angular Momentum, by R. N. Zare, John Wiley and Sons, New York, N.Y., (1988)) that this matrix can be shown to be unitary so that the CG coefficients obey:

$$\sum_{m, m'} \langle j, m; j' m' | J, M \rangle^* \langle j, m; j' m' | J', M' \rangle = \delta_{J, J'} \delta_{M, M'}$$

and

$$\sum_{J, M} \langle j, n; j' n' | J, M \rangle \langle j, m; j' m' | J, M \rangle^* = \delta_{n, m} \delta_{n', m'}.$$

This unitarity of the CG coefficient matrix allows the inverse of the relation giving coupled functions in terms of the product functions:

$$|J, M\rangle = \sum_{m, m'} \langle j, m; j' m' | J, M \rangle |j, m\rangle |j', m'\rangle$$

to be written as:

$$\begin{aligned} |j, m\rangle |j', m'\rangle &= \sum_{J, M} \langle j, m; j' m' | J, M \rangle^* |J, M\rangle \\ &= \sum_{J, M} \langle J, M | j, m; j' m' \rangle |J, M\rangle. \end{aligned}$$

This result expresses the product functions in terms of the coupled angular momentum functions.

### C. Generation of the CG Coefficients

The CG coefficients can be generated in a systematic manner; however, they can also be looked up in books where they have been tabulated (e.g., see Table 2.4 of Zare's book on angular momentum; the reference is given above). Here, we will demonstrate the technique by which the CG coefficients can be obtained, but we will do so for rather limited cases and refer the reader to more extensive tabulations.

The strategy we take is to generate the  $|J,J\rangle$  state (i.e., the state with maximum M-value) and to then use  $J_-$  to generate  $|J,J-1\rangle$ , after which the state  $|J-1,J-1\rangle$  (i.e., the state with one lower J-value) is constructed by finding a combination of the product states in terms of which  $|J,J-1\rangle$  is expressed (because both  $|J,J-1\rangle$  and  $|J-1,J-1\rangle$  have the same M-value  $M=J-1$ ) which is orthogonal to  $|J,J-1\rangle$  (because  $|J-1,J-1\rangle$  and  $|J,J-1\rangle$  are eigenfunctions of the Hermitian operator  $J^2$  corresponding to different eigenvalues, they must be orthogonal). This same process is then used to generate  $|J,J-2\rangle$   $|J-1,J-2\rangle$  and (by orthogonality construction)  $|J-2,J-2\rangle$ , and so on.

#### 1. The States With Maximum and Minimum M-Values

We begin with the state  $|J,J\rangle$  having the highest M-value. This state must be formed by taking the highest m and the highest m' values (i.e.,  $m=j$  and  $m'=j'$ ), and is given by:

$$|J,J\rangle = |j,j\rangle |j',j'\rangle.$$

Only this one product is needed because only the one term with  $m=j$  and  $m'=j'$  contributes to the sum in the above CG series. The state

$$|J,-J\rangle = |j,-j\rangle |j',-j'\rangle$$

with the minimum M-value is also given as a single product state.

Notice that these states have M-values given as  $\pm(j+j')$ ; since this is the maximum M-value, it must be that the J-value corresponding to this state is  $J=j+j'$ .

#### 2. States With One Lower M-Value But the Same J-Value

Applying  $J_-$  to  $|J,J\rangle$ , and expressing  $J_-$  as the sum of lowering operators for the two individual angular momenta:

$$J_- = J_-(1) + J_-(2)$$

gives

$$\begin{aligned} J_-|J,J\rangle &= \hbar\{J(J+1) - J(J-1)\}^{1/2} |J,J-1\rangle \\ &= (J_-(1) + J_-(2)) |j,j\rangle |j',j'\rangle \\ &= \hbar\{j(j+1) - j(j-1)\}^{1/2} |j,j-1\rangle |j',j'\rangle + \hbar\{j'(j'+1) - j'(j'-1)\}^{1/2} |j,j\rangle |j',j'-1\rangle. \end{aligned}$$

This result expresses  $|J,J-1\rangle$  as follows:

$$|J,J-1\rangle = \left[ \{j(j+1) - j(j-1)\}^{1/2} |j,j-1\rangle |j',j'\rangle \right]$$

$$+ \{j'(j'+1)-j(j-1)\}^{1/2} |j,j\rangle |j',j'-1\rangle \{J(J+1) - J(J-1)\}^{-1/2};$$

that is, the  $|J,J-1\rangle$  state, which has  $M=J-1$ , is formed from the two product states  $|j,j-1\rangle |j',j'\rangle$  and  $|j,j\rangle |j',j'-1\rangle$  that have this same  $M$ -value.

### 3. States With One Lower J-Value

To find the state  $|J-1,J-1\rangle$  that has the same  $M$ -value as the one found above but one lower  $J$ -value, we must construct another combination of the two product states with  $M=J-1$  (i.e.,  $|j,j-1\rangle |j',j'\rangle$  and  $|j,j\rangle |j',j'-1\rangle$ ) that is orthogonal to the combination representing  $|J,J-1\rangle$ ; after doing so, we must scale the resulting function so it is properly normalized. In this case, the desired function is:

$$|J-1,J-1\rangle = \{j(j+1)-j'(j'-1)\}^{1/2} |j,j\rangle |j',j'-1\rangle - \{j'(j'+1)-j(j-1)\}^{1/2} |j,j-1\rangle |j',j'\rangle \{J(J+1) - J(J-1)\}^{-1/2} .$$

It is straightforward to show that this function is indeed orthogonal to  $|J,J-1\rangle$ .

### 4. States With Even One Lower J-Value

Having expressed  $|J,J-1\rangle$  and  $|J-1,J-1\rangle$  in terms of  $|j,j-1\rangle |j',j'\rangle$  and  $|j,j\rangle |j',j'-1\rangle$ , we are now prepared to carry on with this stepwise process to generate the states  $|J,J-2\rangle$ ,  $|J-1,J-2\rangle$  and  $|J-2,J-2\rangle$  as combinations of the product states with  $M=J-2$ . These product states are  $|j,j-2\rangle |j',j'\rangle$ ,  $|j,j\rangle |j',j'-2\rangle$ , and  $|j,j-1\rangle |j',j'-1\rangle$ . Notice that there are precisely as many product states whose  $m+m'$  values add up to the desired  $M$ -value as there are total angular momentum states that must be constructed (there are three of each in this case).

The steps needed to find the state  $|J-2,J-2\rangle$  are analogous to those taken above:

a. One first applies  $J$  to  $|J-1,J-1\rangle$  and to  $|J,J-1\rangle$  to obtain  $|J-1,J-2\rangle$  and  $|J,J-2\rangle$ , respectively as combinations of  $|j,j-2\rangle |j',j'\rangle$ ,  $|j,j\rangle |j',j'-2\rangle$ , and  $|j,j-1\rangle |j',j'-1\rangle$ .

b. One then constructs  $|J-2,J-2\rangle$  as a linear combination of the  $|j,j-2\rangle |j',j'\rangle$ ,  $|j,j\rangle |j',j'-2\rangle$ , and  $|j,j-1\rangle |j',j'-1\rangle$  that is orthogonal to the combinations found for  $|J-1,J-2\rangle$  and  $|J,J-2\rangle$ .

Once  $|J-2,J-2\rangle$  is obtained, it is then possible to move on to form  $|J,J-3\rangle$ ,  $|J-1,J-3\rangle$ , and  $|J-2,J-3\rangle$  by applying  $J$  to the three states obtained in the preceding application of the process, and to then form  $|J-3,J-3\rangle$  as the combination of  $|j,j-3\rangle |j',j'\rangle$ ,  $|j,j\rangle |j',j'-3\rangle$ ,  $|j,j-2\rangle |j',j'-1\rangle$ ,  $|j,j-1\rangle |j',j'-2\rangle$  that is orthogonal to the combinations obtained for  $|J,J-3\rangle$ ,  $|J-1,J-3\rangle$ , and  $|J-2,J-3\rangle$ .

Again notice that there are precisely the correct number of product states (four here) as there are total angular momentum states to be formed. In fact, the product states and the total angular momentum states are equal in number and are both members of orthonormal function sets (because  $J^2(1)$ ,  $J_z(1)$ ,  $J^2(2)$ , and  $J_z(2)$  as well as  $J^2$  and  $J_z$  are Hermitian operators). This is why the CG coefficient matrix is unitary; because it maps one set of orthonormal functions to another, with both sets containing the same number of functions.

## D. An Example

Let us consider an example in which the spin and orbital angular momenta of the Si atom in its  $^3P$  ground state can be coupled to produce various  $^3P_J$  states. In this case, the specific values for  $j$  and  $j'$  are  $j=S=1$  and  $j'=L=1$ . We could, of course take  $j=L=1$  and

$j'=S=1$ , but the final wavefunctions obtained would span the same space as those we are about to determine.

The state with highest M-value is the  ${}^3P(M_S=1, M_L=1)$  state. As shown in Chapter 10 which deals with electronic configurations and states, this particular product wavefunction can be represented by the product of an spin function (representing  $S=1, M_S=1$ ) and a  $3p_13p_0$  spatial function (representing  $L=1, M_L=1$ ), where the first function corresponds to the first open-shell orbital and the second function to the second open-shell orbital. Thus, the maximum M-value is  $M=2$  and corresponds to a state with  $J=2$ :

$$|J=2, M=2\rangle = |2, 2\rangle = 3p_13p_0 .$$

Clearly, the state  $|2, -2\rangle$  would be given as  $3p_{-1}3p_0$ .

The states  $|2, 1\rangle$  and  $|1, 1\rangle$  with one lower M-value are obtained by applying  $J_+ = S_+ + L_+$  to  $|2, 2\rangle$  as follows:

$$\begin{aligned} J_+ |2, 2\rangle &= \hbar\{2(3)-2(1)\}^{1/2} |2, 1\rangle \\ &= (S_+ + L_+) 3p_13p_0 . \end{aligned}$$

To apply  $S_+$  or  $L_+$  to  $3p_13p_0$ , one must realize that each of these operators is, in turn, a sum of lowering operators for each of the two open-shell electrons:

$$\begin{aligned} S_+ &= S_+(1) + S_+(2), \\ L_+ &= L_+(1) + L_+(2). \end{aligned}$$

The result above can therefore be continued as

$$\begin{aligned} (S_+ + L_+) 3p_13p_0 &= \hbar\{1/2(3/2)-1/2(-1/2)\}^{1/2} 3p_13p_0 \\ &+ \hbar\{1/2(3/2)-1/2(-1/2)\}^{1/2} 3p_13p_0 \\ &+ \hbar\{1(2)-1(0)\}^{1/2} 3p_03p_0 \\ &+ \hbar\{1(2)-0(-1)\}^{1/2} 3p_13p_{-1}. \end{aligned}$$

So, the function  $|2, 1\rangle$  is given by

$$\begin{aligned} |2, 1\rangle &= [ 3p_13p_0 + 3p_13p_0 + \{2\}^{1/2} 3p_03p_0 \\ &+ \{2\}^{1/2} 3p_13p_{-1} ]/2, \end{aligned}$$

which can be rewritten as:

$$|2, 1\rangle = [( + )3p_13p_0 + \{2\}^{1/2} (3p_03p_0 + 3p_13p_{-1})]/2.$$

Writing the result in this way makes it clear that  $|2,1\rangle$  is a combination of the product states  $|S=1, M_S=0\rangle |L=1, M_L=1\rangle$  (the terms containing  $|S=1, M_S=0\rangle = 2^{-1/2}(\quad + \quad)$ ) and  $|S=1, M_S=1\rangle |L=1, M_L=0\rangle$  (the terms containing  $|S=1, M_S=1\rangle = \quad$ ).

To form the other function with  $M=1$ , the  $|1,1\rangle$  state, we must find another combination of  $|S=1, M_S=0\rangle |L=1, M_L=1\rangle$  and  $|S=1, M_S=1\rangle |L=1, M_L=0\rangle$  that is orthogonal to  $|2,1\rangle$  and is normalized. Since

$$|2,1\rangle = 2^{-1/2} [|S=1, M_S=0\rangle |L=1, M_L=1\rangle + |S=1, M_S=1\rangle |L=1, M_L=0\rangle],$$

we immediately see that the requisite function is

$$|1,1\rangle = 2^{-1/2} [|S=1, M_S=0\rangle |L=1, M_L=1\rangle - |S=1, M_S=1\rangle |L=1, M_L=0\rangle].$$

In the spin-orbital notation used above, this state is:

$$|1,1\rangle = [(\quad + \quad)3p_13p_0 - \{2\}^{1/2} \quad (3p_03p_0 + 3p_13p_{-1})]/2.$$

Thus far, we have found the  ${}^3P_J$  states with  $J=2, M=2$ ;  $J=2, M=1$ ; and  $J=1, M=1$ .

To find the  ${}^3P_J$  states with  $J=2, M=0$ ;  $J=1, M=0$ ; and  $J=0, M=0$ , we must once again apply the  $J$ -tool. In particular, we apply  $J$  to  $|2,1\rangle$  to obtain  $|2,0\rangle$  and we apply  $J$  to  $|1,1\rangle$  to obtain  $|1,0\rangle$ , each of which will be expressed in terms of  $|S=1, M_S=0\rangle |L=1, M_L=0\rangle$ ,  $|S=1, M_S=1\rangle |L=1, M_L=-1\rangle$ , and  $|S=1, M_S=-1\rangle |L=1, M_L=1\rangle$ . The  $|0,0\rangle$  state is then constructed to be a combination of these same product states which is orthogonal to  $|2,0\rangle$  and to  $|1,0\rangle$ . The results are as follows:

$$|J=2, M=0\rangle = 6^{-1/2}[2 |1,0\rangle |1,0\rangle + |1,1\rangle |1,-1\rangle + |1,-1\rangle |1,1\rangle],$$

$$|J=1, M=0\rangle = 2^{-1/2}[|1,1\rangle |1,-1\rangle - |1,-1\rangle |1,1\rangle],$$

$$|J=0, M=0\rangle = 3^{-1/2}[|1,0\rangle |1,0\rangle - |1,1\rangle |1,-1\rangle - |1,-1\rangle |1,1\rangle],$$

where, in all cases, a short hand notation has been used in which the  $|S, M_S\rangle |L, M_L\rangle$  product states have been represented by their quantum numbers with the spin function always appearing first in the product. To finally express all three of these new functions in terms of spin-orbital products it is necessary to give the  $|S, M_S\rangle |L, M_L\rangle$  products with  $M=0$  in terms of these products. For the spin functions, we have:

$$|S=1, M_S=1\rangle = \quad ,$$

$$|S=1, M_S=0\rangle = 2^{-1/2}(\quad + \quad).$$

$$|S=1, M_S=-1\rangle = \quad .$$

For the orbital product function, we have:

$$|L=1, M_L=1\rangle = 3p_13p_0 ,$$

$$|L=1, M_L=0\rangle = 2^{-1/2}(3p_03p_0 + 3p_13p_{-1}),$$

$$|L=1, M_L=-1\rangle = 3p_03p_{-1}.$$

### E. CG Coefficients and 3-j Symbols

As stated above, the CG coefficients can be worked out for any particular case using the raising and lowering operator techniques demonstrated above. Alternatively, as also stated above, the CG coefficients are tabulated (see, for example, Zare's book on angular momentum the reference to which is given earlier in this Appendix) for several values of  $j$ ,  $j'$ , and  $J$ .

An alternative to the CG coefficients is provided by the so-called 3-j coefficients (see Sec. 2.2 of Zare's book) which are defined in terms of the CG coefficients as follows:

$$\frac{j \quad j' \quad J}{m \quad m' \quad -M} = (-1)^{j-j'-M} \langle j, m; j', m' | J, M \rangle (2J+1)^{-1/2}.$$

Clearly, these coefficients contain no more or less information than do the CG coefficients. However, both sets of symbols have symmetries under interchange of the  $j$  and  $m$  quantum number that are more easily expressed in terms of the 3-j symbols. In particular, odd permutations of the columns of the 3-j symbol leave the magnitude unchanged and change the sign by  $(-1)^{j+j'+J}$ , whereas even permutations leave the value unchanged. Moreover, replacement of all of the  $m$ -values ( $m$ ,  $m'$ , and  $M$ ) by their negatives leave the magnitude the same and changes the sign by  $(-1)^{j+j'+J}$ . Table 2.5 in Zare's book (see above for reference) contains 3-j symbols for  $J=0, 1/2, 1, 3/2, \text{ and } 2$ .

## IV. How Angular Momentum Arises in Molecular Quantum Chemistry

### A. The Hamiltonian May Commute With Angular Momentum Operators

As is illustrated throughout this text, angular momentum operators often commute with the Hamiltonian of the system. In such cases, the eigenfunctions of the Hamiltonian can be made to also be eigenfunctions of the angular momentum operators. This allows one to label the energy eigenstates by quantum numbers associated with the angular momentum eigenvalues.

#### 1. Electronic Atomic Hamiltonia Without Spin-Orbit Coupling

For example, the electronic Hamiltonian of atoms, as treated in Chapters 1 and 3 in which only kinetic and coulombic interaction energies are treated, commutes with  $L^2$ , and  $L_z$ , where

$$L_z = \sum_j L_z(j)$$

and

$$L^2 = L_z^2 + L_x^2 + L_y^2.$$

The fact that  $H$  commutes with  $L_z$ ,  $L_x$ , and  $L_y$  and hence  $L^2$  is a result of the fact that the total coulombic potential energies among all the electrons and the nucleus is invariant to rotations of all electrons about the  $z$ ,  $x$ , or  $y$  axes ( $H$  does not commute with  $L_z(j)$  since if

only the  $j^{\text{th}}$  electron's coordinates are so rotated, the total coulombic potential is altered (because inter-electronic distances change). The invariance of the potential to rotations of all electrons is, in turn, related to the spherical nature of the atom. As a result, atomic energy levels for such a Hamiltonian can be labeled by their total L and M quantum numbers.

## 2. Electronic Linear-Molecule Hamiltonia Without Spin-Orbit Coupling

For linear molecules, the coulombic potential is unchanged (because the set of all inter-particle distances are unchanged) by rotations about the molecular axis (the z axis); hence H commutes with  $L_z$ . H does not commute with  $L_x$  or  $L_y$ , and thus not  $L^2$ , because the potential is altered by rotations about the x or y axes. As a result, linear-molecule energy levels for such a Hamiltonian can be labeled by their total M quantum number, which in this context is usually replaced by the quantum number  $\ell = |M|$ .

## 3. Spin-Orbit Effects

When spin-orbit couplings are added to the electrostatic Hamiltonian considered in the text, additional terms arise in H. These terms have the form of a one-electron additive operator:

$$H_{\text{SO}} = \sum_j \left\{ \frac{g_e}{2m_e c^2} \right\} r_j^{-1} \nabla / R_j \mathbf{S}(j) \cdot \mathbf{L}(j)$$

where V is the total coulombic potential that electron j feels due to the presence of the other electrons and the nuclei.  $\mathbf{S}(j)$  and  $\mathbf{L}(j)$  are the spin- and orbital- angular momentum operators of electron j, and  $g_e$  is the electron magnetic moment in Bohr magneton units ( $g_e = 2.002319$ ). For atoms in which these spin-orbit terms are considered (they are important for "heavy atoms" because  $r_j^{-1} \nabla / R_j$  varies as  $Z r_j^{-3}$  for atoms, whose expectation value varies as  $Z^4$ ), it turns out that neither  $L^2$  nor  $S^2$  commute with  $H_{\text{SO}}$ . However, the "combined" angular momentum

$$\mathbf{J} = \mathbf{L} + \mathbf{S}$$

$$J_z = L_z + S_z$$

$$J^2 = J_z^2 + J_x^2 + J_y^2$$

does commute with  $H_{\text{SO}}$ , and hence with the full  $H + H_{\text{SO}}$  Hamiltonian including spin-orbit coupling. For this reason, the eigenstates of atoms in which spin-orbit coupling is important can not be labeled by L, M, S, and  $M_S$ , but only by J and  $M_J$ .

## B. The Hamiltonian May Contain Angular Momentum Operators

### 1. Electronic Hamiltonia for Atoms Without Spin-Orbit Effects

There are cases in which the angular momentum operators themselves appear in the Hamiltonian. For electrons moving around a single nucleus, the total kinetic energy operator T has the form:

$$\begin{aligned} T &= \sum_j \left\{ -\frac{\hbar^2}{2m_e} \nabla_j^2 \right\} \\ &= \sum_j \left\{ -\frac{\hbar^2}{2m_e} \left[ \frac{\partial^2}{\partial r_j^2} + \frac{2}{r_j} \frac{\partial}{\partial r_j} + \frac{L_j^2}{r_j^2} \right] \right\} \end{aligned}$$

$$- (r_j^2 \sin^2 \theta_j)^{-1} / j(j+1) \\ - (r_j \sin \theta_j)^{-2} / j^2 \}.$$

The factor  $\hbar^2 [( \sin \theta_j )^{-1} / j(j+1) + ( \sin \theta_j )^{-2} / j^2 ]$  is  $L^2(j)$ , the square of the angular momentum for the  $j^{\text{th}}$  electron. In this case, the Hamiltonian contains  $L^2(j)$  for the individual electrons, not the total  $L^2$ , although it still commutes with the total  $L^2$  (which thus renders  $L$  and  $M$  good quantum numbers).

## 2. Linear Rigid-Molecule Rotation

The rotational Hamiltonian for a diatomic molecule as given in Chapter 3 is

$$H_{\text{rot}} = \hbar^2 / 2\mu \{ (R^2 \sin^2 \theta)^{-1} / j(j+1) + (R^2 \sin^2 \theta)^{-1} / j^2 \},$$

where  $\mu$  is the reduced mass of the molecule, and  $R$  is its bond length.

Again, the square of the total rotational angular momentum operator appears in  $H_{\text{rot}}$

$$H_{\text{rot}} = L^2 / 2\mu R^2.$$

In this case, the Hamiltonian both contains and commutes with the total  $L^2$ ; it also commutes with  $L_z$ , as a result of which  $L$  and  $M$  are both good quantum numbers and the spherical harmonics  $Y_{L,M}(\theta, \phi)$  are eigenfunctions of  $H$ . These eigenfunctions obey orthogonality relations:

$$\int_0^{2\pi} \int_0^\pi (Y_{L,M}^*(\theta, \phi) Y_{L',M'}(\theta, \phi) \sin \theta d\theta d\phi) = \delta_{L,L'} \delta_{M,M'}$$

because they are eigenfunctions of two Hermitian operators ( $L^2$  and  $L_z$ ) with (generally) different eigenvalues.

## 3. Non-Linear Molecule Rotation

For non-linear molecules, when treated as rigid (i.e., having fixed bond lengths, usually taken to be the equilibrium values or some vibrationally averaged values), the rotational Hamiltonian can be written in terms of rotation about three axes. If these axes ( $X, Y, Z$ ) are located at the center of mass of the molecule but fixed in space such that they do not move with the molecule, then the rotational Hamiltonian can be expressed as:

$$H_{\text{rot}} = 1/2 \sum_{K,K'} I_{K,K'} \omega_K \omega_{K'}$$

where  $\omega_K$  is the angular velocity about the  $K^{\text{th}}$  axis and

$$I_{K,K} = \sum_j m_j (R_j^2 - R_{K,j}^2) \quad (\text{for } K = K')$$

$$I_{K,K'} = - \sum_j m_j R_{K,j} R_{K',j} \quad (\text{for } K \neq K')$$

are the elements of the so-called moment of inertia tensor. This tensor has components along the axes labeled  $K$  and  $K'$  (each of which runs over  $X$ ,  $Y$ , and  $Z$ ). The  $m_j$  denote the masses of the atoms (labeled  $j$ ) in the molecule,  $R_{K,j}$  is the coordinate of atom  $j$  along the  $K$ -axis relative to the center of mass of the molecule, and  $R_j$  is the distance of atom  $j$  from the center of mass ( $R_j^2 = \sum_K (R_{K,j})^2$ ).

Introducing a new set of axes  $x$ ,  $y$ ,  $z$  that also have their origin at the center of mass, but that rotate with the molecule, it is possible to reexpress  $H_{\text{rot}}$  in terms of motions of these axes. It is especially useful to choose a particular set of such molecule-fixed axes, those that cause the moment of inertia tensor to be diagonal. This symmetric matrix can, of course, be made diagonal by first computing  $I_{k,k'}$  (where  $k$  and  $k'$  run over  $x$ ,  $y$ , and  $z$ ) for an arbitrary  $x$ ,  $y$ ,  $z$  axis choice and then finding the orthogonal transformation (i.e., the eigenvectors of the  $\mathbf{I}$  matrix) that brings  $\mathbf{I}$  to diagonal form. Such molecule-fixed axes (which we denote as  $a$ ,  $b$ , and  $c$ ) in which  $\mathbf{I}$  is diagonal are called **principal axes**; in terms of them,  $H_{\text{rot}}$  becomes:

$$H_{\text{rot}} = 1/2 [ I_a \dot{\alpha}^2 + I_b \dot{\beta}^2 + I_c \dot{\gamma}^2 ].$$

The angular momentum conjugate to each of these three angular coordinates (each  $\dot{\alpha}$  is the time rate of change of an angle of rotation about an axis:  $\dot{\alpha} = d(\text{angle})/dt$ ) is obtained, as usual, from the Lagrangian function  $L = T - V$  of classical mechanics:

$$p = \partial L / \partial \dot{q} = (\text{Kinetic Energy} - \text{Potential Energy}) / (dq/dt)$$

or (using  $J_a$  to denote the angular momentum conjugate to  $\alpha$  and realizing that since this free rotational motion has no potential energy,  $L = T = H_{\text{rot}}$ )

$$J_a = \partial H_{\text{rot}} / \partial \dot{\alpha} = I_a \dot{\alpha}$$

$$J_b = I_b \dot{\beta}$$

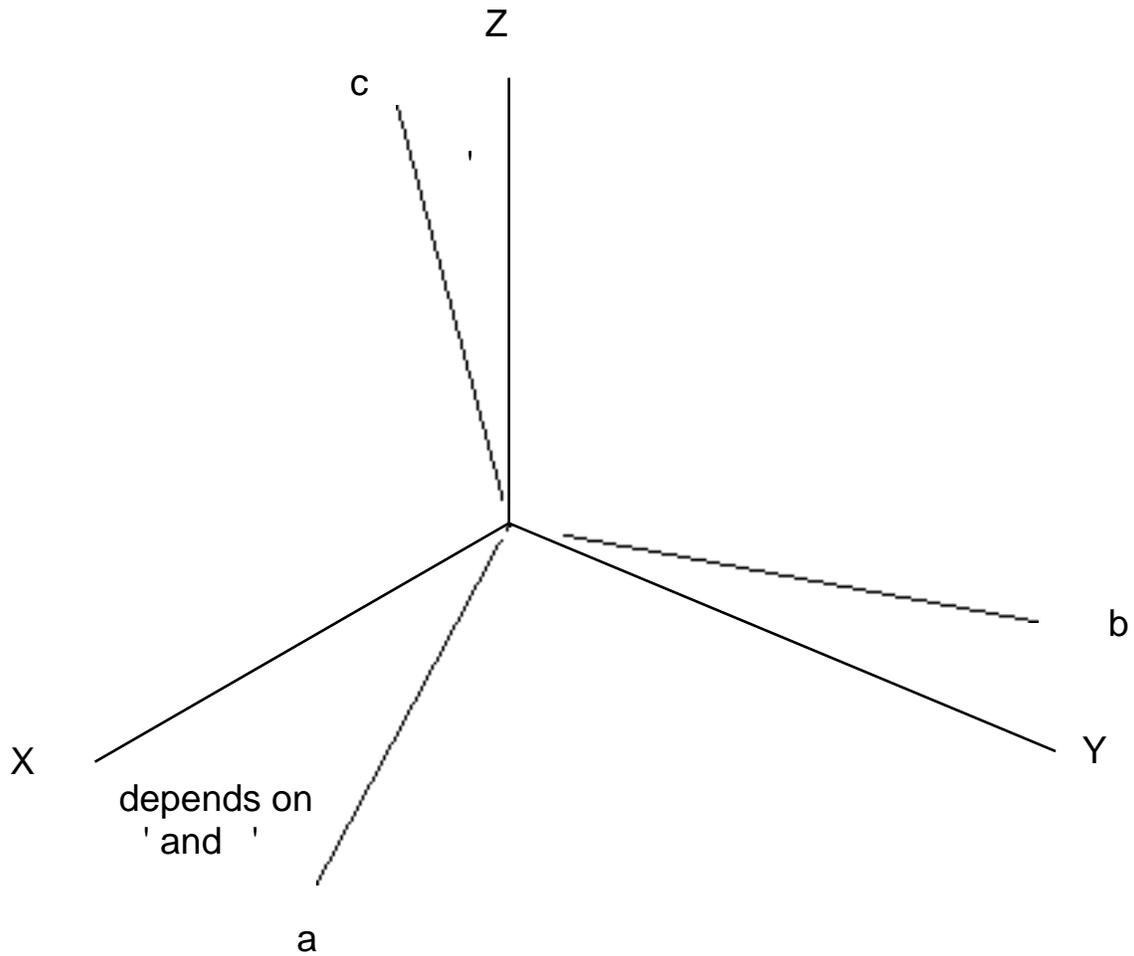
$$J_c = I_c \dot{\gamma}.$$

The rotational Hamiltonian can then be written in terms of angular momenta and principal-axis moments of inertia as:

$$H_{\text{rot}} = J_a^2/2I_a + J_b^2/2I_b + J_c^2/2I_c.$$

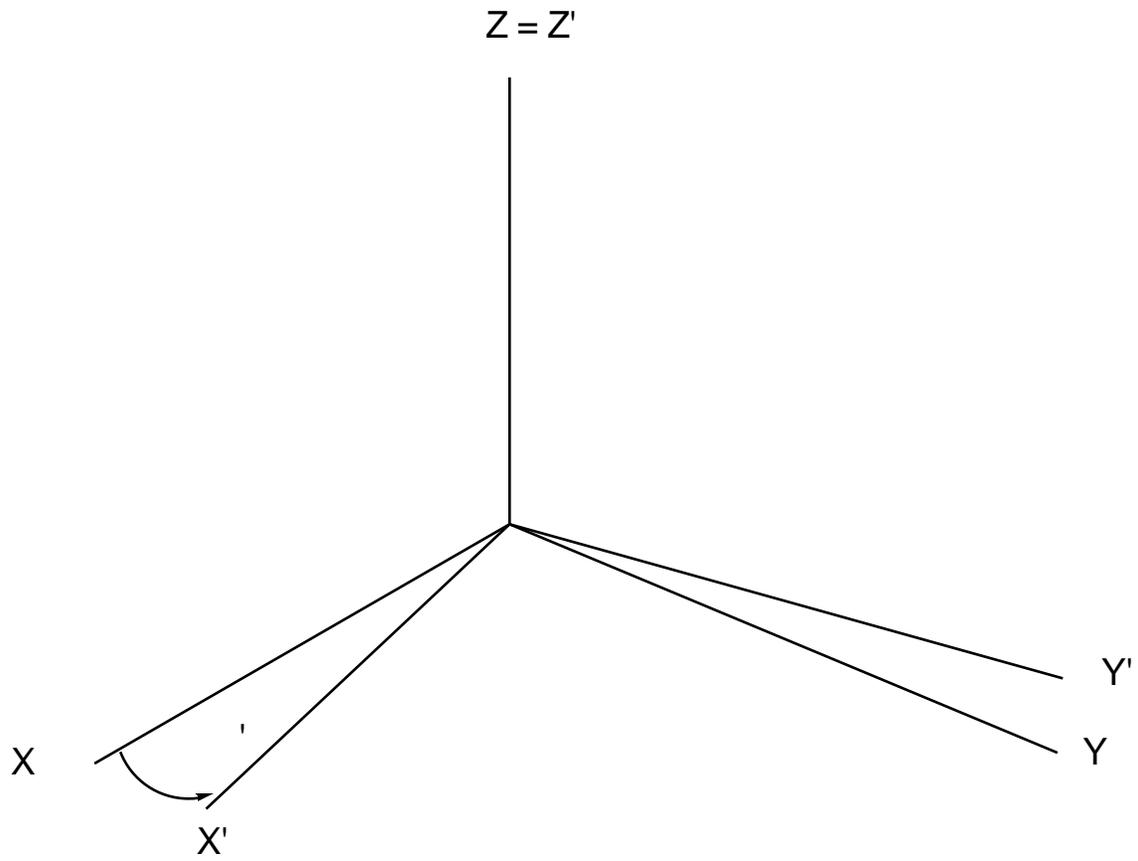
With respect to this principal axis point of view, the rotation of the molecule is described in terms of three angles (it takes three angles to specify the orientation of such a rigid body) that detail the movement of the  $a$ ,  $b$ , and  $c$  axes relative to the lab-fixed  $X$ ,  $Y$ , and  $Z$  axes. It is convention to call these angles  $\theta$  (which can be viewed as the angle

between the lab-fixed Z axis and one of the principal axes- say c- in the molecule),  $\theta$ , and  $\phi$ . The volume element for integration over these three angles is  $\sin \theta \, d\theta \, d\phi \, d\chi$ , with  $\theta$  and  $\phi$  running between 0 and  $2\pi$ , and  $\chi$  going from 0 to  $\pi$ . These coordinates are described visually below.

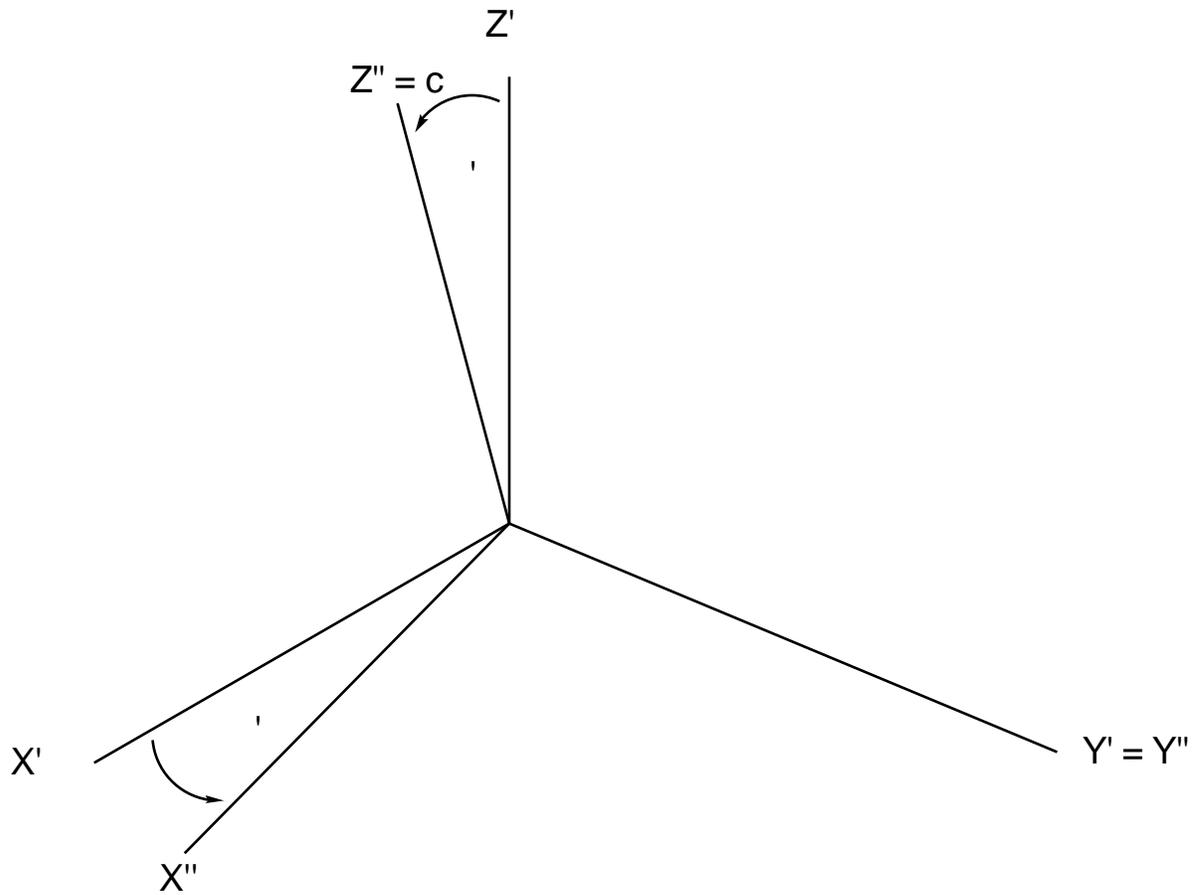


The a, b, c, coordinate system can be formed by beginning with the original X, Y, Z system and sequentially:

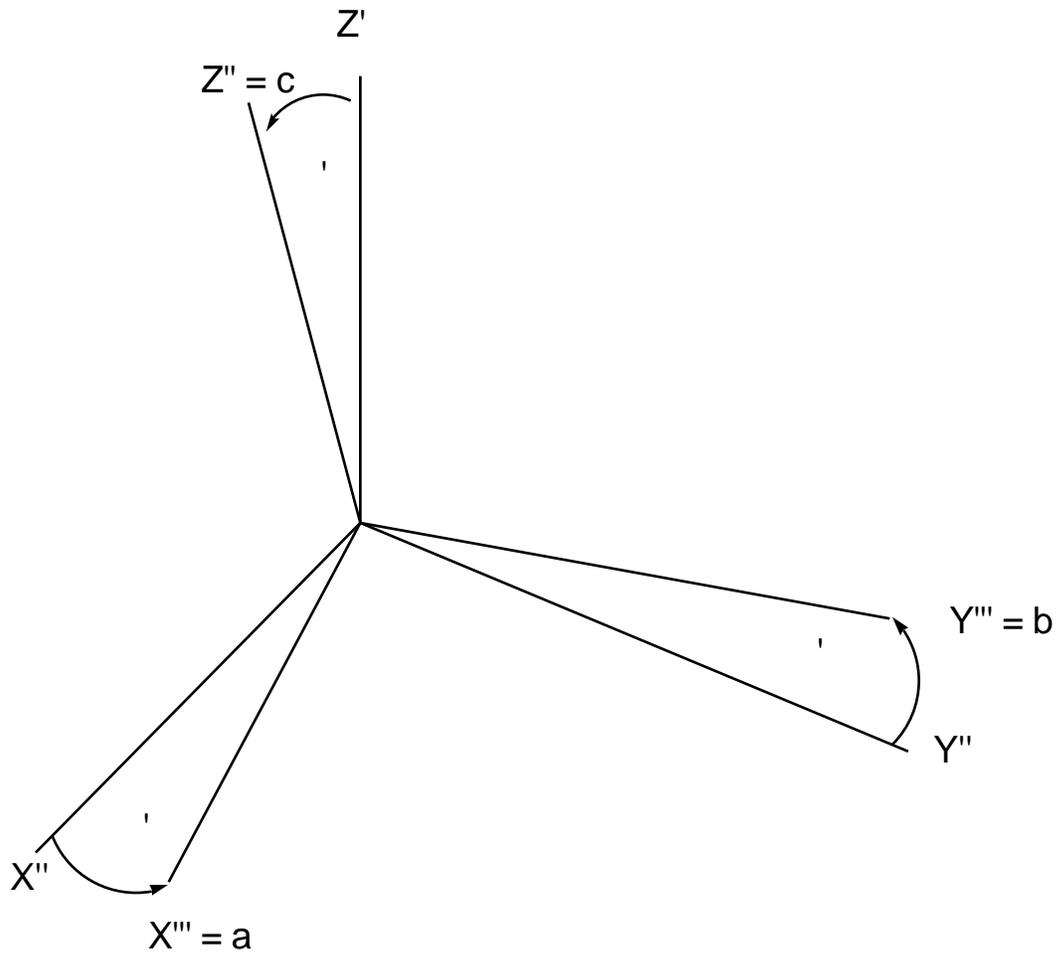
- (i) rotating about the Z axis by an amount  $\phi$ , to generate intermediate X' Y', and Z = Z' axes (X' and Y' being rotated by  $\phi$  relative to X and Y);



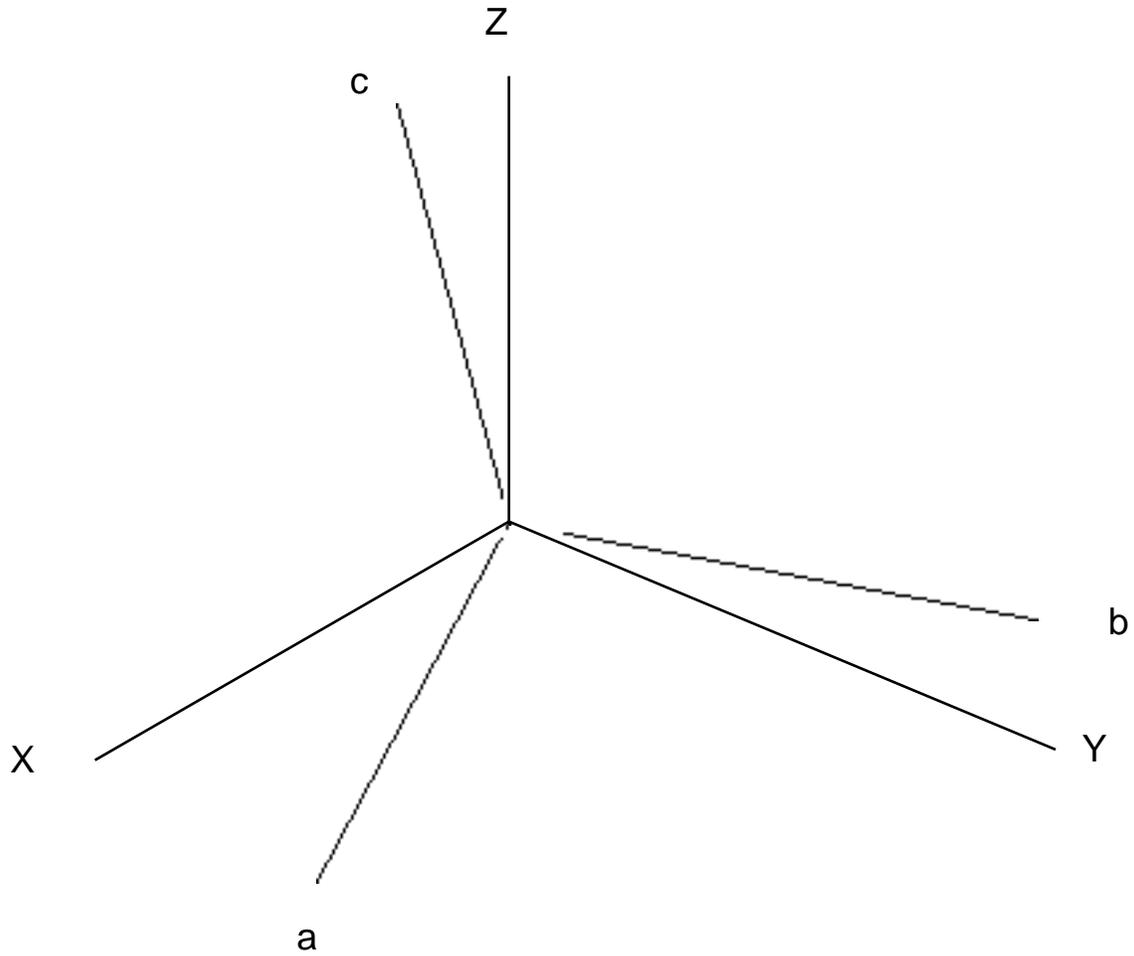
(ii) next rotating about the  $Y'$  axis by an amount  $\phi$ , to generate  $X''$ ,  $Y'$ , and  $Z'' = z$  axes



(iii) and finally rotating about the new  $Z'' = c$  axis by an amount  $\beta$  to generate the final  $X''' = a$  and  $Y''' = b$  axes



Thus, the original and final coordinates can be depicted as follows:



The explicit expressions for the components of the quantum mechanical angular momentum operators along the three new axes are:

$$J_a = -i\hbar \cos \theta \left[ \cot \theta \frac{\partial}{\partial \phi} - (\sin \theta)^{-1} \frac{\partial}{\partial \alpha} \right] - i\hbar \sin \theta \frac{\partial}{\partial \beta}$$

$$J_b = i\hbar \sin \theta \left[ \cot \theta \frac{\partial}{\partial \phi} - (\sin \theta)^{-1} \frac{\partial}{\partial \alpha} \right] - i\hbar \cos \theta \frac{\partial}{\partial \beta}$$

$$J_c = -i\hbar \frac{\partial}{\partial \alpha}$$

The corresponding total angular momentum operator  $J^2$  can be obtained as

$$J^2 = J_a^2 + J_b^2 + J_c^2$$

$$= \hbar^2 \left[ \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \alpha^2} + \frac{\partial^2}{\partial \beta^2} - \cot \theta \frac{\partial}{\partial \theta} \right]$$

$$- (1/\sin \theta) ( \frac{2}{I_a} J_a^2 + \frac{2}{I_c} J_c^2 - 2 \cos \theta \frac{2}{I_a} J_a J_c ),$$

and the component along the original Z axis  $J_Z$  is still  $- i\hbar / \theta$ .

Returning now to the rigid-body rotational Hamiltonian shown above, there are two special cases for which exact eigenfunctions and energy levels can be found using the general properties of angular momentum operators.

#### a. Spherical and Symmetric Top Energies

The special cases for which  $I_a = I_b = I_c$  (the **spherical top**) and for which  $I_a = I_b > I_c$  (the oblate symmetric top) or  $I_a > I_b = I_c$  (the prolate symmetric top) are covered in Chapter 3. In the former case, the rotational Hamiltonian can be expressed in terms of  $J^2 = J_a^2 + J_b^2 + J_c^2$  because all three moments of inertia are identical:

$$H_{rot} = J^2/2I,$$

as a result of which the eigenfunctions of  $H_{rot}$  are those of  $J^2$  (and  $J_a$  as well as  $J_Z$  both of which commute with  $J^2$  and with one another;  $J_Z$  is the component of  $\mathbf{J}$  along the lab-fixed Z-axis and commutes with  $J_a$  because  $J_Z = - i\hbar / \theta$  and  $J_a = - i\hbar / \theta$  act on different angles). The energies associated with such eigenfunctions are

$$E(J,K,M) = \hbar^2 J(J+1)/2I^2,$$

for all K (i.e.,  $J_a$  quantum numbers) ranging from  $-J$  to  $J$  in unit steps and for all M (i.e.,  $J_Z$  quantum numbers) ranging from  $-J$  to  $J$ . Each energy level is therefore  $(2J + 1)^2$  **degenerate** because there are  $2J + 1$  possible K values and  $2J + 1$  M values for each J.

In the **symmetric top** cases,  $H_{rot}$  can be expressed in terms of  $J^2$  and the angular momentum along the axis with the unique moment of inertia (denoted the a-axis for prolate tops and the c-axis of oblate tops):

$$H_{rot} = J^2/2I + J_a^2\{1/2I_a - 1/2I\}, \text{ for prolate tops}$$

$$H_{rot} = J^2/2I + J_c^2\{1/2I_c - 1/2I\}, \text{ for oblate tops.}$$

$H_{rot}$ , along with  $J^2$  and  $J_a$  (or  $J_c$  for oblate tops) and  $J_Z$  (the component of  $\mathbf{J}$  along the lab-fixed Z-axis) form a mutually commutative set of operators.  $J_Z$ , which is  $- i\hbar / \theta$ , and  $J_a$  (or c), which is  $- i\hbar / \theta$ , commute because they act on different angles. As a result, the eigenfunctions of  $H_{rot}$  are those of  $J^2$  and  $J_a$  or  $J_c$  (and of  $J_Z$ ), and the corresponding energy levels are:

$$E(J,K,M) = \hbar^2 J(J+1)/2I^2 + \hbar^2 K^2 \{1/2I_a - 1/2I\},$$

for prolate tops

$$E(J,K,M) = \hbar^2 J(J+1)/2I^2 + \hbar^2 K^2 \{1/2I_c - 1/2I\},$$

for oblate tops,

again for K and M (i.e.,  $J_a$  or  $J_c$  and  $J_z$  quantum numbers, respectively) ranging from -J to J in unit steps. Since the energy now depends on K, these levels are only **2J + 1 degenerate** due to the 2J + 1 different M values that arise for each J value.

### b. Spherical and Symmetric Top Wavefunctions

The eigenfunctions of  $J^2$ ,  $J_a$  (or  $J_c$ ) and  $J_z$  clearly play important roles in polyatomic molecule rotational motion; they are the eigenstates for spherical-top and symmetric-top species, and they can be used as a basis in terms of which to expand the eigenstates of asymmetric-top molecules whose energy levels do not admit an analytical solution. These eigenfunctions  $|J,M,K\rangle$  are given in terms of the set of so-called "**rotation matrices**" which are denoted  $D_{J,M,K}$ :

$$|J,M,K\rangle = \sqrt{\frac{2J+1}{8}} D_{J,M,K}^*(\alpha, \beta, \gamma).$$

They obey

$$J^2 |J,M,K\rangle = \hbar^2 J(J+1) |J,M,K\rangle,$$

$$J_a \text{ (or } J_c \text{ for oblate tops)} |J,M,K\rangle = \hbar K |J,M,K\rangle,$$

$$J_z |J,M,K\rangle = \hbar M |J,M,K\rangle.$$

It is demonstrated below why the symmetric and spherical top wavefunctions are given in terms of these  $D_{J,M,K}$  functions.

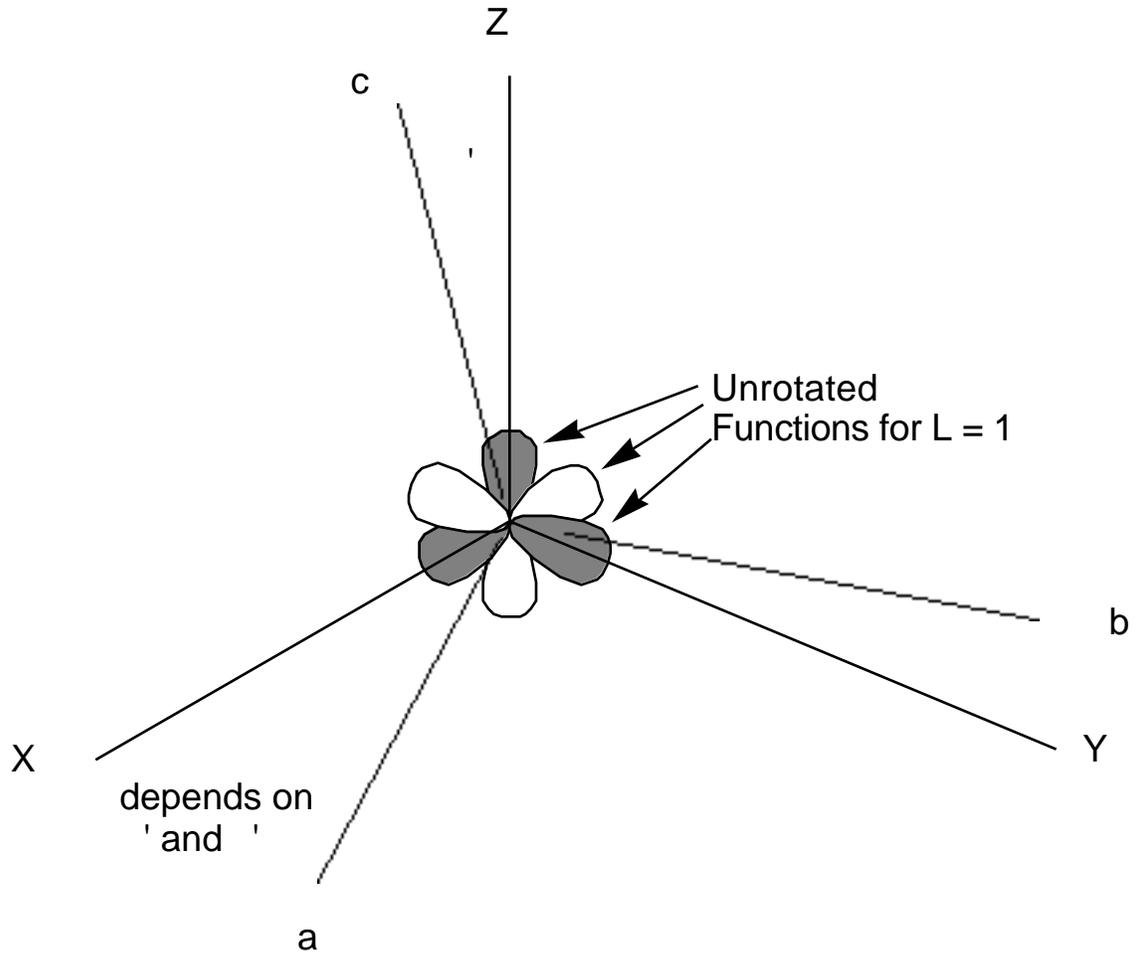
### c. Rotation Matrices

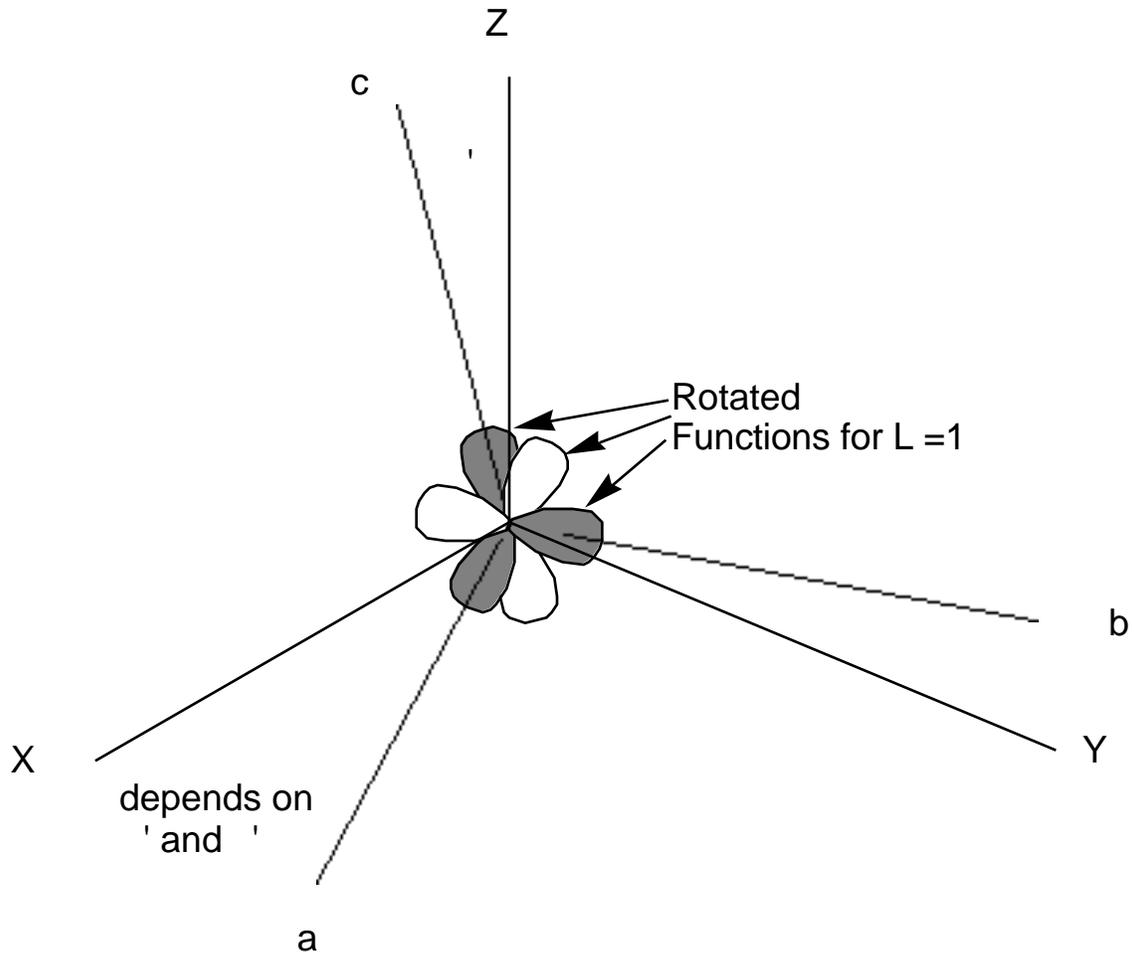
These same rotation matrices arise when the transformation properties of spherical harmonics are examined for transformations that rotate coordinate systems. For example, given a spherical harmonic  $Y_{L,M}(\alpha, \beta)$  describing the location of a particle in terms of polar angles  $\alpha, \beta$  within the X, Y, Z axes, one might want to rotate this function by Euler angles  $\alpha', \beta', \gamma'$  and evaluate this rotated function at the same physical point. As shown in Zare's text on angular momentum, the rotated function  $Y_{L,M}$  evaluated at the angles  $\alpha, \beta$  can be expressed as follows:

$$Y_{L,M}(\alpha, \beta) = \sum_{M'} D_{L,M',M}(\alpha', \beta', \gamma') Y_{L,M'}(\alpha, \beta).$$

In this form, one sees why the array  $D_{J,M',M}$  is viewed as a unitary matrix, with  $M'$  and  $M$  as indices, that describes the effect of rotation on the set of functions  $\{Y_{L,M}\}$ . This mapping from the unrotated set  $\{Y_{L,M}\}$  into the rotated set of functions  $\{Y_{L,M}\}$  must be

**unitary** if the sets  $\{Y_{L,M}\}$  and  $\{Y'_{L,M}\}$  are both orthonormal. The unitary matrix carries an additional index ( $L$  in this example) that details the dimension ( $2L + 1$ ) of the space of functions whose transformations are so parameterized. An example, for  $L = 1$ , of a set of unrotated and rotated functions is shown below.





#### d. Products of Rotation Matrices

An identity that proves very useful when treating coupled angular momenta that are subjected to rotations of the axes with respect to which their eigenfunctions are quantized can be derived by combining the above result:

$$Y_{L,M}(\theta, \phi) = \sum_{M'} D_{L,M',M}(\theta, \phi, \psi) Y_{L,M'}(\theta, \phi)$$

and the expression for coupling two angular momenta:

$$|J,M\rangle = \sum_{m,n} \langle j,m;l,n|J,M\rangle |j,m\rangle |l,n\rangle.$$

Applying the rotation to the left and right sides of the equation defining  $|J,M\rangle$ , gives:

$$\begin{aligned} & \sum_{M'} D_{J,M,M'}(\theta, \phi, \psi) |J,M'\rangle \\ &= \sum_{m',n'} \langle j,m';l,n'|J,M\rangle D_{j,m,m'}(\theta, \phi, \psi) D_{l,n,n'}(\theta, \phi, \psi) |j,m'\rangle |l,n'\rangle. \end{aligned}$$

Multiplying both sides of this equation by  $\langle J, M' |$  and using the orthonormality of the angular momentum eigenfunctions gives:

$$D_{J, M, M'} = \sum_{m', n'} \langle j, m; l, n | J, M \rangle D_{j, m, m'} D_{l, n, n'} \langle J, M' | j, m'; l, n' \rangle.$$

This result expresses one  $D_{J, M, M'}$  in terms of sums of products of D matrix elements for angular momenta  $j, m, m'$  and  $l, n, n'$  that can be coupled to for  $J, M, M'$ .

If the above series of operations is applied to the angular momentum coupling expression in the form:

$$|j, m\rangle |l, n\rangle = \sum_{J, M} \langle J, M | j, m; l, n \rangle |J, M\rangle,$$

one can obtain:

$$\begin{aligned} & \sum_{m', n'} D_{j, m, m'}( \quad ) D_{l, n, n'}( \quad ) |j, m'\rangle |l, n'\rangle \\ &= \sum_{J, M} \langle J, M | j, m; l, n \rangle \sum_{M'} D_{J, M, M'}( \quad ) |J, M'\rangle. \end{aligned}$$

Multiplying by  $\langle j, m' | \langle l, n' |$  then yields:

$$\begin{aligned} & D_{j, m, m'} D_{l, n, n'} \\ &= \sum_{J, M, M'} \langle J, M | j, m; l, n \rangle \langle j, m'; l, n' | J, M' \rangle D_{J, M, M'} \end{aligned}$$

which expresses the product of two D matrices as a sum of D matrices whose angular momentum indices are related to those of the product.

#### e. Rigid Body Rotational Wavefunctions as Rotation Matrices

This same analysis can be used to describe how a set of functions  $Y_{J, M}(\theta, \phi)$  (labeled by a total angular momentum quantum number that determines the number of functions in the set and an  $M$  quantum number that labels the Z-axis projection of this angular momentum) that are functions of three coordinates  $\theta, \phi, \chi$ , transform under rotation. In particular, one obtains a result analogous to the spherical harmonic expression:

$$Y_{J, M}(\theta, \phi) = \sum_{M'} D_{J, M', M}(\alpha, \beta, \gamma) Y_{J, M'}(\theta, \phi).$$

Here  $Y_{J, M}(\theta, \phi)$  is the original unrotated function evaluated at a point whose angular coordinates are  $\theta, \phi$ ;  $\alpha, \beta, \gamma$  are the Euler angles through which this function is rotated to obtain the rotated function  $Y_{J, M}$  whose value at the above point is denoted  $Y_{J, M}(\theta, \phi)$ .

Now, if the angles  $\alpha, \beta, \gamma$  through which the original function is rotated were chosen to equal the angular coordinates  $\theta, \phi$  of the point discussed here, then the rotated function  $Y_{J, M}$  evaluated at this point could easily be identified. Its value would be

nothing more than the unrotated function  $Y_{J,M}$  evaluated at  $\alpha = 0, \beta = 0, \gamma = 0$ . In this case, we can write:

$$Y_{J,M}(\alpha, \beta, \gamma) = Y_{J,M}(0, 0, 0) = \sum_{M'} D_{J,M',M}(\alpha, \beta, \gamma) Y_{J,M'}(\alpha, \beta, \gamma).$$

Using the unitary nature of the  $D_{L,M',M}$  array, this equation can be solved for the  $Y_{J,M'}(\alpha, \beta, \gamma)$  functions:

$$Y_{J,M'}(\alpha, \beta, \gamma) = \sum_M D_{J,M',M}^*(\alpha, \beta, \gamma) Y_{J,M}(0, 0, 0).$$

This result shows that the functions that describe the rotation of a rigid body through angles  $\alpha, \beta, \gamma$  must be a combination of rotation matrices (actually  $D_{L,M',M}^*(\alpha, \beta, \gamma)$  functions). Because of the normalization of the  $D_{L,M,M'}(\alpha, \beta, \gamma)$  functions:

$$\begin{aligned} & (D_{L',M',K'}^*(\alpha, \beta, \gamma) D_{L,M,K}(\alpha, \beta, \gamma) \sin^2 \alpha \sin^2 \beta \sin^2 \gamma) \\ &= \frac{8}{2L+1} \delta_{L,L'} \delta_{M,M'} \delta_{K,K'} \end{aligned}$$

the properly normalized rotational functions that describe spherical or symmetric tops are:

$$|J,M,K\rangle = \sqrt{\frac{2J+1}{8}} D_{J,M,K}^*(\alpha, \beta, \gamma)$$

as given above. For asymmetric top cases, the correct eigenstates are combinations of these  $\{|J,M,K\rangle\}$  functions:

$$Y_{J,M}(\alpha, \beta, \gamma) = \sum_K \sqrt{\frac{2J+1}{8}} D_{J,M,K}^*(\alpha, \beta, \gamma) C_K$$

with amplitudes  $\{C_K\}$  determined by diagonalizing the full  $H_{\text{rot}}$  Hamiltonian within the basis consisting of the set of

$$\sqrt{\frac{2J+1}{8}} D_{J,M,K}^*(\alpha, \beta, \gamma)$$

functions.

#### 4. Electronic and Nuclear Zeeman Interactions

When magnetic fields are present, the intrinsic spin angular momenta of the electrons  $\mathbf{S}(j)$  and of the nuclei  $\mathbf{I}(k)$  are affected by the field in a manner that produces additional energy contributions to the total Hamiltonian  $H$ . The Zeeman interaction of an external magnetic field (e.g., the earth's magnetic field of 4. Gauss or that of a NMR

machine's magnet) with such intrinsic spins is expressed in terms of the following contributions to  $H$ :

$$H_{\text{Zeeman}} = (g_e e/2m_e c) \sum_j S_z(j) H - (e/2m_p c) \sum_k g_k I_z(k) H.$$

Here  $g_k$  is the so-called nuclear g-value of the  $k^{\text{th}}$  nucleus,  $H$  is the strength of the applied field,  $m_p$  is the mass of the proton,  $g_e$  is the electron magnetic moment, and  $c$  is the speed of light. When chemical shieldings (denoted  $\sigma_k$ ), nuclear spin-spin couplings (denoted  $J_{k,l}$ ), and electron-nuclear spin couplings (denoted  $a_{j,k}$ ) are considered, the following spin-dependent Hamiltonian is obtained:

$$H = (g_e e/2m_e c) \sum_j S_z(j) H - (e/2m_p c) \sum_k g_k (1 - \sigma_k) I_z(k) H + h \sum_{j,k} (a_{j,k}/\hbar^2) \mathbf{I}(k) \cdot \mathbf{S}(j) + h \sum_{k,l} (J_{k,l}/\hbar^2) \mathbf{I}(k) \cdot \mathbf{I}(l).$$

Clearly, the treatment of electron and nuclear spin angular momenta is essential to analyzing the energy levels of such Hamiltonians, which play a central role in NMR and ESR spectroscopy.

## QMIC program descriptions

### Appendix H. QMIC Programs

The Quantum Mechanics in Chemistry (QMIC) programs whose source and executable versions are provided along with the text are designed to be pedagogical in nature; therefore they are not designed with optimization in mind, and could certainly be improved by interested students or instructors. The software is actually a suite of programs allowing the student to carry out many different types of *ab initio* calculations. The student can perform Hartree-Fock, MP2, or CI calculations, in a single step or by putting together a series of steps, by running the programs provided. The software can be found on the world wide web in several locations:

- at the University of Utah, located at: <http://www.chem.utah.edu>
- at the Pacific Northwest National Laboratory, located at:  
[http://www.emsl.pnl.gov:2080/people/bionames/nichols\\_ja.html](http://www.emsl.pnl.gov:2080/people/bionames/nichols_ja.html)
- at the Oxford University Press, located at: <http://www.oup-usa.org>

These programs are designed to run in very limited environments (e.g. memory, disk, and CPU power). With the exception of "integral.f" all are written in single precision and use minimal memory (less than 640K) in most instances. The programs are designed for simple systems e.g., only a few atoms (usually less than 8) and small basis sets. They do not use group symmetry, and they use simple Slater determinants rather than spin-adapted configuration state functions to perform the CI. The programs were all originally developed and run on an IBM RISC System 6000 using AIX v3.2 and Fortran compilers xlf v2 and v3. All routines compile untouched with gnu compilers and utilities for workstations and PCs. The gnu utilities were obtained from the ftp server: <ftp://ftp.coast.net> in directory: [simtel/vendors/gnu](ftp://ftp.coast.net/simtel/vendors/gnu). Except for very minor modifications all run untouched when compiled using Language Systems Fortran for the Macintosh. The intrinsics "and", "xor", and "rshift" have to be replaced by their counterparts "iand", "ixor", and "ishft". These intrinsic functions are only used in program hamilton.f and their replacement functions are detailed and commented in the hamilton program source. No floating point unit has been turned on in the compilation. Because of this, computations on chemical systems with lots of basis functions performed on an old Mac SE can be tiring (the N<sup>5</sup> processes like the transformation can take as long as a half hour on these systems). Needless to say all of these run in less than a minute on the fancier workstations. Special thanks goes to Martin Feyereisen (Cray Research) for supplying us with very compact subroutines which evaluate one- and two-electron integrals in a very simple and straight forward manner. Brief descriptions of each of the programs in QMIC follow:

#### Current QMIC program limits:

- Maximum number of atoms: 8
- Maximum number of orbitals: 26
- Maximum number of shells: 20
- Maximum number of primitives per shell: 7
- Maximum orbital angular momentum: 1
- Maximum number of active orbitals in the CI: 15
- ! xt• Maximum number of determinants: 350
- Maximum matrix size (row or column): 350

Program INTEGRAL This program is designed to calculate one- and two-electron AO integrals and to write them out to disk in canonical order (in Dirac <12|12> convention). It is designed to handle only S and P orbitals. With the program limitations described above, INTEGRAL memory usage is 542776 bytes.

Program MOCOEF This program is designed to read in (from the keyboard) the LCAO-MO coefficient matrix and write it out to disk. Alternatively, you can choose to have a unit matrix (as your initial guess) put out to disk. With the program limitations described above, MOCOEF memory usage is 2744 bytes.

Program FNCT\_MAT This program is designed to read in a real square matrix, perform a function on it, and return this new array. Possible functions, using X as the input matrix, are:

- (1)  $X^{(-1/2)}$ , NOTE: X must be real symmetric, and positive definite.
- (2)  $X^{(+1/2)}$ , NOTE: X must be real symmetric, and positive definite.
- (3)  $X^{(-1)}$ , NOTE: X must be real symmetric, and have non-zero eigenvalues.
- (4) a power series expansion of a matrix to find the transformation matrix:  

$$U = \exp(X) = 1 + X + X^{**2}/2! + X^{**3}/3! + \dots + X^{**N}/N!$$

With the program limitations described above, FNCT\_MAT memory usage is 1960040 bytes.

Program FOCK This program is designed to read in the LCAO-MO coefficient matrix, the one- and two-electron AO integrals and to form a closed shell Fock matrix (i.e., a Fock matrix for species with all doubly occupied orbitals). With the program limitations described above, FOCK memory usage is 255256 bytes.

Program UTMATU This program is designed to read in a real matrix, A, a real transformation matrix, B, perform the ! transformation:  $X = B(\text{transpose}) * A * B$ , and output the result. With the program limitations described above, UTMATU memory usage is 1960040 bytes.

Program DIAG This program is designed to read in a real symmetric matrix (but as a square matrix on disk), diagonalize it, and return all eigenvalues and corresponding eigenvectors. With the program limitations described above, DIAG memory usage is 738540 bytes.

Program MATXMAT This program is designed to read in two real matrices; A and B, and to multiply them together:  $AB = A * B$ , and output the result. With the program limitations described above, MATXMAT memory usage is 1470040 bytes.

Program FENERGY This program is designed to read in the LCAO-MO coefficient matrix, the one- and two-electron AO integrals (in Dirac <12|12> convention), and the Fock orbital energies. Upon transformation of the one- and two-electron integrals from the AO to the MO basis, the closed shell Hartree - Fock energy is calculated in two ways. First, the energy is calculated with the MO integrals,

$$\sum_k (2 * \langle k|h|k \rangle + \sum_{(k,l)} (2 * \langle k,l|k,l \rangle - \langle k,l|l,k \rangle) + \sum_u Z_u Z_v / R_{uv}).$$

Secondly, the energy is calculated with the Fock orbital energies and one electron energies in the MO basis,

$$\text{Sum}(k) (\text{eps}(k) + \langle k|h|k\rangle) + \sum_k Z_k V_k / R_{uv}$$

With the program limitations described above, FENERGY memory usage is 1905060 bytes..

Program TRANS This program is designed to read in the LCAO-MO coefficient matrix, the one- and two-electron AO integrals (in Dirac  $\langle 12|12\rangle$  convention), and to transform the integrals from the AO to the MO basis, and write these MO integrals to a file. With the program limitations described above, TRANS memory usage is 1905060 bytes.

Program SCF This program is designed to read in the LCAO-MO coefficient matrix (or generate one), the one- and two-electron AO integrals and form a closed shell Fock matrix (i.e., a Fock matrix for species with all doubly occupied orbitals). It then solves the Fock equations; iterating until convergence to six significant figures in the energy expression. A modified damping algorithm is used to insure convergence. With the program limitations described above, SCF memory usage is 259780 bytes.

Program MP2 This program is designed to read in the transformed one- and two-electron integrals and the Fock orbital energies after which it will compute the second order Moller Plesset perturbation theory energy (MP2). With the program limitations described above, MP2 memory usage is 250056 bytes.

Program HAMILTON This program is designed to generate or read in a list of determinants. You can generate determinants for a CAS (Complete Active Space) of orbitals or you can input your own list of determinants. Next, if you wish, you may read in the one- and two-electron MO integrals and form a Hamiltonian matrix over the determinants. Finally, if you so choose, you may diagonalize the Hamiltonian matrix constructed over the determinants generated. With the program limitations described above, HAMILTON memory usage is 988784 bytes.

Program RW\_INTS This program is designed to read the one- and two-electron AO integrals (in Dirac  $\langle 12|12\rangle$  convention) from user input and put them out to disk in canonical order. There are no memory limitations associated with program RW\_INTS.

QMLIB This is a library of subroutines and functions which are used by the QMIC programs.

"limits.h" This is an include file containing ALL the parameters which determine memory requirements for the QMIC programs.

Makefile There are a few versions of Makefiles available: a generic Makefile (Makefile.gnu) which works with Gnu make on a unix box, a Makefile (Makefile.486) which was used to make the programs on a 486 PC using other Gnu utilities like "f2c", "gcc", etc. and a Makefile (Makefile.mac) which was used on the Macintosh.

BasisLib This is a library file which contains gaussian atomic orbital basis sets for Hydrogen - Neon. The basis sets available to choose from are:

- 1.) STO3G by Hehre, Stewart, and Pople, JCP, 51, 2657 (1969).
- 2.) 3-21G by Brinkley, Pople, and Hehre, JACS, 102, 939 (1980).
- 3.) [3s2p] by Dunning and Hay in: Modern Theoretical Chemistry Vol 3, Henry F. Schaefer III, Ed., 1977, Plenum Press, NY.



