Modeling and Visualizing Surfaces

Thesis by

Robert Edward Donnelly, Jr.

In partial fulfillment of the requirements
for the degree of Doctor of Philosophy

California Institute of Technology

Pasadena, California

1992

(Defended February 17, 1992)

To my daughters Kelly and Jennifer, my best work as a graduate student, I dedicate this thesis, and I dedicate myself to spending more time with them now that this thesis is done.

Acknowledgements

I acknowledge that I do not have time to write a proper **Acknowledgements** section, but I wish to briefly thank several people. Charles Musgrave and Jason Perry were very helpful in providing and discussing the quantum chemistry calculations on hydrogen exchange reactions for Chapter 1. Several graduate student system managers (Bill: get a full timer) have been very kind in putting up with all of my problems and questions, especially Mark Brusich from the time when VAXen reigned supreme and Kian-Tat "KT" Lim who masterminded the current rule of UNIX.

It is hard to describe briefly the many facets of a relationship with your thesis advisor, especially one as interesting as Bill Goddard, so I just say: **THNKSWAG.**

My family life has changed much since my Cornell graduation. At that time, I was the youngest of my family. My grandparents have always been very kind and generous. My sister and brother-in-law have been good friends despite the distance between us. Most importantly, throughout many years of higher education, my parents have been most supportive, and being unable to describe everything they have done for me, I just say: **Thanks, Mom and Dad.**

Just two weeks before our arrival in Pasadena, I entered a new phase of my family life with my marriage with Karen. Her love, support, and companionship mean the world to me. The blessed arrivals of Kelly and Jennifer in the last two plus years have greatly enriched our lives. Now I am the "head" of my own family, or, more accurately, "co-chairs" with my wife. The words "**thank you**" cannot come close to expressing my gratitude for my wife, but, alas, they will have to do.

Thesis Abstract

"Modeling and Visualizing Surfaces" denotes several developments aimed at increasing our ability to model and understand molecular surfaces. The Generalized London Potential is a new method of modeling potential energy surfaces of reactions. The London Equation assumptions of zero overlap and no three-body interactions are discarded and a more general potential is derived via valence bond theory and careful substitution of two-body terms into three-body energy expressions. Three-body corrections for dispersion energy are also introduced. Using the lowest order forms of overlap and dispersion corrections, a much improved potential energy surface is found for H_3. Input is limited to H_2 potentials and information only about the H_3 saddle point region, the latter determining the two or three parameters used. The predicted surface is shown to be stable with respect to varied input. A straightforward method of extending the method to make use of additional input is discussed. The method is applied to hydrogen abstraction from terminal carbons. The development of a stable model of exchange reactions will greatly increase the complexity of systems which can be studied with the increasingly accurate force fields of molecular modeling techniques by providing the means of handling reactive dynamics at polymer and crystalline surfaces. This introduces the second major theme of modeling and visualizing molecular surfaces.

Common to most definitions of a molecule's surface and, in fact, many calculations involving local spherical symmetry is the use of spherical meshes. A method of systematically creating spherical meshes of various sizes is presented. Degrees of freedom built into the mesh design can be optimized for a variety of problems.

The meshes are used in calculating molecular surfaces and determining surface area. They are separately optimized for the integration of spherical harmonics and provide lower error for integration of higher angular momentum functions than previous quadratures. Finally, methods of visualizing molecular surfaces that allow real-time manipulation of complex molecules and yield a better understanding of surface properties are presented.

Contents

List of Figures

List of Tables

Chapter 1

A New Approach to Potential Energy Surfaces for Reactions

I. Introduction

The construction of potential energy surfaces (or force fields) for dynamics simulations or molecular modeling is much more difficult when bonds must be allowed to break and form than for situations involving nonreactive forces such as bending, stretching, van der Waals, etc. In exchange reactions, bonds are concurrently broken and formed, making it kinetically feasible to break strong bonds. However, this requires that the strongly attractive interaction between one bond pair must turn into a repulsive interaction between them as one of them forms a bond with the third body. Obviously, the sum of three bonding interactions resulting in a stable triangular geometry is inadequate. This changing of sign of the forces between various pairs of reactants can only be modeled by careful consideration of the bonding (attractive) and antibonding (repulsive) forces involved as moderated by the Pauli Principle. The fact that the barrier to reaction is often far less (2 to 10%) of the bond energy emphasizes the importance of the interplay of bonding and antibonding during the reaction.

For exchange reactions current one- or few-parameter models are often based on the Sato modification of the London-Eyring-Polanyi model [1] which dates to

1955 and cannot be derived from any set of approximations involving valence bond or other theory. When detailed information is available for a given potential energy surface (such as accurate *ab initio* calculations over much of the surface), researchers have resorted to many (20 or more) parameter fits in order to accurately reproduce the features of the surface. Terms specific to various regions of the surface, e.g., linear-symmetric or bent geometries, are used to decouple the complex least-squares fitting, but such terms often have little or no scientific motivation and must be reconsidered whenever more accurate information becomes available. We seek a middle ground involving a few parameters motivated by the physics of the reaction for the majority of systems where detailed information is not available. Our goal is to be able to incorporate limited information (perhaps just an estimate of the saddle point from experimental or *ab initio* data) into a physically motivated model of the potential energy surface of exchange reactions.

One such reaction is the exchange of an atom between two molecules (or even two sites on the same molecule). The energy barrier is generally far less than that required to break the first bond. We first consider the construction of potential energy surfaces for the transfer of an atom with a nearly complete valence shell from one system to another, i.e., those of the type

$$A \ + \ XB \longrightarrow AX \ + \ B,$$

where X is either hydrogen or a halide and both A and B are atoms or molecules one electron shy of a complete valence shell. Thus A and B could be hydrogen or a halide themselves or a more complex unit such as a Si or C atom with three other bonds. One example of the latter would be hydrogen abstraction on a silicon surface:

$$H \ + \ HSi\overleftrightarrow{\searrow\nearrow} \longrightarrow H_2 \ + \ Si\overleftrightarrow{\searrow\nearrow},$$

where the silicon atom is on the surface of a crystal.

In forming a model for the potential energy surface, it is best to start with a simpler (model) system for which more data is available for comparison. For example, the reactions

$$H + H_2 \longrightarrow H_2 + H,$$

$$F + H_2 \longrightarrow HF + H$$

are well studied and suitable for this task. H_3, especially, has been the testing ground for many *ab initio* and semi-empirical studies. Even so, recent measurements of both integral [2] and differential [3] cross sections for

$$H + H_2\,(v = 0, j = 0, para) \longrightarrow H_2\,(v' = 1, j' = 1, 3) + H,$$

show a much greater energy dependence of the effects of Feshbach resonances than do theoretical studies [4-6]. This has prompted additional study of the H_3 potential energy surface in higher energy regions (>20 kcal/mole), especially bent configurations [7]. $F + H_2$ has been given much attention since about 1970 due to application to chemical lasers. The amount known about this system rivals H_3, but the estimates of the classical barrier height still vary from 0.86 to 1.18 kcal/mole (experimental) [8] and 1.4 to 4.6 kcal/mole (theoretical) [9]. The validity of including Davidson and correlation corrections and the importance of zero-point energy and tunneling effects are still undecided. Thus even for the two most closely scrutinized exchange reactions, there are still a few unknowns. Information on other exchange systems is more limited, so a general model of potential energy surfaces should not require a large amount of information.

For other hydrogen and halide exchange reactions, we can look for input from several sources. Experimental reaction rates can yield estimates of the barrier height, perhaps to ± 1 kcal/mole. Quantum chemistry calculations can be done at a few points to gain additional energy and saddle point geometry information. The relative

energies of the reactant and product channels and the AX, XB, and AB two-body potentials are usually attainable. An ideal model for the potential energy would require only the information listed in this paragraph yet correctly describe the initial and final states and approximate the important features of the transition state. This chapter seeks to answer the question: "How do we formulate a potential energy surface using this information?" The next section looks at several historically important models for exchange reactions, Section III derives the London Equation from the new perspective of spin coupling, and the following sections correct the London Equation to account for the approximations it makes and then apply the new methods to several systems.

II. Important Models of Exchange Reactions

This section lists some of the important representations of the potential energy of H_3 from a historical perspective. Our goal is not to review the modeling of the H_3 potential energy surface [10-12] but to understand the assumptions and approximations underlying several of the techniques and thus their relation to the current work.

By setting overlap and three-body terms to zero, the valence bond potential energy surface for three atoms with only one valence s electron each is given by the London Equation,

$$W = A + B + C - \sqrt{\alpha^2 + \beta^2 + \gamma^2 - \alpha\beta - \beta\gamma - \gamma\alpha}, \tag{1.1}$$

where A and α are functions of one interatomic distance, only, as B and β, C and γ are functions of the other two interatomic distances. A, B, and C are the Coulombic or classical terms, and α, β, and γ are the resonant or exchange terms of the three pairs of atoms. The original derivation of the London Equation is based on solving

the secular equation for this system [13-18]. While the London Equation is quite useful by reducing the three-atom problem to the Coulombic and resonant terms of two-atom systems, functions must be assigned to these terms before one can proceed.

Before overlap has been set to zero, Heitler and London [19] determined the energy of H_2 to be[1]

$$W = \frac{A \pm \alpha}{1 \pm S^2}. \tag{1.2}$$

With α negative for all but very short distances, the upper sign corresponds to a bonding state and the lower sign to an antibonding state. Sugiura [20] was able to obtain equations for these Coulomb and resonant terms for H_2 based on strict "1s" orbitals. Thus he obtained potential energy curves for the singlet ground state and the triplet excited state. The Coulomb and resonant terms and the resulting binding energy from Sugiura's equations are shown in Figure 1.1. His equilibrium bond distance is 1.51 bohr and bond energy is 74 kcal/mole compared to the known values of 1.40 bohr and 109.5 kcal/mole (accounting for zero-point energy), respectively. The results of this simple calculation actually can be improved by using an optimally scaled "1s" orbital instead of that of a hydrogen atom. In Figure 1.2, we see that the Coulomb energy is a small portion (10 to 15%) of the total binding energy for distances above the H_2 bond distance. While the calculated energies of Sugiura for H_2 are not accurate enough to estimate barrier energies for H_3, the fact that the H_2 Coulomb energy is on the order of 0.12 times the binding energy is useful.

Eyring and Polanyi [21,22] created the original semi-empirical method by using a Morse Potential [23] for the binding energy of H_2. Initially, they assumed that since the Coulomb portion of the binding energy is small compared to the resonant portion, the Morse Potential can be used for α in a first approximation. By using

[1]We define S as the overlap of two orbitals instead of the overlap of two wavefunctions as done in the early papers dealing with this subject.

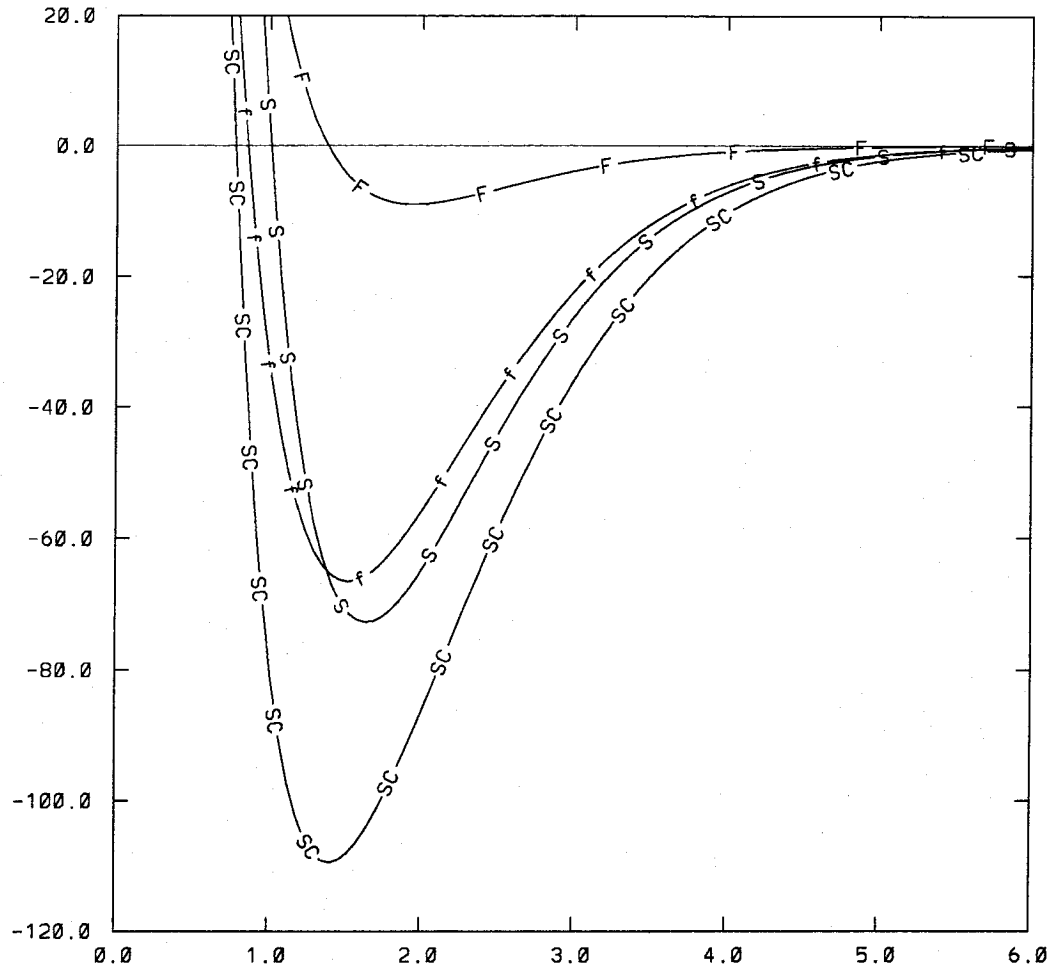


Figure 1.1. Sugiura's Coulomb and Resonant Energies. The binding (S), Coulomb (F), and resonant (f) energies of Sugiura are compared with the singlet energy as currently known (SC). The Coulomb and resonant (exchange) energies reflect Sugiura's functional form (and notation) of including the correction for overlap, i.e., $F(r) = \frac{A}{1+S^2}$ and $f(r) = \frac{\alpha}{1+S^2}$, so that $E^s = F + f$.

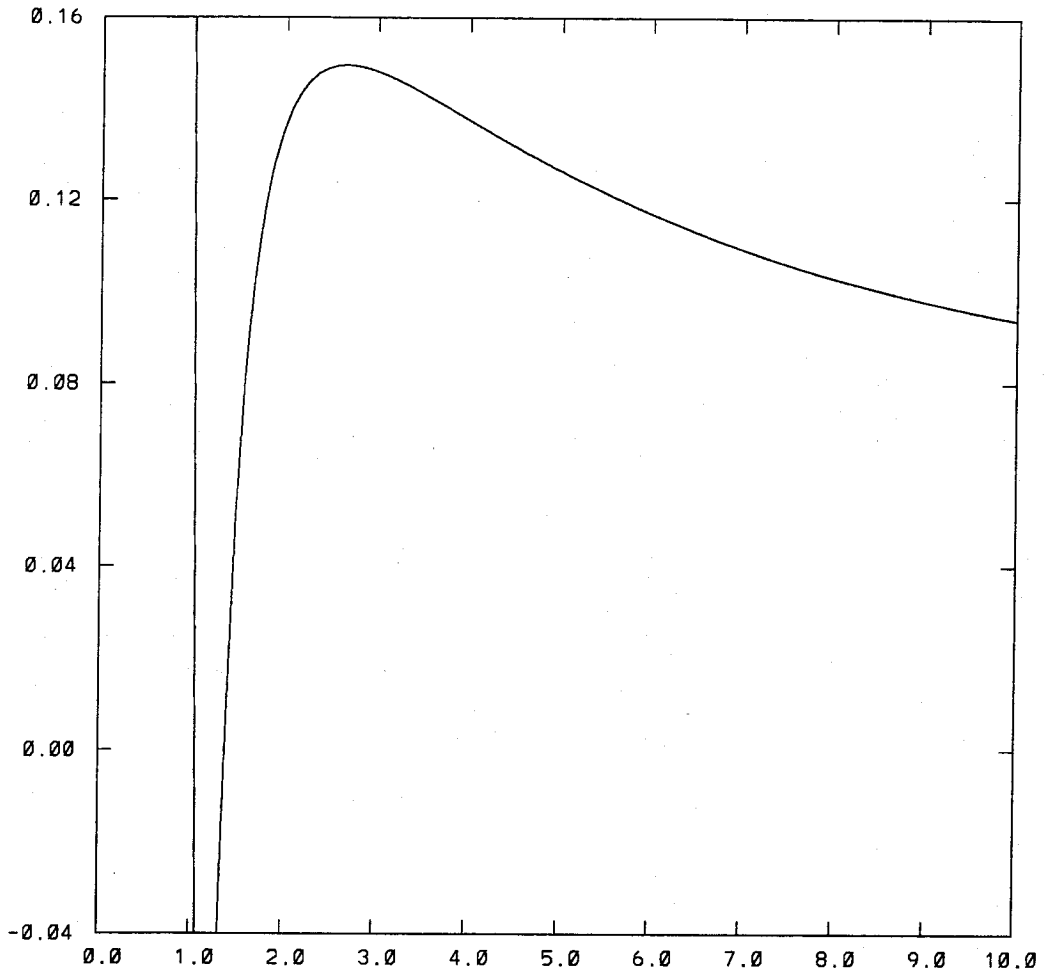


Figure 1.2. Coulomb to Binding Energy Ratio. Sugiura's Coulomb energy $F(r)$ divided by his binding energy is plotted.

the Coulomb potential of Sugiura for A, a potential energy surface can be plotted for H_3 using the London Equation. This leads to a barrier energy of about 20 kcal/mole and, more importantly, a shallow (< 1.5 kcal/mole) well at the top of the reaction path. Following up on an earlier suggestion [24] of his that the Coulomb energy be taken as a fixed fraction (10%) of the binding energy of the Morse function, Eyring [25] determined another potential energy surface that has a lower barrier energy (13.0 kcal/mole) but a similar 1.6 kcal/mole well so that the minimum energy for a linear symmetric geometry is 11.4 kcal/mole. This method of assuming a constant Coulomb to binding energy ratio is known as the LEP (London, Eyring, Polanyi) method. The well is now known to be merely an artifact of the LEP method, probably caused by the fast changing Coulomb/binding energy ratio near the saddle point distance.

Sato [1] used estimates of both the ground state singlet and the lowest energy triplet curves to determine the Coulomb and resonant energies of H_2. He used a Morse Potential for the singlet, proposed an Antimorse Potential for the triplet, and used a form of Equation (1.2) where the distance-dependent S^2 is replaced by an adjustable constant k. He also added a multiplicative factor of $\frac{1}{1+k}$ to the London Equation. This last modification (even with variable overlap) should be treated as merely a useful functional form since it cannot be obtained from any approximate treatment of valence bond theory. Using constant overlap to scale both the H_2 and H_3 energies does reproduce the correct dissociation limit of one pair in a pure singlet state as the third atom is removed. Weston implemented Sato's method by adjusting k to fit the experimental activation energy [26]. No well at the top of the reaction path is found. This method of using an adjustable Sato parameter k is known as the LEPS (LEP plus Sato) method.

Many applications of the LEP and LEPS methods exist (see those mentioned in

References 10 and 12), but we note only one additional implementation of the London Equation and its relatives. Cashion and Hershbach [27] started with the then best available estimates of the actual singlet and triplet potential energy curves. Using the Heitler-London energy of H_2 with overlap set to zero ($E = A \pm \alpha$) and the original London Equation, they obtained a potential energy surface for H_3 with a barrier of 8.9 kcal/mole. This method with improved H_2 curves is discussed in Section IV.

The accurate configuration interaction (CI) calculations of Liu [28] on linear H_3 and Siegbahn and Liu [29] on bent H_3 give us a set of points covering a large range of geometries for comparison and well define the properties at the barrier to reaction. The calculated saddle point energy was 9.80 kcal/mole with an estimate of 9.68 $\pm$ 0.12 kcal/mole adjusting for correlation error. Truhlar and Horowitz presented an accurate fit to these calculations [30]. They start with the London Equation using a singlet curve that is a spline fit to *ab initio* energies and a parameterized "triplet" curve that is reduced from the actual triplet, especially at shorter R. The four parameters of that curve are chosen to minimize the least squares deviation of the linear symmetric points under the constraints of exact E_{sp} and R_{sp}. One five parameter term is added to fit the additional linear points in a least-squares sense, and five bending terms comprising fourteen parameters are added for nonlinear geometries. This fitted surface is called the LSTH (Liu, Siegbahn, Truhlar, Horowitz) surface. While such a functional form is too complicated to be useful here, the LSTH was considered for many years to be the most accurate representation of any reaction surface and is quite useful (along with the CI points themselves) for comparison.

Liu published a revised saddle point energy [31] and (with Blomberg) additional bent geometries of higher energy [32]. The calculated saddle point energy is 9.65 kcal/mole with an estimate of 9.59 $\pm$ 0.06 kcal/mole. This agrees well with quantum Monte Carlo calculations [33,34]. A. J. C. Varandas et al. [35] adjust the Liu et al.

[28,29,32] points such that the barrier is 9.65 kcal/mole and present a new many-parameter fit (called DMBE for double many-body expansion) that has a higher root-mean-square error than LSTH, but may more accurately represent higher energy regions.

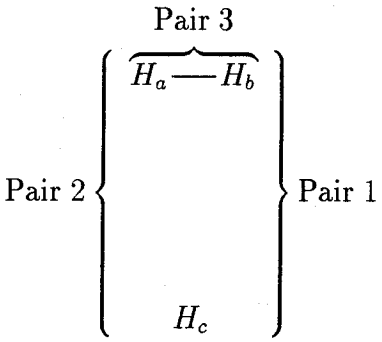
To summarize, the LSTH and DMBE surfaces are both accurate representations of the H_3 potential energy surface and differ by at most only a few tenths of a kcal/mole for all but high energy regions. The error bounds on any other exchange reaction potential surface will be greater than this difference. While the forms of these equations are unique to H_3 and such a large number of parameters cannot be used for less accurately known surfaces, they provide a unique opportunity for detailed comparison of experimental and theoretical dynamical measurements. The arrival of accurate *ab initio* energies over a large range of configuration space suspended the search for few-parameter fits (such as the LEPS surface) to the H_3 potential energy surface. We renew that search in the hope not of finding a more accurate H_3 surface but of finding a physically based model capable of handling various exchange reactions.

III. Derivation of London Potential in Terms of Spin Coupling

The original derivation of the London Equation [13-18] is based on solving the secular equation for three atoms with one s electron each. We find it worthwhile to consider two special cases explicitly and to derive the general potential by considering the optimal spin coupling of these special cases. Before proceeding, the approximations and assumptions of the London Equation are stated. First, it is assumed that a single valence bond orbital on each hydrogen atom can represent the total wavefunction.

Second, both orbital overlap and explicit three-body terms are neglected. Third, it is assumed that the remaining energy terms can be replaced by their counterparts in H_2. We make use of the results of two appendices. Appendix A details the derivation of energy expressions that are exact under the valence bond description, i.e., no terms are set to zero; Appendix B gives approximate energy expressions under several sets of assumptions where various terms are assumed to be zero. Each set basically varies in its treatment of overlap (and quantities which exponentially decay like overlap). The final set in which all overlap terms are set to zero is that of the London Equation. The method of obtaining the usual London Potential from the spin-coupling representation used here and the relation of the London Potential and its spin coupling are also given in Appendix B.

The first limiting (special) case is that of a bonded pair with the third atom far removed:



Since we know that the spins of H_a and H_b (pair 3) are singlet paired, we can assign either up (α) or down (β) spin to H_c. The appropriate spin function is thus

$$\chi_1 = \frac{1}{\sqrt{2}}(\alpha\beta - \beta\alpha)\alpha. \tag{1.3}$$

Under the approximations of the London Potential, Equation (1.138) from Appendix B tells us that the energy is

$$\mathcal{H}_{11} = E_3^S + (\tfrac{3}{4}E_2^T + \tfrac{1}{4}E_2^S) + (\tfrac{3}{4}E_1^T + \tfrac{1}{4}E_1^S). \tag{1.4}$$

Here, E_i^T and E_i^S are the triplet (antibonding) and singlet (bonding) energies, respectively, of pair i. A second limiting case is that of the transition state ($R_1 = R_2 < R_3$ for bent, $R_1 = R_2 = \frac{1}{2}R_3$ for linear, where R_i is the distance between pair i):

$$H_a \text{————} H_c \text{————} H_b.$$

Since we know that the spins of atoms H_a and H_b are triplet paired, this state has spin function

$$\chi_2 = \frac{1}{\sqrt{6}}(\alpha\beta\alpha + \beta\alpha\alpha - 2\alpha\alpha\beta). \tag{1.5}$$

The energy expression per Equation (1.139) is

$$\mathcal{H}_{22} = E_3^T + (\tfrac{1}{4}E_2^T + \tfrac{3}{4}E_2^S) + (\tfrac{1}{4}E_1^T + \tfrac{3}{4}E_1^S). \tag{1.6}$$

In an equilateral triangle arrangement of the three atoms ($R_1 = R_2 = R_3$), both of these limits reduce to the same energy of

$$E = \tfrac{3}{2}(E_i^T + E_i^S). \tag{1.7}$$

For nonequivalent atoms, the more general condition for this reduction is that $\frac{1}{2}(E_i^T - E_i^S)$ be the same for every pair. This degeneracy will be important later in limiting the form of corrections to the London Equation. Note that the overlap between the wavefunctions associated with these two spin functions is zero under the approximations of the London Equation. Per Equation (1.140) the energy cross term between these two states is

$$\mathcal{H}_{12} = \tfrac{\sqrt{3}}{4}(E_2^T - E_2^S - E_1^T + E_1^S). \tag{1.8}$$

Note that the overlap between the wavefunctions corresponding to these two spin functions is zero under the London approximations. The general energy expression is found by taking a normalized combination of these two spin functions. The energy

corresponding to the general spin function

$$\chi = \chi_1 \sin\tfrac{1}{2}\gamma + \chi_2 \cos\tfrac{1}{2}\gamma \tag{1.9}$$

is

$$\begin{aligned} E_L &= \sin^2\tfrac{1}{2}\gamma\mathcal{H}_{11} + \cos^2\tfrac{1}{2}\gamma\mathcal{H}_{22} + 2\sin\tfrac{1}{2}\gamma\cos\tfrac{1}{2}\gamma\mathcal{H}_{12} \\ &= \tfrac{1}{2}(1-\cos\gamma)\mathcal{H}_{11} + \tfrac{1}{2}(1+\cos\gamma)\mathcal{H}_{22} + \sin\gamma\mathcal{H}_{12}. \end{aligned} \tag{1.10}$$

The London Equation as found using the secular equation is independent of the spin coupling γ. To obtain it, E_L is minimized with respect to γ yielding

$$\tan\gamma = \frac{\mathcal{H}_{12}}{\frac{1}{2}(\mathcal{H}_{22}-\mathcal{H}_{11})} = \frac{\frac{\sqrt{3}}{2}(E_2^x - E_1^x)}{(E_3^x - \frac{1}{2}E_2^x - \frac{1}{2}E_1^x)}, \tag{1.11}$$

where we have defined $E_i^x \equiv \frac{1}{2}(E_i^T - E_i^S)$ and $E_i^{cl} \equiv \frac{1}{2}(E_i^T + E_i^S)$. Note that these definitions change when overlap is no longer neglected. Substituting this solution for γ in Equation (1.10) (see Appendix B for details) and using the above expressions for $\mathcal{H}_{ij}$ gives us

$$E_L = E_1^{cl} + E_2^{cl} + E_3^{cl} - (E_1^{x2} + E_2^{x2} + E_3^{x2} - E_1^x E_2^x - E_2^x E_3^x - E_1^x E_3^x)^{\frac{1}{2}}, \tag{1.12}$$

the standard form of the London Potential.

Previous manipulations of the London Equation (such as the LEPS potential) start with the London Potential without reference to the mixing of the underlying limiting states. It is our belief that information is lost in only using the standard form of the London Equation (Equation (1.12) as derived per the secular equation) instead of the spin-coupling form (Equation (1.10) as derived here, i.e., before substituting for γ). γ tells us the mixing of the states χ_1 and χ_2, and thereby gains its name of spin coupling. See Figure 1.5 in the next section for a plot of γ for linear H_3 to see how it uniformly changes along the reaction path. Table 1.6 in Appendix B further

illustrates the usefulness of γ as a reaction coordinate as it specifies pure singlet or pure triplet pairing between a certain pair of atoms at each multiple of $\frac{\pi}{3}$ over a 2π range.

While the London Equation is developed above by concentrating on the coupling limits of pair 3 (atoms a and b), it is worthwhile to consider further manipulation of the London Equation (before the substitution for γ) in order to emphasize the symmetry of the London Equation with respect to exchange of atoms. Begin by rewriting Equation (1.10) after substituting for $\mathcal{H}_{ij}$:

$$\begin{aligned} E_L = \tfrac{1}{2}[\quad & (1+\cos\gamma)E_3^T + (1-\cos\gamma)E_3^S \\ & + (1-\tfrac{1}{2}\cos\gamma + \tfrac{\sqrt{3}}{2}\sin\gamma)E_2^T + (1+\tfrac{1}{2}\cos\gamma - \tfrac{\sqrt{3}}{2}\sin\gamma)E_2^S \\ & + (1-\tfrac{1}{2}\cos\gamma - \tfrac{\sqrt{3}}{2}\sin\gamma)E_1^T + (1+\tfrac{1}{2}\cos\gamma + \tfrac{\sqrt{3}}{2}\sin\gamma)E_1^S \quad]. \end{aligned} \tag{1.13}$$

Next note that

$$\cos(\gamma \pm \tfrac{2\pi}{3}) = -\tfrac{1}{2}\cos\gamma \mp \tfrac{\sqrt{3}}{2}\sin\gamma. \tag{1.14}$$

Thus if we define $\gamma_3 \equiv \gamma$, $\gamma_2 \equiv \gamma - \frac{2\pi}{3}$, and $\gamma_1 \equiv \gamma + \frac{2\pi}{3}$, we can succinctly write

$$E_L = \sum_{i=1}^{3}[f^T(\gamma_i)E_i^T + f^S(\gamma_i)E_i^S], \tag{1.15}$$

by defining

$$f^T(\gamma) \equiv \tfrac{1}{2}(1+\cos\gamma), \tag{1.16}$$

$$f^S(\gamma) \equiv \tfrac{1}{2}(1-\cos\gamma). \tag{1.17}$$

This symmetric form of the London Equation emphasizes the importance of the change of singlet and triplet character of each pair of atoms as the spin coupling γ varies over its domain of 2π. One final version which emphasizes the distribution of exchange energy among the pairs of atoms is

$$E_L = \sum_{i=1}^{3}[E_i^{cl} + \cos\gamma_i E_i^x]. \tag{1.18}$$

With the sum of the three $\cos\gamma_i$'s being zero, this illustrates how the London Potential is a distribution of bonding ($\cos\gamma_i < 0$) and antibonding ($\cos\gamma_i > 0$) among the three pairs of atoms so that the net bonding is zero.

IV. London Potential and H_3

While the H_2 calculations of Kolos and Wolniewicz [36] greatly improved the accuracy to which the ground singlet and first excited triplet potentials are known, the current authors know of no reporting of their use in a strict London Potential. This section investigates the accuracies and limitations of doing thus that. Spline fits to their CI points are plotted in Figure 1.3. This gives us the best possible starting point for our two-body potentials in order to evaluate our methods, but we cannot expect to have starting potentials of this accuracy for other systems. While it is necessary to use the actual singlet curve to ensure that the correct $H + H_2$ limit is obtained, the triplet curve could be adjusted to fit certain features of the H_3 surface. We prefer to use the best available estimates for both the actual singlet and triplet curves to avoid the extra parameters of a fitted triplet, since for systems other than H_3 we can more easily find two-body triplet curves than extra three-body data to which to fit the triplet. Instead, we look for physically motivated methods to account for the approximations of the London Equation in the next section. Figure 1.4 shows a contour diagram of the London Potential using the spline fits to the H_2 CI points of Kolos and Wolniewicz [36]. The variation of γ over the same linear space is shown in Figure 1.5. The facility of γ as a reaction coordinate is seen in its smooth change from $-\frac{\pi}{3}$ to $\frac{\pi}{3}$ along the reaction path. We use the fit of Truhlar and Horowitz [30] to the H_3 CI points of Liu et al. [28,29] (the LSTH surface) for comparison to the London surface. The main reason for this is to retain the concurrent ability of direct

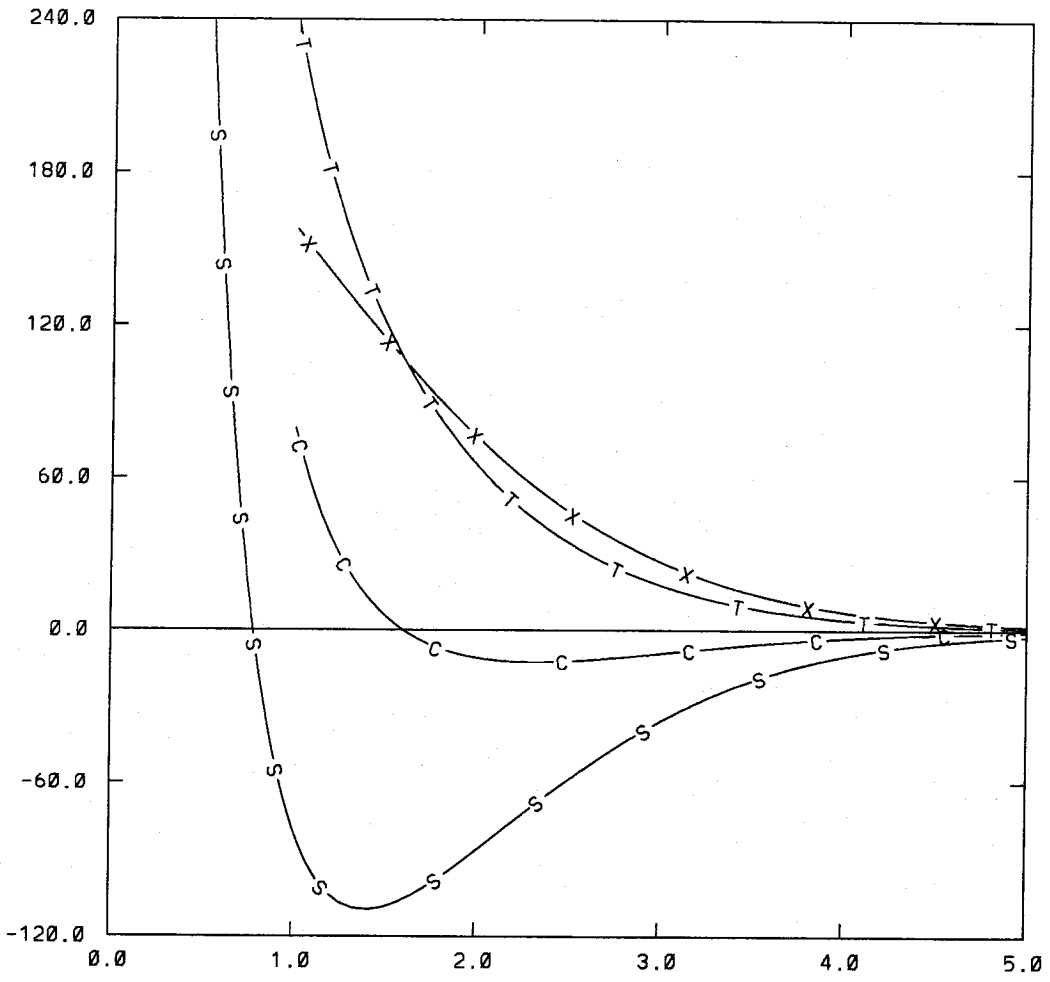


Figure 1.3. H_2 Spline Fits to CI Points. T is spline of E^T, the triplet; S is spline of E^S, the singlet. C for E^{cl} and X for E^x are half of the sum and difference, respectively, of the triplet and singlet splines. Energy is in kcal/mole, and distance is in atomic units.

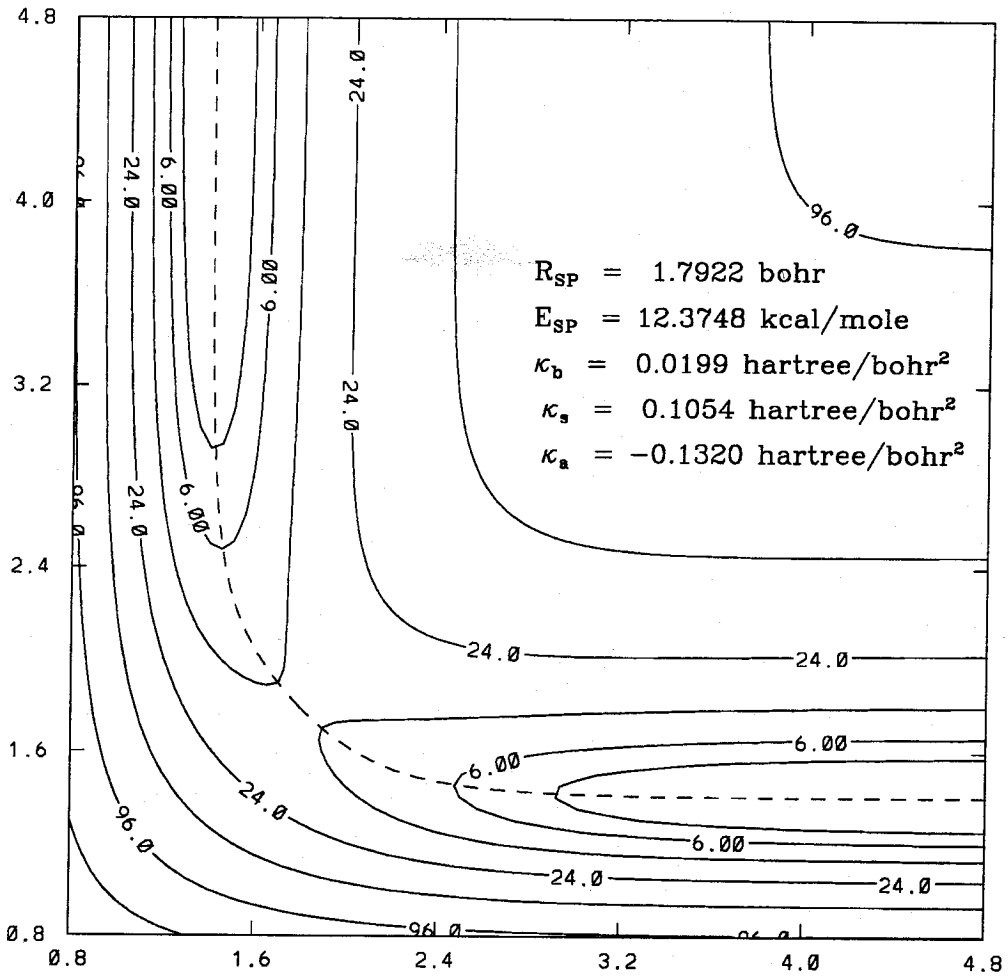


Figure 1.4. London Potential Contour Diagram. Splines of H_2 CI energies are used to provide two-body terms. The contour values shown are the powers of two times 3 kcal/mole. Axes are distances in atomic units.

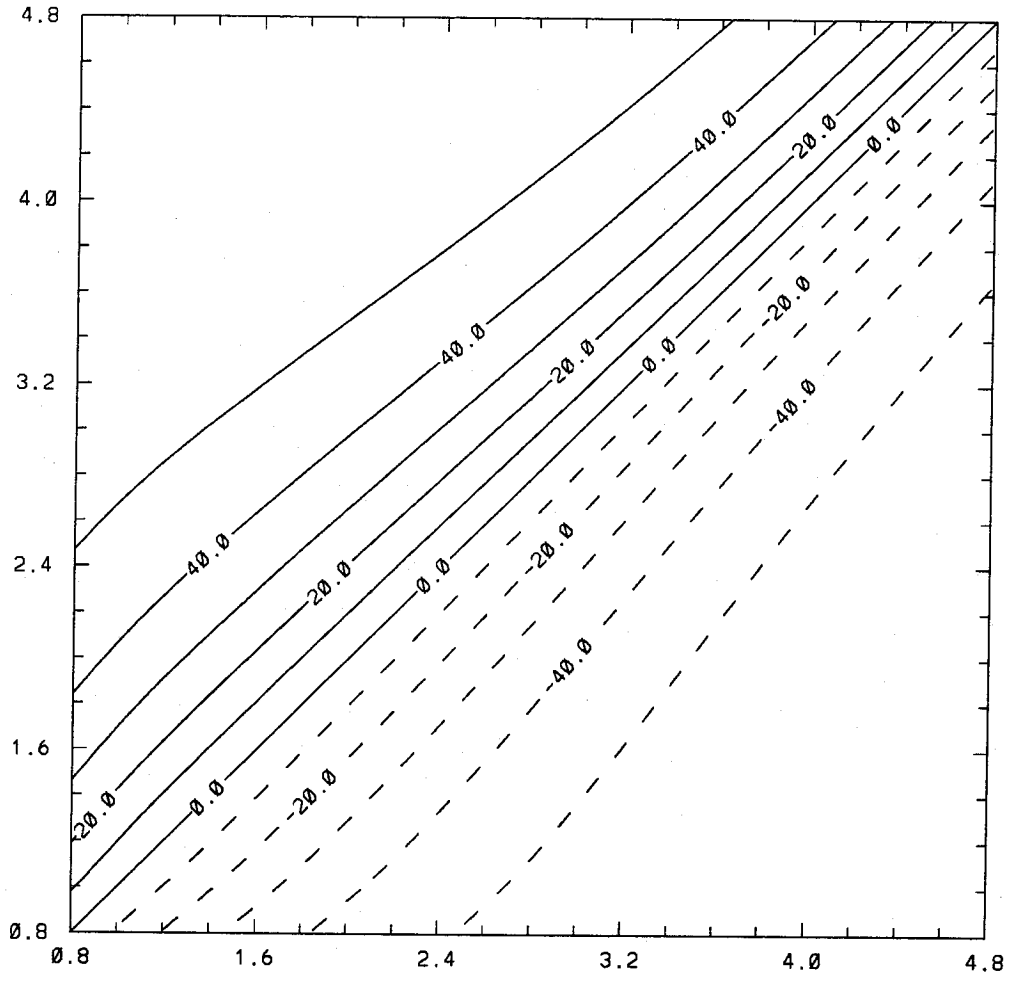


Figure 1.5. Contour Plot of γ for Linear H_3. Contours are spaced every 10 degrees with negative contours endashed.

comparison with the unadjusted Liu et al. points, and the at most few tenths of a kcal/mole difference between the LSTH and DMBE surfaces is not significant for our considerations. Contrast the LSTH contour plot of Figure 1.6 with the London plot of Figure 1.4.

We now examine what the London Potential is able to do and what it cannot do for H_3. Most importantly, it provides the correct dissociation limits ($H_2 + H$ and $H + H + H$) and has all of the correct features such as a barrier energy of the correct order of magnitude and a minimum energy reaction path of a reasonable shape. However, the barrier energy (E_{sp}) is not to "chemical accuracy" and the position of the saddle point (R_{sp}) is off by two percent. Two of the three curvatures at the saddle point are reasonable, but the third is off by over a factor of two. The curvatures, κ_b, κ_s, and κ_a, are the second derivatives of the energy with respect to the variables:

$$x_b \equiv \left(\frac{R_{1\mathrm{sp}}R_{2\mathrm{sp}}}{R_{3sp}}\right)(\theta_{sp} - \theta), \tag{1.19}$$

$$x_s \equiv \frac{\sqrt{3}}{2}[(R_1 - R_{1sp}) + (R_2 - R_{2sp})], \tag{1.20}$$

$$x_a \equiv \frac{1}{2}[(R_1 - R_{1sp}) - (R_2 - R_{2sp})], \tag{1.21}$$

where an sp subscript indicates the value at the saddle point and θ is the angle opposite pair 3. While the London Potential does a remarkable job for linear configurations given that it requires only H_2 curves for input, Figure 1.7 shows that there are two major errors along the reaction path, the most important region to fit. One is that the barrier is too high as seen near the origin; two is that the barrier is too "thin" (κ_i is too negative) as seen for longer distances. To summarize, our task now is to generalize the London Potential in such a manner as to require only one or two parameters but to fulfill the following goals:

- To be extendible to other exchange reactions.

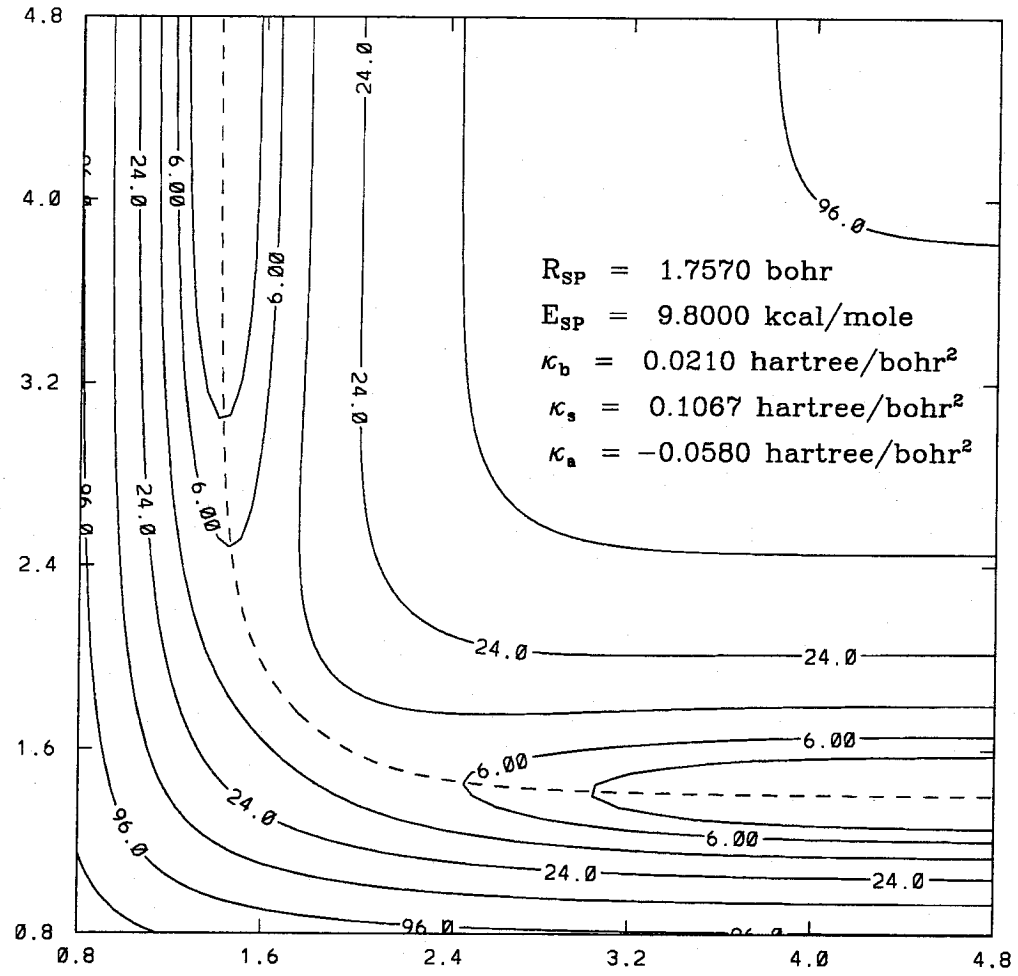


Figure 1.6. Truhlar and Horowitz (LSTH) Contour Diagram. Reported saddle point properties are actually those of CI calculations of Liu et al.

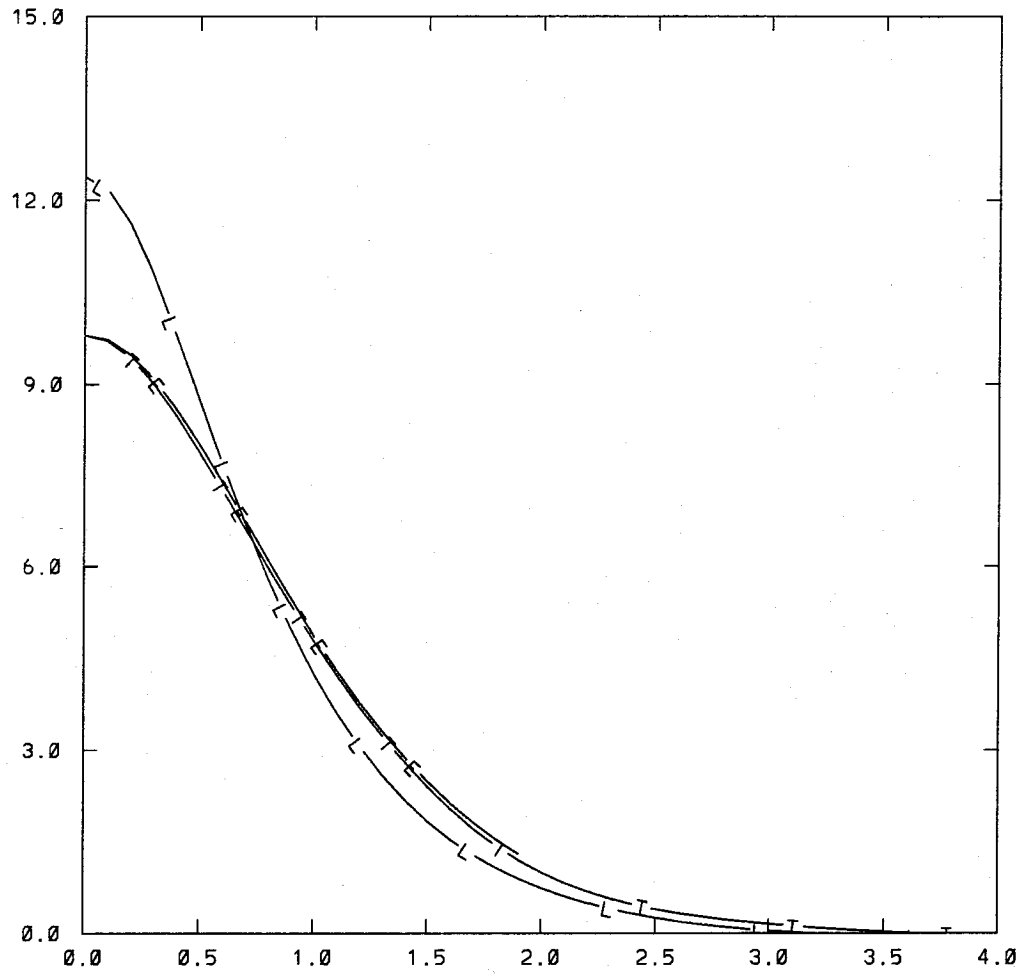


Figure 1.7. Reaction Path Energy: London vs. LSTH. Energy in kcal/mole is given along the reaction path with the saddle point as the distance origin. Distance (in atomic units) is integrated along the minimum energy path.

- To retain the correct dissociation limits of the London Potential.
- To correct for the barrier being too high (fit E_{sp} better).
- To correct for the barrier being too thin (fit κ_a better).
- To require limited input, however be extendible as more input is available.
- To do all of the above based on sound physical reasoning.

Such a method is proposed in the next section.

V. Generalized London Potential

The approximations used in deriving the London Equation include the neglect of overlap, the neglect of terms involving all three atoms, and the use of H_2 values for similar terms in the H_3 energy expression. Only the last of these approximations is not included in the valence bond formulation derived in Appendix A. Valence bond methodology does not include dynamic electron correlation and charge transfer. Thus the use of H_2 expressions in H_3 implicitly assumes that the H_3 dispersion energy is the pairwise sum of H_2 dispersion energies. In this section we look at correcting these approximations by generalizing the form of the spin-coupled London Equation. Since the corrections will always involve at least two pairs of atoms and thus are always three-body in nature, the London Equation can be considered exact through one pair interactions. We distinguish between corrections involving two pairs and those involving three pairs in the belief that this is the proper hierarchy of corrections. Note that Appendices A and B give new equations including the corrections but not the form of the corrections themselves. To help to understand the magnitude and nature of the required corrections, Figures 1.8 and 1.9

plot the difference between the London and LSTH potentials for linear and bent H_3, respectively. The overestimate by the London Potential in the lower left of each plot is due to the neglect of both overlap and three-body terms plus the lack of correlation energy. This effect increases as θ goes from 180 to 60 degrees as all three bond lengths become shorter. The disagreement is further enhanced near equilateral triangle geometries as three-body interactions increase. The underestimate by the London Potential at longer distances is due to the use of H_2 terms which include a greater amount of dispersion energy than possible for H_3. When one pair of atoms is primarily singlet coupled with the third atom somewhat removed, the full amount of interaction of orbitals is expected for that pair. However, the interaction of this pair of orbitals with each other reduces the amount with which they can interact with the third atom. This section provides methods for correcting for the terms explicitly dropped in the valence bond derivation of the London Potential and for accounting for the interaction of pairwise dispersion energies.

V.A. Overlap Correction

Presumably the most important approximation in the derivation of the London Potential is the neglect of overlap. The Heitler-London theory for H_2 (which evolved into valence bond theory) traditionally gives the energies of the singlet and triplet states as

$$E_{ab}^S = \frac{E_{ab}^{cl} - E_{ab}^x}{1 + S_{ab}^2}, \tag{1.22}$$

$$E_{ab}^T = \frac{E_{ab}^{cl} + E_{ab}^x}{1 - S_{ab}^2}, \tag{1.23}$$

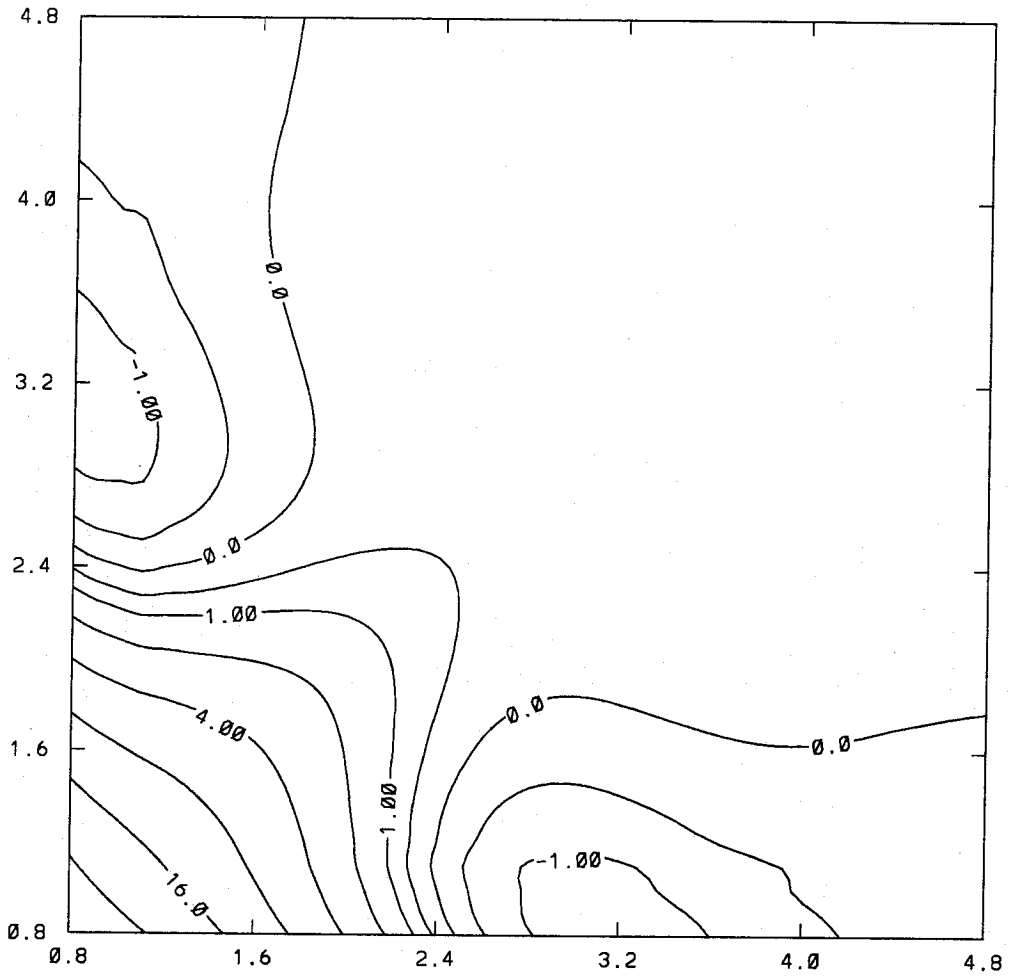


Figure 1.8. Linear London Minus LSTH Contour Plot. Contours (in kcal/mole) are zero and $\pm 2^l$, $l = -1, 0, 1, 2, \ldots$. Axes are distances in atomic units.

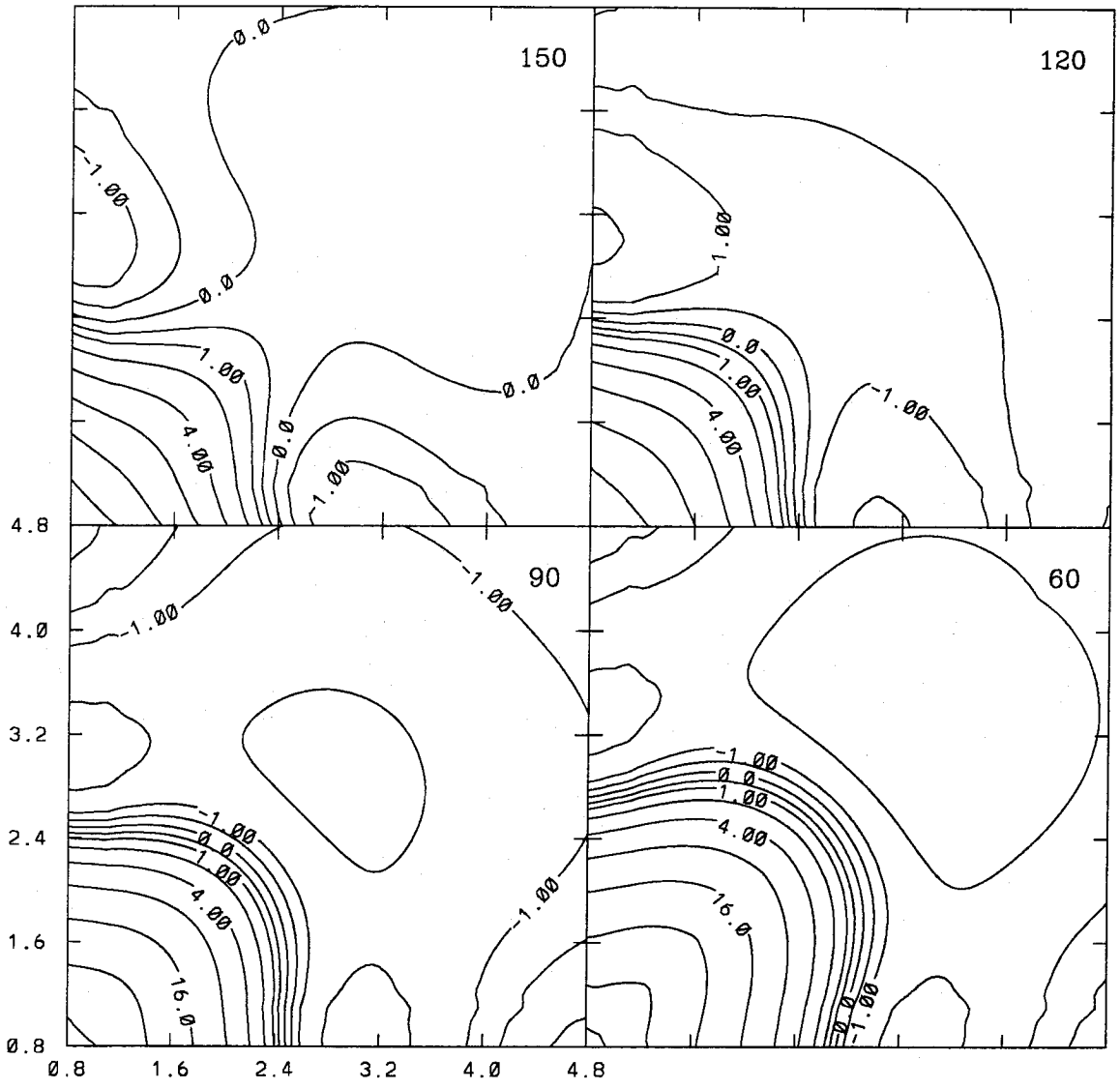


Figure 1.9. Bent London Minus LSTH Contour Plot. Contours (in kcal/mole) are zero and $\pm 2^l$, $l = -1, 0, 1, 2, \ldots$. Axes are distances in atomic units per lower left plot. Angles are marked in degrees in upper right corner.

where E_{ab}^{cl} and $-E_{ab}^{x}$ replace the A and α of Equation (1.2). Note that we use the expressions

$$E_{ab}^{S} = E_{ab}^{cl} - \frac{E_{ab}^{x}}{1+S_{ab}^{2}}, \tag{1.24}$$

$$E_{ab}^{T} = E_{ab}^{cl} + \frac{E_{ab}^{x}}{1-S_{ab}^{2}} \tag{1.25}$$

in order to make the substitution into the H_3 equations easier, but this necessitates only a slight change in definition of E_{ab}^{x}. The advantage of these definitions is that E_{ab}^{cl} follows the nuclear repulsion $\frac{1}{r_{ab}}$ instead of the Coulomb integral J_{ab}. This matters little for H_2 by itself but is important when substitution into H_3 energy expressions is made. Starting with the secular determinant, it is difficult to derive a similar energy expression including overlap for H_3. Starting with the spin-coupling formulation used here, we have done just that in the appendices. There we progress with increasing orders of approximation leading up to the London Equation, but here we reverse that order and start with the London Equation. We generically refer to the methods discussed here as overlap correction of London (OCL) potentials which form part of the Generalized London Potential (GLP) methodology.

The H_3 potential corrected for overlap through second order (2-OCL) is given by Equation (1.137) in Appendix B, which can be rewritten as

$$E = \sum_{i=1}^{3}[f_i^T E_i^T + f_i^S E_i^S - \tfrac{1}{4}(S_j^2 + S_k^2)(E_i^T - E_i^S)], \quad i \neq j \neq k, \tag{1.26}$$

by abbreviating $f_i^T = f^T(\gamma_i)$ and $f_i^S = f^S(\gamma_i)$. This simplest first correction illustrates that (within this approximation) the overlap between one pair of atoms has no effect on the potential due to that pair of atoms, i.e., a pair's overlap only effects the energy terms of other pairs. This results in the correct dissociation limits and the overall good fit of the London Potential except in that region of configuration space where all three atoms are nearby. Another important observation is that this

correction does not involve spin coupling. Thus it can be considered as a correction to the classical energy:

$$E = \sum_{i=1}^{3}[E_i^{cl} + (S_i^2 - \tfrac{1}{2}S_j^2 - \tfrac{1}{2}S_k^2)E_i^x + \cos\gamma_i E_i^x]. \tag{1.27}$$

Remember that when including S^2 terms in the H_3 energy expression that the classical and exchange energies are given by Equations (1.97) and (1.98) leading to the cancellation of the $S_i^2 E_i^x$ term with part of E_i^{cl}. Thus the London spin coupling has not been changed by adding overlap through second order. At this point we require an estimate of overlap. Inspecting the scaled "1s" H_2 integrals of Appendix A, one finds that both S_{ab}^2 and E_{ab}^x fall off at the same exponential rate. In fact, Figure 1.10 shows that one obtains a reasonable estimate of overlap from $S_{ab}^2 = \delta^o E_{ab}^x$ by taking δ^o as an adjustable parameter. It is varied to adjust the saddle point energy, i.e., the barrier to reaction, to a known value. It is convenient to adjust δ^o to fit four possible energy barriers:

9.59 kcal/mole The best known estimate [31].

9.65 kcal/mole The best calculated value [31].

9.80 kcal/mole The best calculated value for which a larger group of linear points at the same level exist [28].

9.9 kcal/mole The best calculated value for which a larger group of bent points at the same level exist [29].

The results are given in Table 1.1, showing that just this one parameter does a remarkable job in bringing the London R_{sp} and E_{sp} close to their true values. κ_a is also improved but remains too large in magnitude. Also important is slow and smooth variation of the rest of the surface as δ^o changes to fit various barrier lengths.

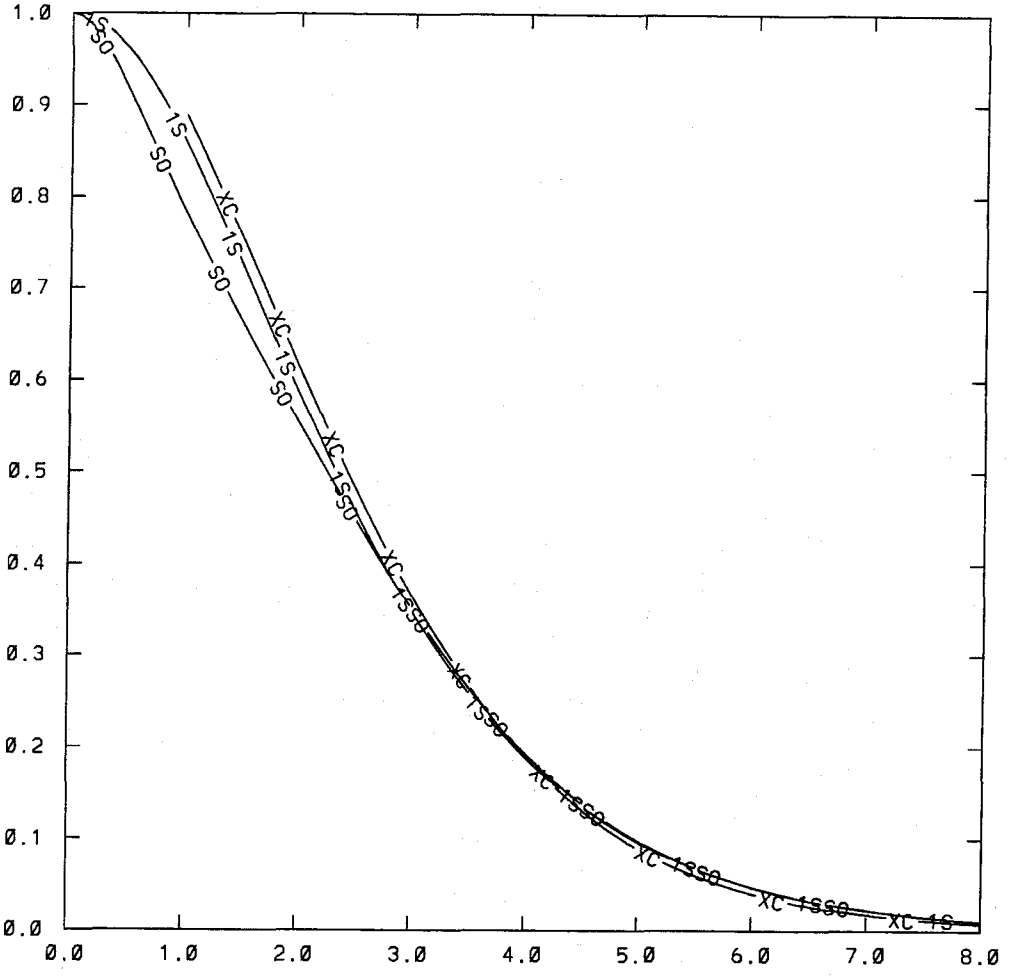


Figure 1.10. Estimate of Overlap for H_2. 1S is the standard "1s" overlap, SO uses an optimized (scaled) "1s" orbital, and XC uses $S^2 = \delta^o \frac{1}{2}(E_{CI}^T - E_{CI}^S)$ with $\delta^o = 0.0025$ in $(\text{kcal/mole})^{-1}$.

Method	δ^o	R_{sp}	E_{sp}	κ_b	κ_s	κ_a
Reference [31]		1.759	9.65		.1092	
Reference [28]		1.757	9.80		.1067	-.058
Reference [29]		1.757	9.90	.021		
London	0	1.7922	12.375	.01991	.1054	-.1320
2-OCL	2.5686 -4	1.7638	9.590	.02027	.1101	-.1079
2-OCL	2.5148 -4	1.7644	9.650	.02026	.1100	-.1084
2-OCL	2.3801 -4	1.7659	9.800	.02025	.1098	-.1097
2-OCL	2.2900 -4	1.7669	9.900	.02024	.1096	-.1106
LEPS		1.7732	9.800	.02010	.1092	-.1064

Table 1.1. One Parameter 2-OCL Fit to E_{sp}. δ^o is in floating point in (kcal/mole)$^{-}$1. *Rsp* and κ's are in atomic units while E_{sp} is in kcal/mole. LEPS surface with Sato parameter of 0.02636 is also given.

To get a better idea of the correction involved over the entire surface, the difference of the resulting energy and the London Equation is plotted in Figures 1.11 (linear) and 1.12 (bent). This negative correction most accurately represents linear geometries and is less accurate for bent geometries where explicit three-body terms are more important.

As mentioned above, the exchange energy ($\frac{1}{2}(E^T - E^S)$) and overlap squared (S^2) are known to have the same exponential decay. However, inspection of the H^2 integrals of Subsection A.II shows that each has a dependence on various powers of R. The correspondence between the two might well include terms in $\frac{1}{R}$, R, R^2, etc. To give a simple example of the extendibility of the OCL part of the GLP, a fit of the form:

$$S^2 = (\delta_0^o + \delta_1^o R)E^x \tag{1.28}$$

was made. This enables a simultaneous fit of E_{sp} and R_{sp}. Fixing $E_{sp} = 9.8$ kcal/mole and $R_{sp} = 1.757a_o$ yielded a surface similar to the above one-parameter fits with $\kappa_b = .02119$, $\kappa_s = 0.1089$, and $\kappa_a = -0.1028$. The optimal form of S^2

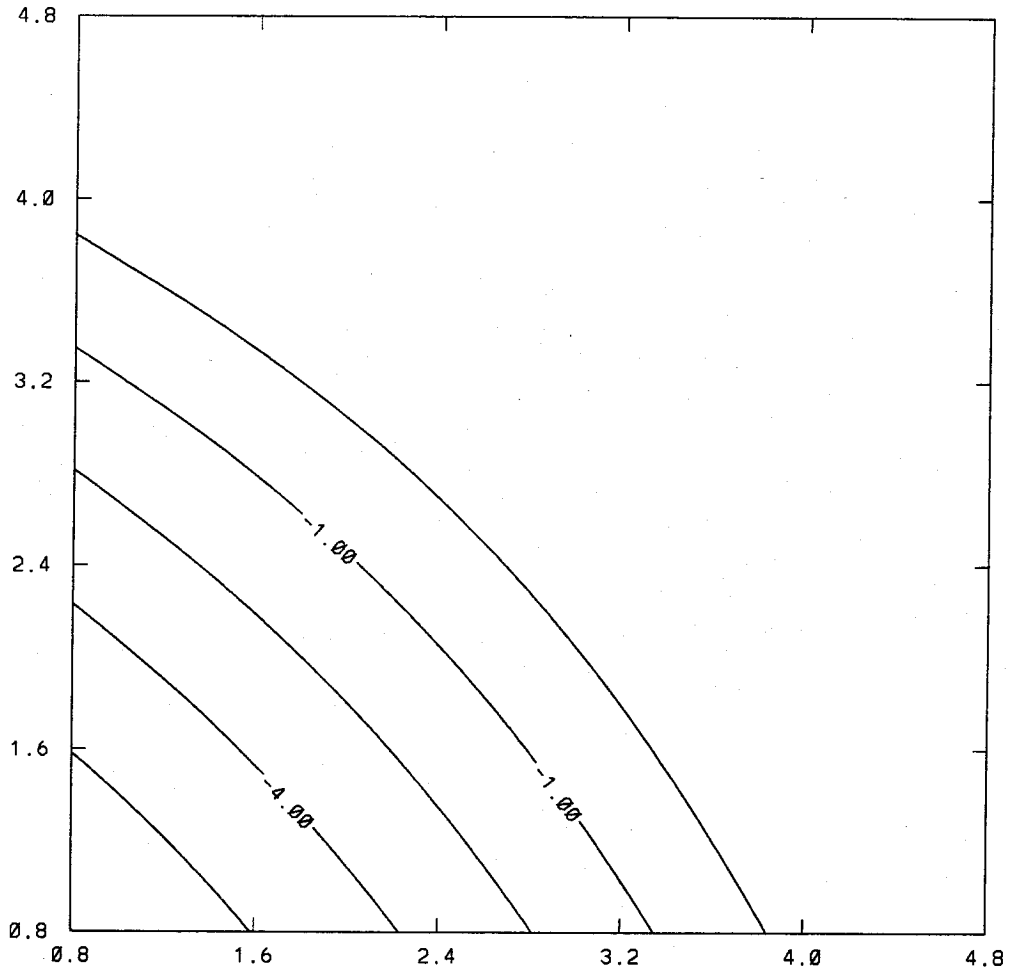


Figure 1.11. Linear 2-OCL Minus London Contour Plot. δ° set so $E_{sp} = 9.8$ kcal/mole.

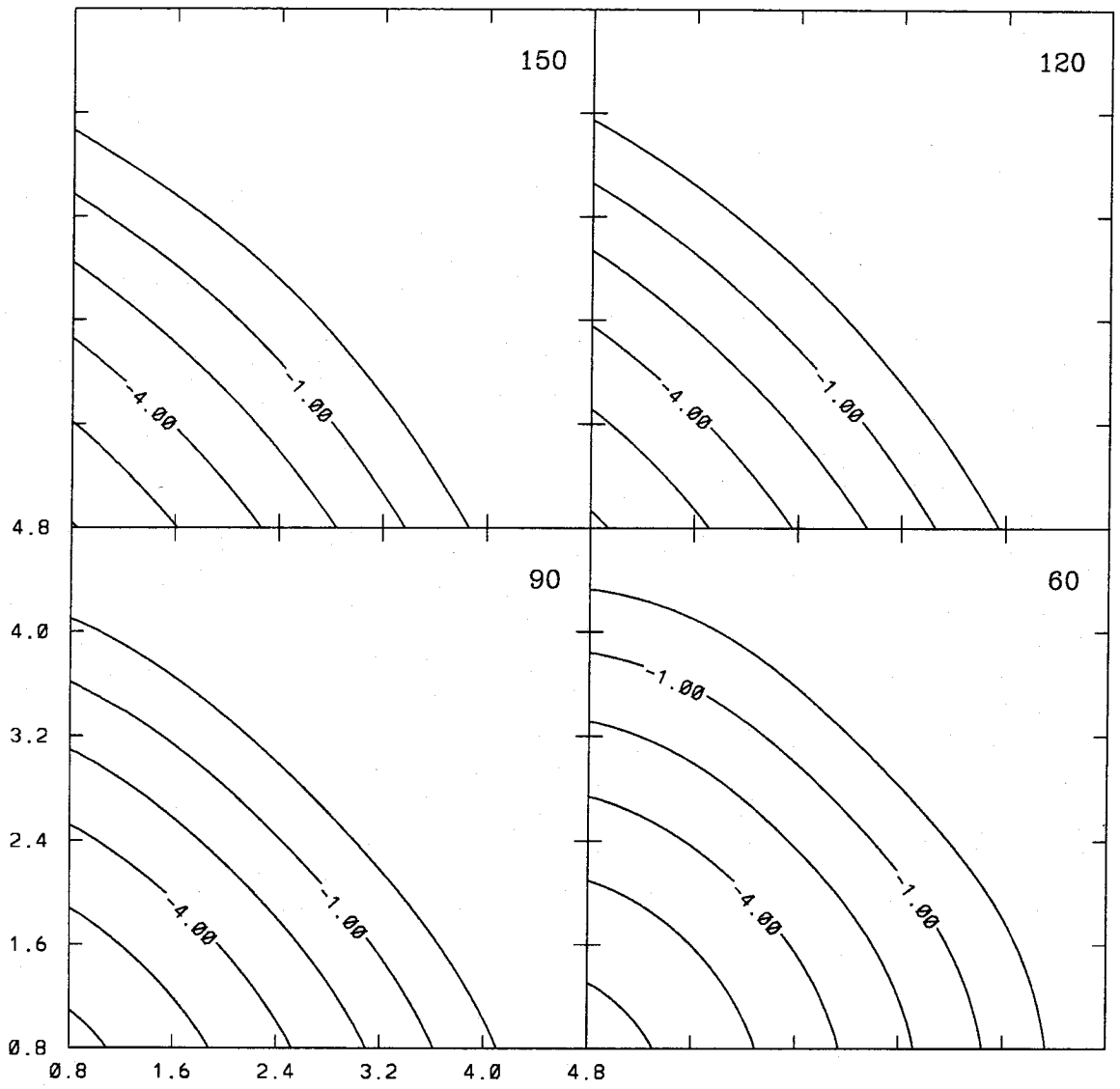


Figure 1.12. Bent 2-OCL Minus London Contour Plot. δ^o set so $E_{sp} = 9.8$ kcal/mole.

in terms of E^x for the entire potential energy surface may well differ, however the ability to do simple one-parameter fits if just an energy barrier is available and to extend the form of even just this one part of the Generalized London Potential to account for additional information is demonstrated.

In this section we have concentrated on the saddle point region since it is the most studied part of any reaction and overlap is the dominant correction there. We must also consider the consequences along the reaction path. Figure 1.13 shows that our underestimate of the energy along the reaction path (too "thin" a barrier) has grown worse by including an overlap correction. The LEPS surface suffers the same malady. A minor part of the problem is that a correction second order in overlap (involving pair-pair expressions) has been forced to account for all the approximations of the London Potential including corrections third and higher order in overlap (involving 3-pair expressions). This leads to an overcorrection along the reaction path. A major part of the problem is the lack of any correction for the errors of including all of the dispersion energy of three H_2 pairs. This correction is shown to be also second order in overlap and involves pair-pair terms in the next section.

V.B. Dispersion Correction

The remaining corrections derived in Appendices A and B are third order in overlap and involve all three pairs. Not contained in that valence bond derivation, however, is any treatment of dispersion energy. The dispersion energy is due primarily to dynamic (as opposed to static) electron correlation and charge transfer. CI calculations on H_2 do include dispersion energy. Thus their use in H_3 energy expressions amounts to including some treatment of dispersion energy as necessary to achieve the correct $H_2 + H$ limits and implicitly assumes that the dispersion energy of H_3 is the pairwise sum of the dispersion energies of the three H_2 pairs. This assump-

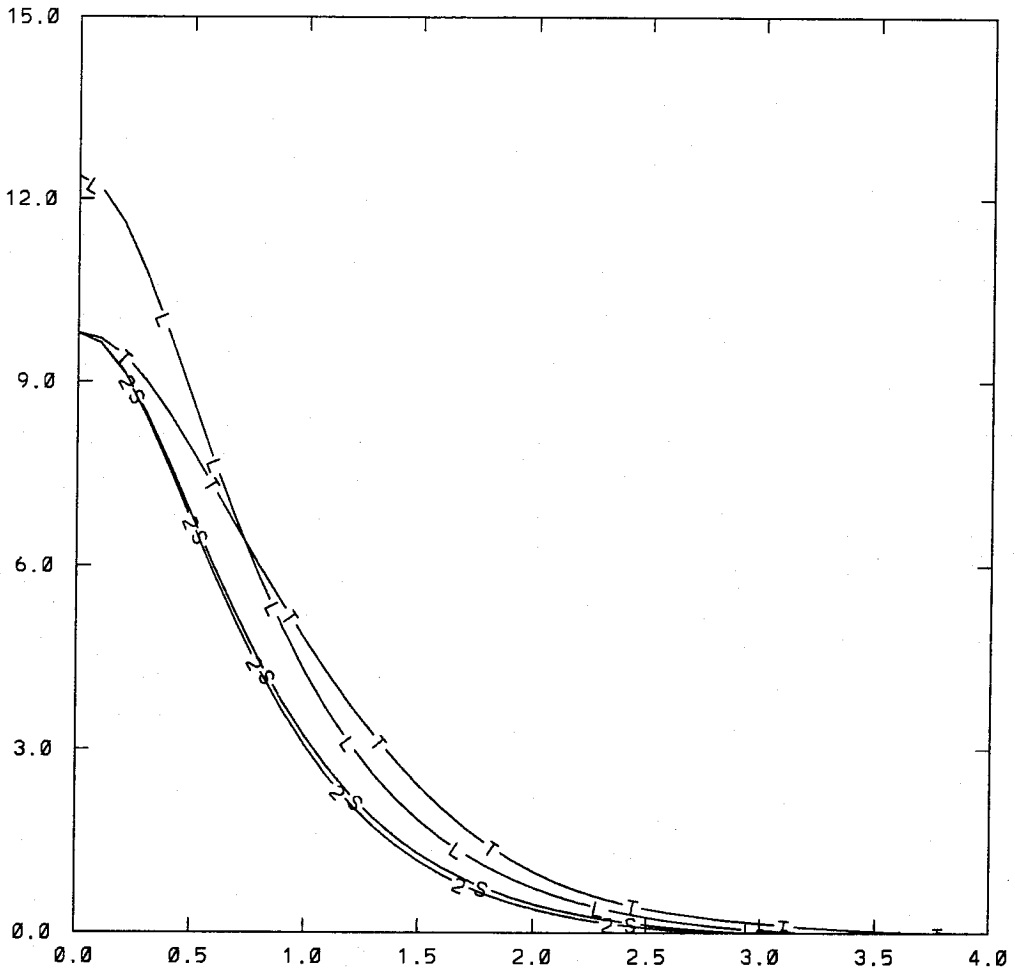


Figure 1.13. Reaction Path Energy: 2-OCL and LEPS.

tion can be considered correct when considering one-pair energies only. Thus it is logical to consider a second-order (pair-pair) dispersion correction similar to the pair-pair overlap correction of the previous subsection. While it might seem a deviation to consider dispersion now and leave the valence bond corrections, dispersion is the second (and final) second-order correction and is thus properly treated next in progressing through the hierarchy of approximations of the London Equation.

There are two equivalent methods of obtaining the dispersion correction, one concentrating on the interaction of exchange energies, the other concentrating on the interaction of singlet and triplet character. The first starts with the Coulomb/exchange version of the spin-coupled London Equation, (1.18). Consider the pairs i and j and their effect on each other. If their exchange corrections to the classical energy are in the same direction ($\cos\gamma_i$ and $\cos\gamma_j$ have the same sign), then the dispersion of the two pairs interacts favorably and net dispersion is increased (energy is more negative). If their exchange corrections are in opposite directions ($\cos\gamma_i$ and $\cos\gamma_j$ have opposite signs), then the dispersion of the two pairs interacts disfavorably and net dispersion is decreased (energy is less negative). Thus the pair i and pair j dispersion interaction should go as $\cos\gamma_i \cos\gamma_j$. An alternative derivation based on singlet and triplet character similarly states that pair-pair dispersion due to triplet-triplet or singlet-singlet combinations are reinforcing while that due to triplet-singlet combinations is conflicting. Equation (1.15) then provides the form of the correction as $f_i^t f_j^t + f_i^s f_j^s - f_i^t f_j^s - f_i^s f_j^t$. However, since the London spin coupling is correct through second order, this is easily reduced to $\cos\gamma_i \cos\gamma_j$.

This spin coupling analysis provides the nature of the correction but still not the full functional form. Having already conjectured a pair-pair interaction, it is appropriate to consider it an overlap squared correction and we use the form $S_i^2 E_j^d$

by defining E_j^d as the dispersion energy of pair j. Thus the full correction is

$$E = E_{OCL} + \delta^d \sum_{i \neq j} \cos\gamma_i \cos\gamma_j S_i^2 E_j^d, \tag{1.29}$$

where E_{OCL} is either the London Potential or the Overlap Corrected London Potential described in the previous subsection. δ^d is an adjustable parameter to account for the amount of dispersion energy that is actually lost or gained. This requires the use of an explicit dispersion energy E_i^d. One operational definition is the difference between the energies of the best CI calculation and the best Generalized Valence Bond-Perfect Pairing (GVB-PP) [37] calculation, i.e.,

$$E_i^d = E_i^{CI} - E_i^{GVB}. \tag{1.30}$$

The CI energies are already needed as input, and the GVB calculations require less effort and computer time so the increase of input is minimal. Since both singlet and triplet character are included in each pair of atoms, the most logical definition of the overall dispersion energy of pair i is

$$E_i^d = f_i^T E_i^{dT} + f_i^S E_i^{dS}, \tag{1.31}$$

where E^{dT} is the difference of triplet CI and GVB energies, i.e., triplet dispersion, and E^{dS} is the difference of singlet CI and GVB energies, i.e., singlet dispersion. We make use of the GVB and CI calculations of Musgrave et al. [38]. The difference between the CI H_2 calculations of Kolos and Wolniewicz [36] and these GVB calculations for H_2 are plotted in Figure 1.14. E_i^{dS} is larger in magnitude than E_i^{dT}, most likely due to the addition of charge transfer energy in addition to the dynamic correlation energy present in both.

We are now ready to implement the entire second order Generalized London Potential (2-GLP) including both pair-pair overlap and dispersion corrections. The

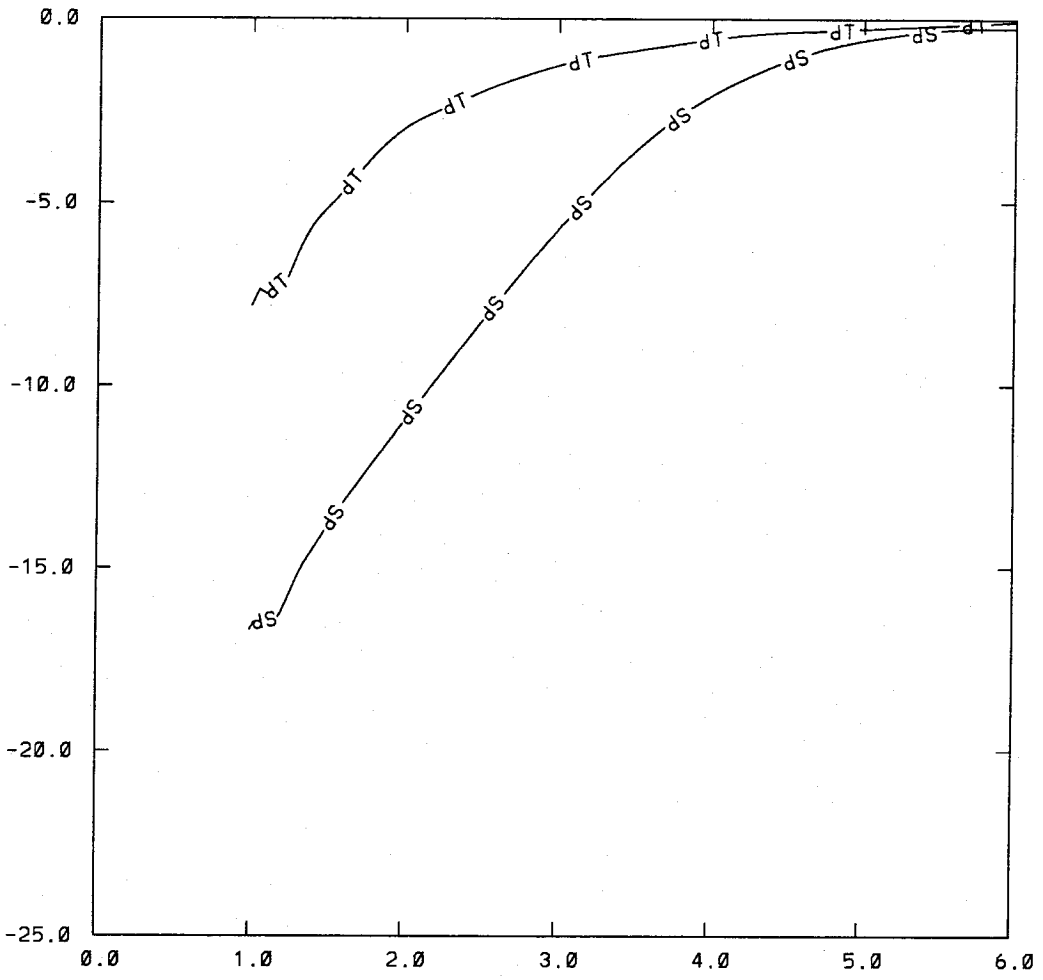


Figure 1.14. Dispersion Energy for Singlet (dS) and Triplet (dT) H_2. Energies are in kcal/mole and distances are in atomic units.

Method	δ^o	δ^d	R_{sp}	E_{sp}	κ_b	κ_s	κ_a
Reference [31]			1.759	9.65		.1092	
Reference [28]			1.757	9.80		.1067	-.058
Reference [29]			1.757	9.90	.021		
London	0	0	1.7922	12.375	.01991	.1054	-.1320
2-GLP far	1.6603 -4	3.0828 -3	1.7734	9.800	.02153	.1102	-.0572
2-GLP half	1.5296 -4	3.6376 -3	1.7748	9.800	.02176	.1102	-.0479
2-GLP near	1.6716 -4	3.0347 -3	1.7733	9.800	.02151	.1102	-.0580
2-GLP near	1.8279 -4	2.9516 -3	1.7716	9.650	.02150	.1104	-.0580

Table 1.2. Two Parameter 2-GLP Fits. E_{sp} and one other point on the LSTH reaction were matched. Again, E_{sp} is in kcal/mole and other quantities are in atomic units. Far used $R_1 = 4.7804$, $R_2 = 1.402$, and $E = 0.1172$; half used $R_1 = 2.6810$, $R_2 = 1.4382$, and $E = 4.791$; near matched κ_a.

simplest form has one parameter for each correction, namely δ^o and δ^d. In order to maintain better separability of the parameters, we can replace S_i^2 with E_i^x in Equation (1.29) at this level. In order to demonstrate the stability of the model, the two parameters are fitted to four separate sets of data. Three have the same barrier height of 9.8 kcal/mole but otherwise energies in distinct regions of the reaction path (of the LSTH surface):

- Far into the channel until the energy is only 0.12 kcal/mole.
- Halfway into the channel until the energy is 4.79 kcal/mole.
- Near the saddle point so that κ_a is matched.

An additional fit with $E_{sp} = 9.65$ kcal/mole is given for the last case. The results are given in Table 1.2. The stability of the surface as the parameters are changed and the stability of the parameters as the input is changed demonstrates the usefulness of the 2-GLP to be used with any available data in any form. See Figure 1.15 for a comparison of the reaction paths and Figure 1.16 for a comparison of the energies

along the reaction path. Part of the deviation of the near (κ_a) and far fits from the LSTH path is again due to overlap being required to handle all three-body corrections at the saddle point, and thus overcorrecting in the channel, and part due to the saddle point distance being too large, and thus the distance traveled along the reaction path is too short compared with the physical coordinates. Figure 1.15 shows that the reaction path is always on the far side of the LSTH path and does not converge as quickly to the LSTH path as does the London path. This is actually beneficial since all distances are shortened when three-body effects are added. Another contribution may come from an overestimate of the dispersion energy. The CI calculations on H_2 and H_3 use a much larger basis set than does the GVB calculation on H_2, and a GVB calculation using the same H_2 basis set would decrease the dispersion energy, possibly effecting the energy along the reaction path. Despite this small deviation from the LSTH reaction path, the small deviation of the 2-GLP paths from each other confirms the stability of the model.

In order to implement the dispersion correction, an estimate of the dispersion energy required an addition of GVB H_2 energies in addition to the CI values. The GVB calculations are simpler than the CI calculations so obtaining them is usually not a problem. However, a very rough estimate of dispersion energy can be obtained by scaling the exchange energy. Not distinguishing between singlet and triplet dispersion and fitting just $E_{sp} = 9.8$ kcal/mole and $\kappa_a = 0.058$ gives the other saddle point properties as $R_{sp} = 1.7757$, $\kappa_b = 0.02225$, and $\kappa_s = 0.1111$. The energy along the reaction path is of similar quality to the fit to E_{sp} and κ_a above.

In summary, this subsection introduced the pair-pair dispersion correction which is the second and final correction which involves pair-pair interactions and is second order in overlap. Using just two parameters, the energy along the reaction path is well represented. While an estimate of the energy barrier is the starting point for

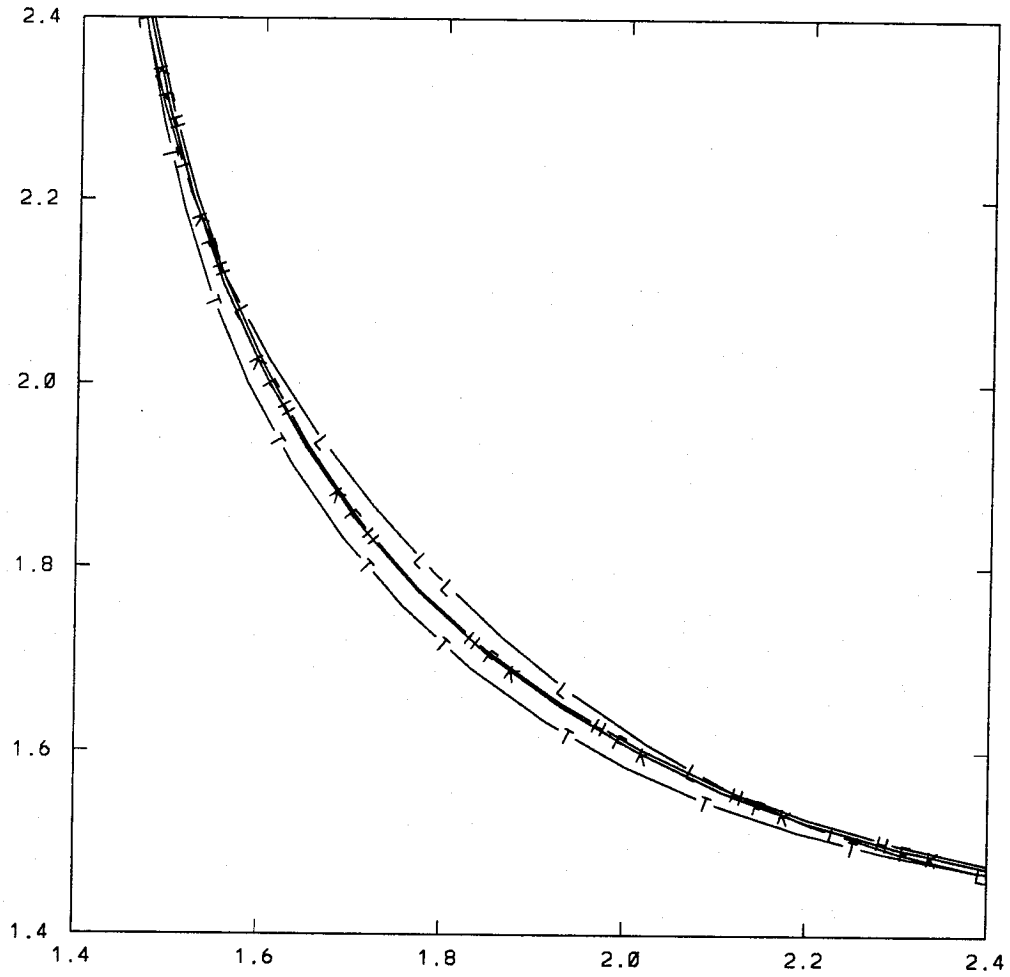


Figure 1.15. Reaction Paths for 2-GLP. L is for London, T is for LSTH, H is for halfway (in energy), F is for far, and K is for κ_a (near) fits.

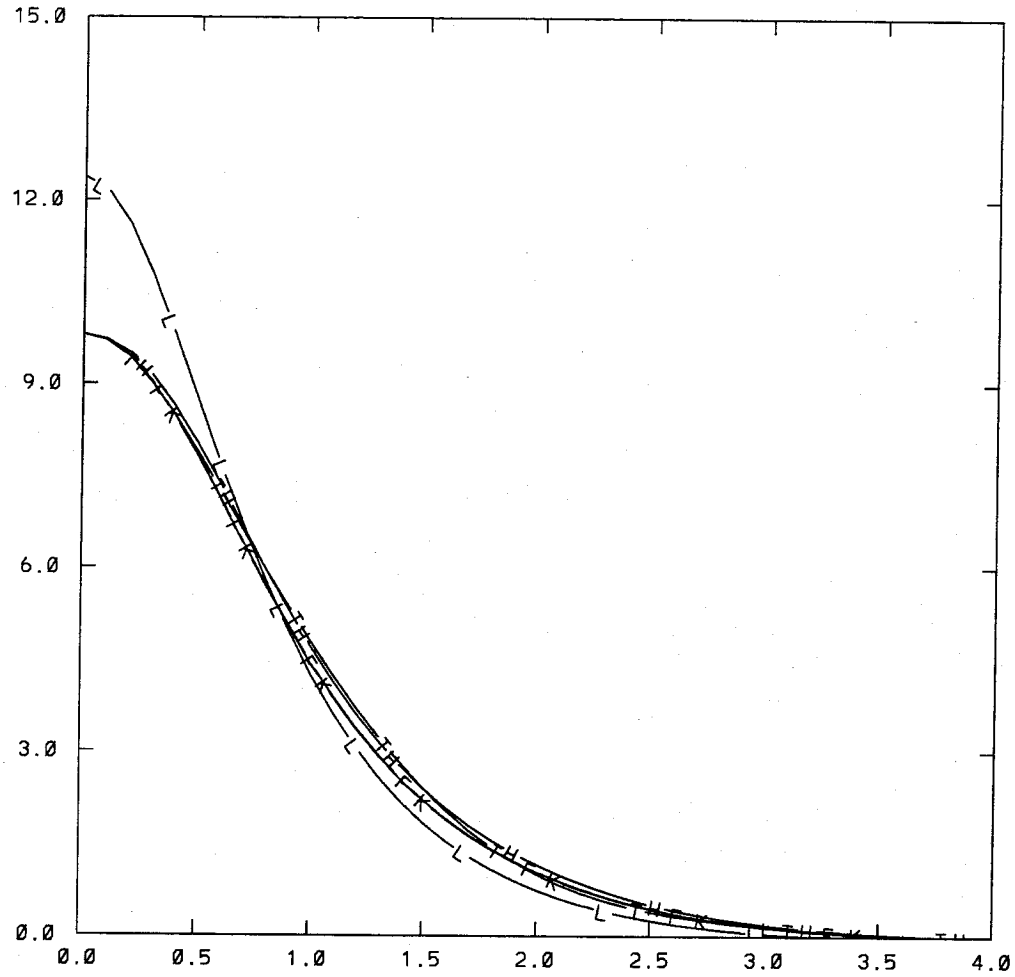


Figure 1.16. Reaction Path Energies for 2-GLP. L is for London, T is for LSTH, H is for halfway (in energy), F is for far, and K is for κ_a (near) fits.

all improvements to the London Potential, stability of the model with respect to various other input has been demonstrated. Our method of starting with the actual CI singlet and triplet curves and correcting for overlap and dispersion with just one parameter each provides a good fit over much of the linear H_3 surface except when all three atoms are close. This presents a strong alterative to using a parameterized triplet which in itself requires four or more parameters and then requires additional asymmetric terms to reasonably represent linear geometries. Using physically motivated and, in the case of overlap, rigorously derivable corrections to the London Potential is shown to have very reasonable results for the amount of input given. Additional parameters could be added in the overlap and dispersion corrections as was done in the previous subsection for just overlap by considering expansions in terms of pair distances. We instead briefly consider higher order corrections to aid the fit of when all three atoms are near and three-body interactions are strongest next.

V.C. Higher Order Corrections

Higher order corrections require additional information, but the goal of this research is to use very limited input to our model. The third order overlap (3-pair) corrections of Appendix Subsection B.I are mainly minor corrections and the gain in accuracy in limited when weighed against the required additional input. However, their appropriate form and their effect on the surface is discussed here in order to gain better insight of the next order valence bond corrections. More important to the saddle point region and to bent geometries is the effect of higher order dispersion. The fit of the London Potential is poorest when all three atoms are within a radius of three to four bohr, primarily due to high electron interaction. A possible method of correcting for this and its implementation using only one more piece of

saddle point information is discussed.

Of the two third order corrections in Appendix Subsection B.I, the one defined as A is the easier to handle. Like all previous corrections, it is independent of spin-coupling which thus retains its London form. The first term of Equation (1.100) is explicitly in terms of $S_1S_2S_3$ times atomic energies. Each of the terms in parentheses in the following terms is known to have the same exponential decay as overlap from Appendix Subsection A.II. Thus, a first order approximation for A is

$$A = A_0 S_1 S_2 S_3, \tag{1.32}$$

where A_0 is an adjustable constant. To gain insight on the effect of this term, it is used alone to provide an energy barrier of E_{sp} = 9.8 kcal/mole. Again, the actual parameter is found after replacing S_i^2 by E_i^x in order to maintain parameter independence. Thus with A_0 =7.705 -4 (kcal/mole)$^{-1}$ we obtain E_{sp} = 9.8 kcal/mole, R_{sp} = 1.7649, κ_b = 0.01860, κ_s = 0.1106, and κ_a = −0.1093. The effect of the correction is shown by Figures 1.17 and 1.18. Comparison to Figures 1.11 and 1.12 shows that the change from the London Potential is seen to be similar to that of the second order overlap correction of Subsection V.A for linear geometries but more highly concentrated in symmetric bent geometries. Thus the gains would be modest in including this correction plus the overlap correction. However, it is an additional extension that is part of the GLP methodology. Even this correction itself is extendible, as Equation (1.100) suggests that a more general fit may be

$$A = A_0 S_1 S_2 S_3 + \sum_{i \le j} S_i S_j S_k f(R_k), \tag{1.33}$$

to account for the deviations of $\Delta_{ab} + A_{ab} + B_{ab}$ from S_{ab}, even though both have the same exponential decay.

The other third order term from valence bond theory, F_{ab} is defined in Equation (1.99) and since it follows the exchange energy E_{ab}, it is the first correction that

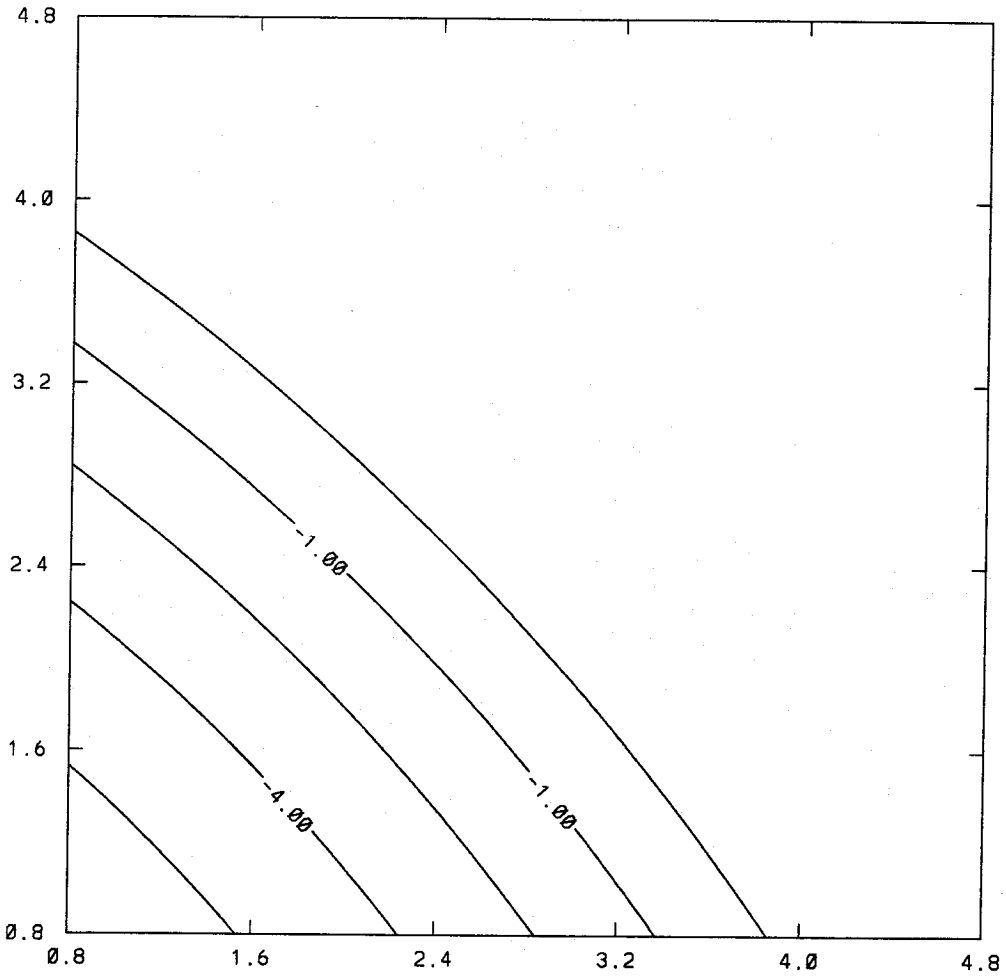


Figure 1.17. Linear London with A_0 Minus London Contour Plot. A_0 set so $E_{sp} = 9.8$ kcal/mole.

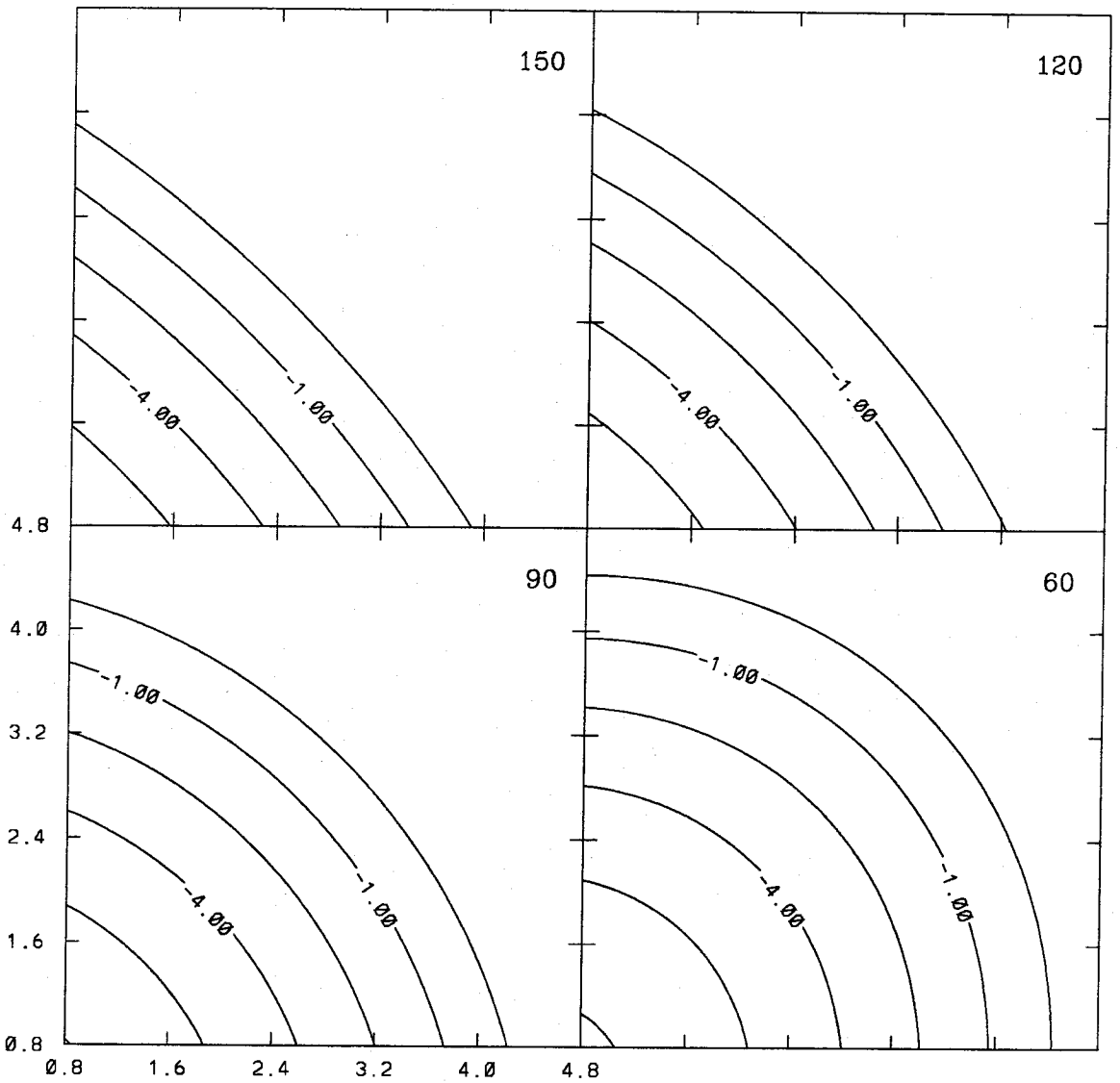


Figure 1.18. Bent London with A_0 Minus London Contour Plot. A_0 set so $E_{sp} = 9.8$ kcal/mole.

can effect the spin coupling. This also has a three-pair dependence, although it is less obvious. Both the C_{ab} and $(ab|cc)$ terms fall off exponentially with both R_{ac} and R_{bc}. Thus the simplest approximation for F_{ab} (in terms of pairs i, j, and k) is $S_3S_2S_1$. However, the H_3 exchange energy term of Equation (1.131) can be written as

$$\overline{X} = \left[\frac{1}{2}\sum_{i<j}(E_i^y - E_j^y)^2\right]. \tag{1.34}$$

Since $E_i^y = E_i^x + F_i$ enters the H_3 energy only as pairwise differences, such a simple F_i form would exactly cancel. Again, terms higher order in R_i could be used to provide flexibility of the Generalized London Potential.

The most difficult correction to make is for higher order dispersion. It is most important for bent geometries, however it is also important in determining the position of the saddle point if not the energy there. One rough estimate is $\delta^3 E_1^x E_2^x E_3^x$. Combining this with the second order overlap and dispersion corrections allows all three of the quantities E_{sp}, κ_a, and R_{sp} to be fit. Note that this information is attainable from only a handful of CI calculations near the saddle point. Fitting $E_{sp} = 9.800$ kcal/mole, $R_{sp} = 1.757$, and $\kappa_a = 0.0580$, leads to $\kappa_b = 0.0197$ and $\kappa_s = 0.1065$. The reaction path and energy along the reaction path are shown for this fit in Figures 1.19 and 1.20. Thus with just three parameters based on information only at the saddle point, we have described the H_3 potential energy surface to remarkable accuracy along the minimum energy reaction path and, in fact, at all lower energy geometries.

VI. Application to Hydrocarbon Reactions

The Generalized London Potential has proven effective in modeling $H_2 + H$ by comparison to accurate *ab initio* calculations. For systems without as much available

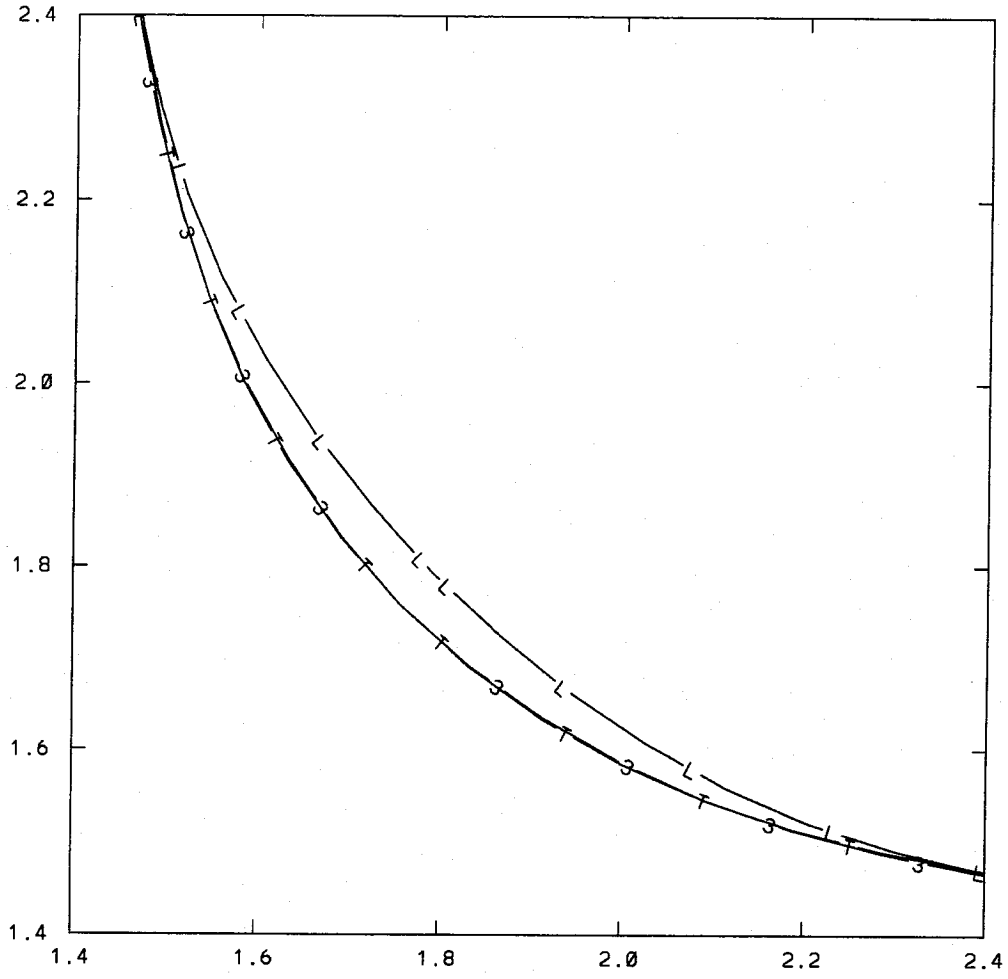


Figure 1.19. Reaction Paths for GLP with 3-Pair Dispersion. L is for London, T is for LSTH, and 3 is for GLP with three-pair dispersion.

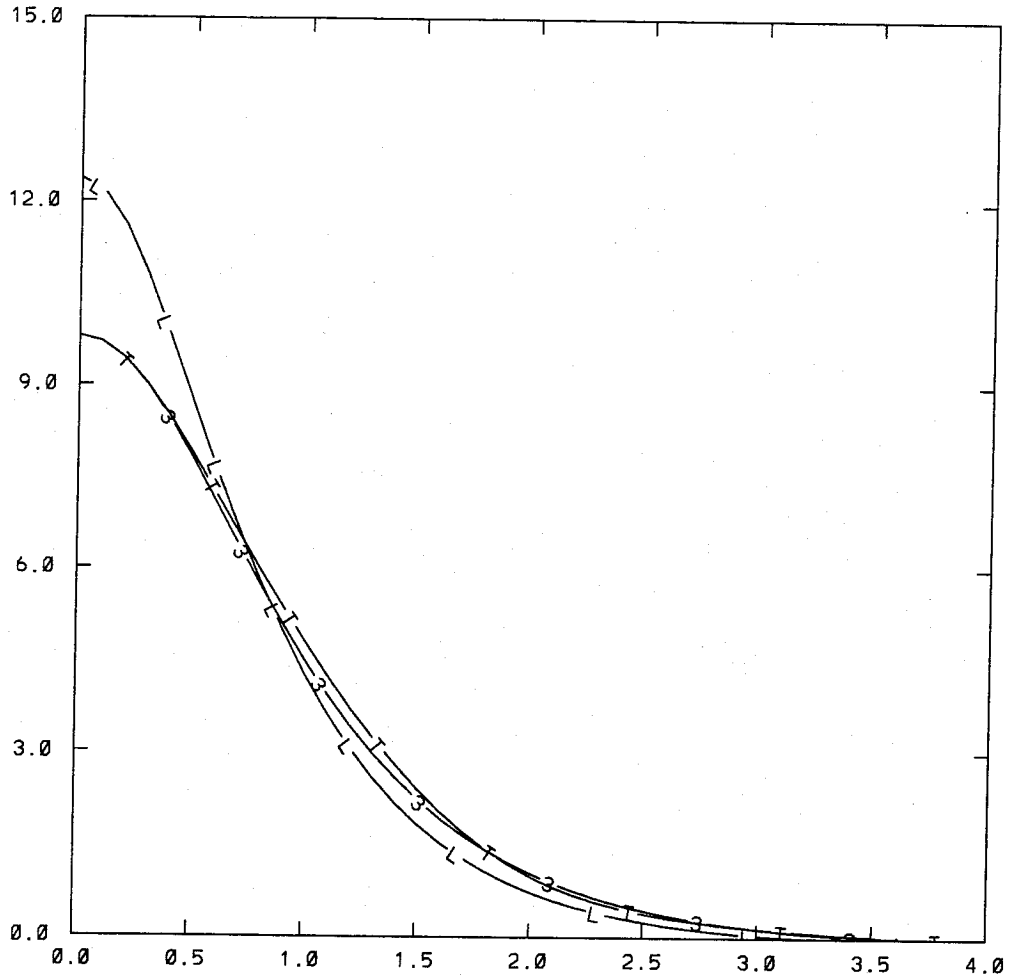


Figure 1.20. Reaction Path Energies for GLP with 3-Pair Dispersion. L is for London, T is for LSTH, and 3 is for GLP with three-pair dispersion.

data, it is hoped that the GLP can predict the entire potential energy surface using only two-body potentials plus a limited amount of three-body information. Actually, the GLP may work best as a partner with *ab initio* calculations in reducing the amount of work required to find the saddle point. An expected sequence would be:

- Calculation of two-body curves.
- Use of London Potential to predict saddle point.
- Use of GLP using parameters estimated from similar systems to improve estimate of saddle point.
- *Ab initio* calculation of several energies near saddle point.
- Incorporation of new information into GLP to improve estimates of entire potential energy surface and saddle point.
- Possible refinement of GLP parameters using a small number of additional *ab initio* points.
- Use of GLP for simulations.

Our overall objective is the introduction of reactive dynamics into molecular modeling. This is especially useful for polymer and crystalline systems, and important carbon-based examples exist for both of these. Thus, our first application is hydrogen exchange among hydrocarbons. We first consider the abstraction of one hydrogen from methane by an atomic hydrogen:

$$CH_4 + H \rightarrow CH_3 + H_2.$$

Actually, this reaction as written is uphill in energy by several kcal/mole, however this is not true for most hydrogen abstractions from hydrocarbons by a lone hydrogen

atom. Thus, we write this equation in this direction, but it is more appropriate to talk of it in the reverse direction for this system only. With the CH_4 bond angles fixed, this is a first approximation of hydrogen abstraction from a diamond surface which can be improved upon by considering abstraction from isobutane or even larger clusters. A variable CH_4 bond angle adds an extra degree of freedom to this system not present in H_3. This is more appropriate for smaller systems and polymers. For either case (fixed or variable angle), the two-body curves are required. We present them next before their use in the GLP.

VI.A. H_2 and CH_4 Potentials

When all of our information is based on *ab initio* calculations, it is best to use a consistent level of calculation for all potential surfaces. The $CH_4 + H$ calculations of Musgrave et al. on the saddle point of $CH_4 + H$ [39] are not to high accuracy, however a similar level calculation is feasible on larger systems. Calculations at a similar level of accuracy were done for H_2 and CH_4 [38] and are reported in Tables 1.3 and 1.4 for H_2 and CH_4, respectively. The tetrahedral geometries are included for the fixed-angle, surface-like atom case. For a lone CH_4, the angle between the H whose bond is being stretched and the other hydrogens relaxes during the course of the reaction with H. For lone CH_3, about 7 kcal/mole is required to go from planar to tetrahedral. Calculations to find bond energies at various angles are possible which can then be used to do the general relaxed angle case. To simplify matters here, we report the energies at 104° which is about the optimized saddle point geometry [39].

R	E^S	E^T
0.4	-0.96387083	
0.5	-1.09183076	-0.57411048
0.6	-1.14695975	-0.68374487
0.7	-1.16604573	-0.75632627
0.74	-1.16781588	
0.8	-1.16641530	-0.80860353
0.9	-1.15698075	-0.84849745
1.0	-1.14262579	-0.87990173
1.2	-1.10924298	-0.92499045
1.4	-1.07765713	
1.5	-1.06391438	-0.96391211
1.6	-1.05181371	
1.8	-1.03262628	-0.98315379
2.0	-1.01952662	-0.99000541
2.5	-1.00460386	-0.99738858
3.0	-1.00090262	-0.99930581
5.0	-0.99988367	-0.99988112
50.0	-0.99988096	-0.99988096

Table 1.3. Two-Body Potentials for H_2. Distances are in Å (note that so far we have dealt mainly in bohr), and energies are in hartrees.

	Tetrahedral		104°	
R	E^S	E^T	E^S	E^T
0.80	-40.25584723	-39.87551304	-40.25047509	-39.86470574
0.90	-40.32346617	-39.94289320	-40.31871254	-39.92518046
1.00	-40.35288348	-39.95152791	-40.34895780	-39.96925500
1.05	-40.35863390	-39.96678868	-40.35518093	-39.98352944
1.09	-40.36023885	-39.97693589	-40.35718587	-39.99279529
1.15	-40.35887945	-39.99059157	-40.35645516	-40.00497786
1.20	-40.35516185	-40.00201942	-40.35328124	-40.01514904
1.25	-40.34974849	-40.01428129	-40.34842353	-40.02631230
1.30	-40.34311129	-40.02727004	-40.34234854	-40.03843393
1.40	-40.32756398	-40.05334920	-40.32792504	-40.06325829
1.50	-40.31062243	-40.07708134	-40.31207546	-40.08607365
2.00	-40.23606332	-40.14743236	-40.24161279	-40.15436516
2.50	-40.19695719	-40.16986183	-40.20385300	-40.17646641
3.00	-40.18315200	-40.17612653	-40.19006031	-40.18274728
4.00	-40.17853794	-40.17815402	-40.18525719	-40.18483343
6.00	-40.17824863	-40.17824772	-40.18493831	-40.18493718
50.00	-40.17824051	-40.17824051	-40.18493026	-40.18493025

Table 1.4. Two-Body Potentials for CH_4. Distances are in Å, and energies are in hartrees.

VI.B. Fixed Angle $CH_4 + H$

As an example of the usefulness of the GLP for crystalline systems, we consider the first approximation of using CH_4 with the bond angle fixed at tetrahedral and describe the resulting potential energy surface. Spline fits to the H_2 and CH_4 energies of the previous subsection are our starting point. The London Potential for this system is shown in Figure 1.21. The zero of energy is taken as zero in the $CH_3 + H_2$ limit. While the CH_4 and H_2 bond energies are only several kcal/mole apart, adding the energy to bend CH_3 to tetrahedral makes the energy difference in the channels about 9 kcal/mole. As expected, the saddle point is closer to the higher energy H_2 channel. The corresponding spin coupling is shown in Figure 1.22. This shows that the "transition state" as defined by the spin coupling ($\gamma = 0$) is even further into the H_2 valley than the saddle point. The smooth variation of γ along the reaction path is again seen as for H_3 in Figure 1.5.

A better approximation of the potential energy surface is obtained by assuming that the parameters of the H_3 GLP are about on the order of those for $CH_4 + H$. Given the similar bond energies, this should be a reasonable approximation. Using the second order overlap and dispersion corrections, we obtain the new surface of Figure 1.23. Thus, we have estimated the entire potential energy surface for $CH_4 + H$ using only two-body curves and our knowledge of the London Equation error for H_3. With more experience in the variation of the GLP parameters across various systems, we can hopefully establish a methodology of creating such surfaces to a reasonable accuracy with little more than two-body input.

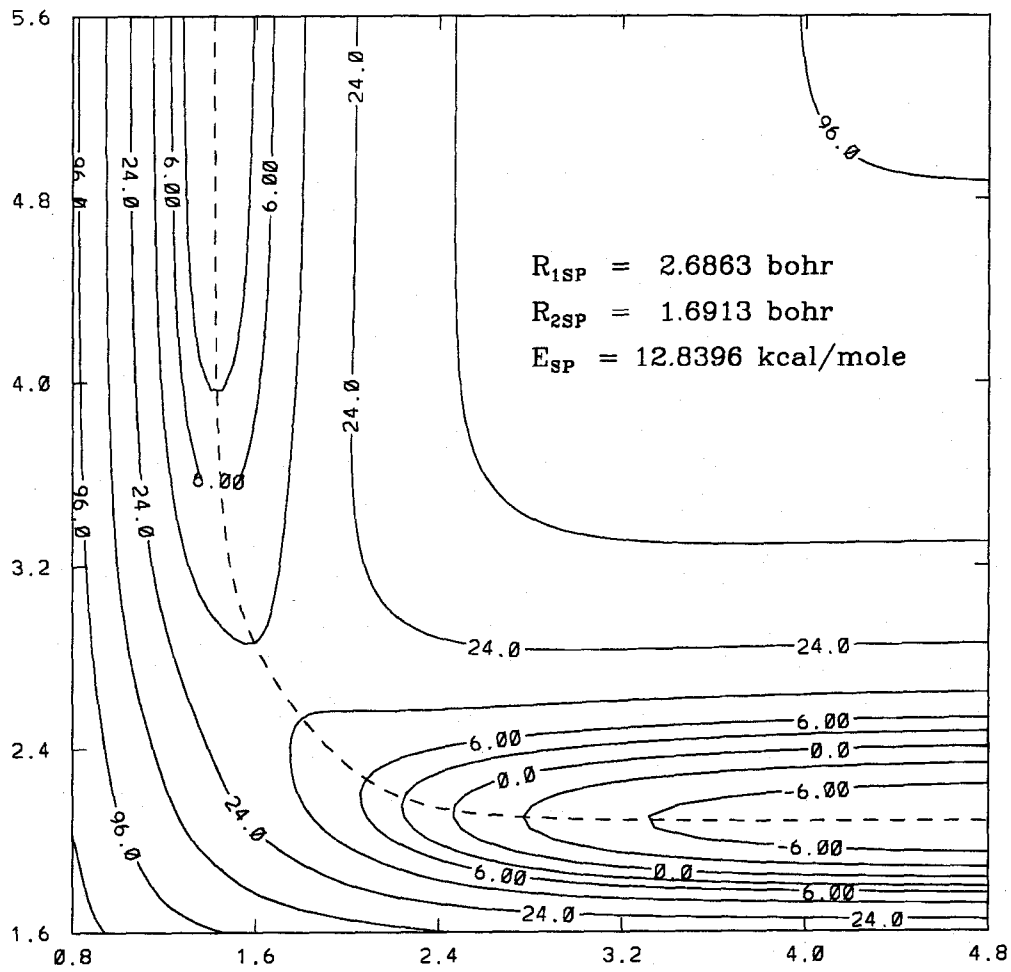


Figure 1.21. $CH_4 + H$ London Potential Contour Plot. The contour values shown are the powers of two times ±3 kcal/mole and zero. Axes are distances in atomic units. H_2 channel is to the upper left and CH_4 channel is to the bottom right.

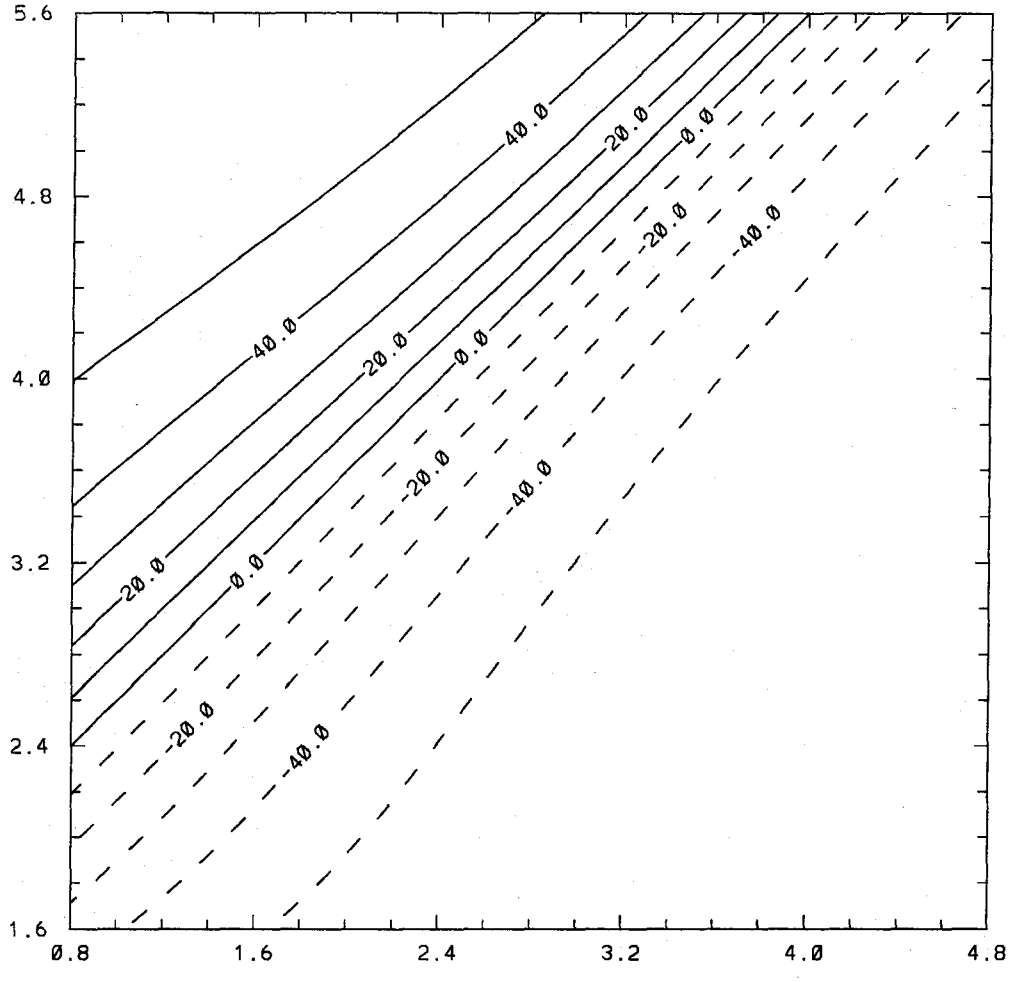


Figure 1.22. Contour Plot of γ for $CH_4 + H$. Contours are spaced every 10 degrees with negative contours endashed.

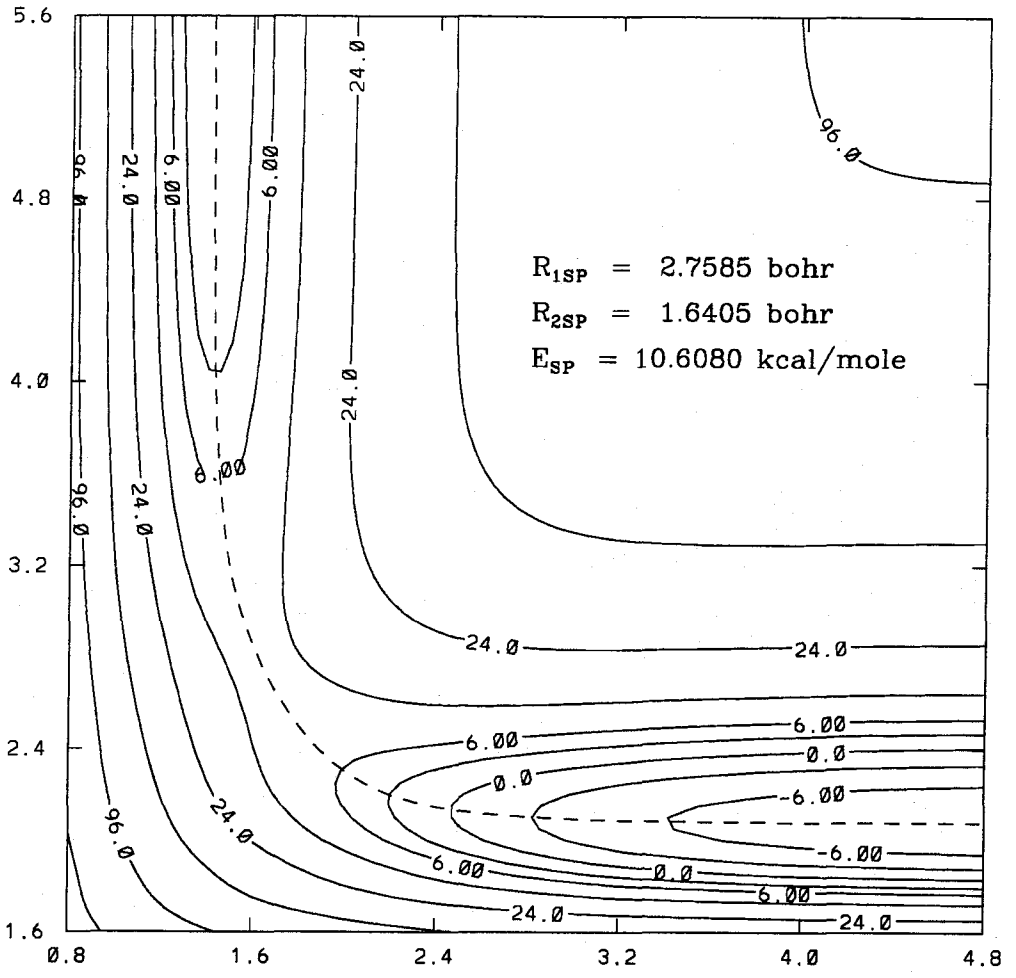


Figure 1.23. $CH_4 + H$ GLP Contour Plot. The contour values shown are the powers of two times ± 3 kcal/mole and zero. Axes are distances in atomic units. H_2 channel is to the upper left and CH_4 channel is to the bottom right.

VI.C. Relaxed Angle $CH_4 + H$

In order to make a comparison with the $CH_4 + H$ results of Reference [39], we need to let the CH_4 bond angle relax as the reaction proceeds. It is possible to calculate the singlet and triplet energies at several angles for each $CH_3 - H$ distance in order to obtain the angular dependence. If the angle relaxed to its singlet minimum, it would be planar at far distances, near tetrahedral at the bond distance, and slightly beyond tetrahedral at shorter distances. At the saddle point distance of about 2.6 bohr, the energy of the singlet is lowest between 106 and 107°. This is in contrast to the optimized angle of 103.8° of Reference [39]. The reason for the discrepancy is that the energy minimum with respect to angle for the triplet is much nearer planar. From the spin coupling γ_i, we know the mix of singlet and triplet character for a given pair of atoms i:

$$E_i = f_i^T E_i^T + f_i^S E_i^S. \tag{1.35}$$

This is the function to be minimized with respect to bond angle at each geometric configuration.

Even without knowledge of the CH_4 triplet and singlet energies at various angles and distances, we can still estimate the energy at the saddle point. Using the 104° energies of CH_4 from Subsection VI.A, the saddle point for the relaxed angle $CH_4 + H$ reaction can be found. This was done for the London Potential and the GLP with second order corrections based on the H_3 parameters. Table 1.5 shows that predicted GLP energy is within 0.3 kcal/mole of the *ab initio* value. Separate H_2 and CH_4 corrections would improve the saddle point position. The information that went into this prediction includes only two-body H_2 and CH_4 potentials, an estimate of two GLP parameters from H_3, and an estimate of the CH_4 bond angle at the saddle point. This last piece of information is also derivable independently

	R_{1sp}	R_{2sp}	E_{sp}
London	2.6095	1.7435	14.9974
GLP	2.6448	1.6998	12.6379
Reference [39]	2.63	1.74	12.91

Table 1.5. Saddle Points for $CH_4 + H$. Distances are in atomic units, and saddle point energies are from H_2 channel in kcal/mole.

by knowing the angular dependence of the separate singlet and triplet curves at the saddle point distance. In addition to this saddle point information, we now have an estimate of the potential energy surface over the entire range of geometries. This can be used to add reactive dynamics to molecular modeling simulations.

VII. Summary

The derivation of the Generalized London Potential is strongly dependent on the spin-coupling formalism. The problem of extending the London Potential to allow for overlap was solved by considering special cases of pure singlet or pure triplet pairing, letting these states mix, correcting for overlap between these states, and carefully replacing the appropriate three-body terms with their corresponding two-body terms. In addition, the spin coupling provides much insight into how an exchange reaction proceeds. Parameterized in terms of γ, spin coupling can be used as a reaction coordinate to describe the transition among various electron pairings. Even without further extension, the spin-coupling treatment of exchange reactions has provided useful insight into the interplay of forces resulting from the Pauli Principle allowing strong bonds to be broken and formed over an energy barrier an order of magnitude smaller. As this is developed within the Heitler-London (valence-bond) formalism, there is no accounting for dispersion energies. The introduction of dispersion terms into the GLP allows for the high accuracy fit of the potential surface for H_3 with

very few parameters.

While the fitting of known surfaces with few parameters is demonstrated, we hope that the GLP finds use in the many more exchange reactions for which little information is known. For the case of $CH_4 + H$, we demonstrated the ability to predict the barrier to reaction using mainly two-body input. We hope to be able to build a database of two-body curves which can be used to model many exchange reactions. With experience with the overlap and dispersion parameters, we may be able to predict GLP potentials without any three-body information. When some three-body information is available, the GLP has shown stability in incorporating that information to improve the entire reaction surface. Letting GLP and quantum chemistry work together may be the best way to systematically find accurate potential energy surfaces. The GLP can be used to estimate saddle points where *ab initio* calculations can be performed. Incorporating this new information into the GLP should allow quick convergence on accurate potential energy surfaces. These surfaces then allow exchange reactions to be introduced into molecular modeling techniques, thereby expanding their usefulness in allowing reactive dynamics for large systems such as polymers and crystals.

A. "Exact" Valence Bond Descriptions

The appendix derives the energies of H_2 and H_3 using the valence bond approach. Here, "exact" means within the valence bond model, i.e., no terms are set to zero.

A.I. Notation

There are many individual terms that comprise the energy expressions derived below. The notation is defined here. In general, the letters a, b, and c denote nuclei and also orbitals centered at that nuclei. ϕ_i is used to refer generically to one of these orbitals. Electrons are specified either generically by i or by the number 1, 2, or 3. Spin functions are denoted by α (up) and β (down). Other notation includes:

$$
\begin{aligned}
\Delta_{\phi_1\phi_2} &\equiv \left\langle \phi_1(i) \left| -\tfrac{1}{2}\nabla_i^2 \right| \phi_2(i) \right\rangle, \\
A_{\phi_1\phi_2} &\equiv \left\langle \phi_1(i) \left| -\frac{1}{r_{ai}} \right| \phi_2(i) \right\rangle, \\
B_{\phi_1\phi_2} &\equiv \left\langle \phi_1(i) \left| -\frac{1}{r_{bi}} \right| \phi_2(i) \right\rangle, \\
C_{\phi_1\phi_2} &\equiv \left\langle \phi_1(i) \left| -\frac{1}{r_{ci}} \right| \phi_2(i) \right\rangle, \\
E_a &\equiv \Delta_{aa} + A_{aa}, \\
(\phi_1\phi_3|\phi_2\phi_4) &\equiv \left\langle \phi_1(1)\phi_2(2) \left| \frac{1}{r_{12}} \right| \phi_3(1)\phi_4(2) \right\rangle, \\
J_{\phi_1\phi_2} &\equiv (\phi_1\phi_1|\phi_2\phi_2), \\
K_{\phi_1\phi_2} &\equiv (\phi_1\phi_2|\phi_2\phi_1), \\
S_{ab} &\equiv \langle a(i)|b(i)\rangle.
\end{aligned}
$$

One must be careful to distinguish orbital overlap S_{ab} from wavefunction overlap S_{12}, the latter being the overlap between Ψ_1 and Ψ_2. Also, E_a is not necessarily the isolated atomic energy of atom a since the orbital $a(i)$ changes if other atoms are near. E_a^0 is used to represent the isolated atomic energy, which, of course, is $-\frac{1}{2}$

hartrees for hydrogen.

Two minor distinctions are made between notation in the appendices and that in the main sections. The first is the substitution of the numerical indices 1, 2, and 3 for the atom pairs bc, ac, and ab, respectively. Further, E^S, E^T, and E^{cl} are defined in the appendices so as to include the atomic energies at dissociation, e.g., $E^S_{ab}(R_{ab}=\infty) = E^0_a + E^0_b$, while the zero of energy is often taken to be zero when all atoms are dissociated in the main sections, e.g., $E^S_1(R_1=\infty)=0$.

A.II. H_2

The energies of the ground and first excited states of H_2 are listed here in terms of the above notation. The wavefunction for the ground singlet state is:

$$^1\Sigma^+_g = \frac{(ab+ba)}{\sqrt{2(1+S^2_{ab})}}\frac{(\alpha\beta-\beta\alpha)}{\sqrt{2}}. \tag{1.36}$$

Using the H_2 Hamiltonian

$$\mathcal{H} = \sum_{i=1}^{2}\left(-\tfrac{1}{2}\nabla^2_i - \tfrac{1}{r_{ai}} - \tfrac{1}{r_{bi}}\right) + \tfrac{1}{r_{12}} + \tfrac{1}{r_{ab}}, \tag{1.37}$$

we find that the singlet energy is:

$$E^S_{ab} = \frac{1}{r_{ab}} + \frac{E_a + B_{aa} + E_b + A_{bb} + J_{ab} + K_{ab} + 2S_{ab}(\Delta_{ab} + A_{ab} + B_{ab})}{1+S^2_{ab}}. \tag{1.38}$$

The wavefunctions for the first excited triplet state are:

$$^3\Sigma^+_u = \frac{(ab-ba)}{\sqrt{2(1-S^2_{ab})}}\begin{Bmatrix} \alpha\alpha \\ \frac{1}{\sqrt{2}}(\alpha\beta+\beta\alpha) \\ \beta\beta \end{Bmatrix}. \tag{1.39}$$

The related triplet energy is:

$$E^T_{ab} = \frac{1}{r_{ab}} + \frac{E_a + B_{aa} + E_b + A_{bb} + J_{ab} - K_{ab} - 2S_{ab}(\Delta_{ab} + A_{ab} + B_{ab})}{1-S^2_{ab}}. \tag{1.40}$$

It is useful to define the classical (Coulomb) and exchange terms as

$$E_{ab}^{cl} \equiv E_a + B_{aa} + E_b + A_{bb} + J_{ab} + \frac{1}{r_{ab}}, \tag{1.41}$$

$$-E_{ab}^{x} \equiv K_{ab} + 2S_{ab}(\Delta_{ab} + A_{ab} + B_{ab}) - S_{ab}^2(E_a + B_{aa} + E_b + A_{bb} + J_{ab}), \tag{1.42}$$

where E_{ab}^x is defined to be a positive quantity. This allows the singlet and triplet energies to be written as:

$$E_{ab}^{S} = E_{ab}^{cl} - \frac{E_{ab}^{x}}{1 + S_{ab}^2}, \tag{1.43}$$

$$E_{ab}^{T} = E_{ab}^{cl} + \frac{E_{ab}^{x}}{1 - S_{ab}^2}. \tag{1.44}$$

The inverted equations are

$$E_{ab}^{cl} = \tfrac{1}{2}[(1 - S_{ab}^2)E_{ab}^{T} + (1 + S_{ab}^2)E_{ab}^{S}], \tag{1.45}$$

$$E_{ab}^{x} = \tfrac{1}{2}(1 - S_{ab}^4)(E_{ab}^{T} - E_{ab}^{S}). \tag{1.46}$$

Note that these definitions of E^{cl} and E^x which include the effect of overlap are distinct from those used in deriving the London Potential when $S_{ab}^2 = 0$.

The terms comprising the classical and exchange energies can be found analytically for scaled "1s" orbitals:

$$a(i) = \sqrt{\frac{\alpha^3}{\pi}} e^{-\alpha r_{ai}}, \tag{1.47}$$

$$b(i) = \sqrt{\frac{\alpha^3}{\pi}} e^{-\alpha r_{bi}}. \tag{1.48}$$

Here, α denotes a variational parameter. Of course, even with α optimized at each nuclear distance r_{ab}, the polarization and dynamic dispersion energies are not included in this simple valence bond model. Precisely that fact is useful, however, since the difference between accurate CI calculations and this description yields an estimate of the dispersion energy. The difference between CI and GVB-PP energies actually yields a better estimate of the dispersion energy since polarization is

included in both. The required integrals in terms of $w = \alpha R_{ab}$ are [40]:

$$\Delta_{aa} = \tfrac{1}{2}\alpha^2, \tag{1.49}$$

$$A_{aa} = -\alpha, \tag{1.50}$$

$$S_{ab} = e^{-w}(1 + w + \tfrac{1}{3}w^2) \tag{1.51}$$

$$B_{aa} = \alpha\left[-\tfrac{1}{w} + e^{-2w}\left(1 + \tfrac{1}{w}\right)\right], \tag{1.52}$$

$$B_{ab} = -\alpha e^{-w}(1 + w), \tag{1.53}$$

$$\Delta_{ab} = \tfrac{1}{2}\alpha^2 e^{-w}(1 + w - \tfrac{1}{3}w^2) = -\alpha B_{ab} - \tfrac{1}{2}\alpha^2 S_{ab}, \tag{1.54}$$

$$J_{ab} = \alpha\left[\tfrac{1}{w} - e^{-2w}\left(\tfrac{1}{w} + \tfrac{11}{8} + \tfrac{3}{2}w + \tfrac{1}{6}w^2\right)\right], \tag{1.55}$$

$$K_{ab} = \tfrac{1}{5}\alpha\left\{-e^{-2w}(\tfrac{25}{8} + \tfrac{23}{4}w + 3w^2 + \tfrac{1}{3}w^3)\right. \tag{1.56}$$

$$\left. + \tfrac{6}{w}[S^2(C + \ln w) + S'^2 E_i(-4w) - 2SS'E_i(-2w)]\right\}, \tag{1.57}$$

where $S' = e^w(1 - w + \frac{1}{3}w^2)$, C is Euler's constant, and E_i is the internal logarithm or exponential integral:

$$E_i(-x) = -\int_x^\infty \frac{e^{-t}}{t}\mathrm{d}t. \tag{1.58}$$

A.III. H_3, Special Cases

Here we consider the energies of two special cases of the doublet ground state of H_3. The first has purely singlet pairing between atoms a and b and thus has one wavefunction

$$\Psi_1 \propto \mathcal{A}[abc\frac{1}{\sqrt{2}}(\alpha\beta\alpha - \beta\alpha\alpha)]. \tag{1.59}$$

The second has purely triplet pairing between atoms a and b and thus has one wavefunction

$$\Psi_2 \propto \mathcal{A}[abc\frac{1}{\sqrt{6}}(\alpha\beta\alpha + \beta\alpha\alpha - 2\alpha\alpha\beta)]. \tag{1.60}$$

Each has a similar wavefunction with two β's and one α to complete the doublet. The antisymmetry operator is given in terms of permutations of the electrons:

$$\mathcal{A} = \frac{1}{\sqrt{6}}[e + (123) + (132)][e - (12)], \tag{1.61}$$

with e the identity operator, $(123)abc = cab$, etc. Let N_1 and N_2 be normalization constants. Expanding the wavefunctions, we get:

$$\Psi_1 = \frac{1}{\sqrt{12}N_1}[e + (123) + (132)][(abc + bac)(\alpha\beta\alpha - \beta\alpha\alpha)], \tag{1.62}$$

$$\begin{aligned} \Psi_1 \;=\; & \frac{1}{\sqrt{12}N_1}[(abc + bac - cab - cba)\alpha\beta\alpha \\ & +(bca + acb - abc - bac)\beta\alpha\alpha \\ & +(cab + cba - bca - acb)\alpha\alpha\beta], \end{aligned} \tag{1.63}$$

$$\begin{aligned} \Psi_1 \;=\; & \frac{1}{\sqrt{12}N_1}\{\alpha\beta\alpha[e + (12) - (123) - (13)]abc \\ & +\beta\alpha\alpha[(132) + (23) - e - (12)]abc \\ & +\alpha\alpha\beta[(123) + (13) - (132) - (23)]abc\}, \end{aligned} \tag{1.64}$$

$$\Psi_2 = \frac{1}{6N_2}[e + (123) + (132)][(abc - bac)(\alpha\beta\alpha + \beta\alpha\alpha - 2\alpha\alpha\beta)], \tag{1.65}$$

$$\begin{aligned} \Psi_2 \;=\; & \frac{1}{6N_2}[(abc - bac + cab - cba - 2bca + 2acb)\alpha\beta\alpha \\ & +(bca - acb + abc - bac - 2cab + 2cba)\beta\alpha\alpha \\ & +(cab - cba + bca - acb - 2abc + 2bac)\alpha\alpha\beta], \end{aligned} \tag{1.66}$$

$$\begin{aligned} \Psi_2 \;=\; & \frac{1}{6N_2}\{\alpha\beta\alpha[e - (12) + (123) - (13) - 2(132) + 2(23)]abc \\ & +\beta\alpha\alpha[(132) - (23) + e - (12) - 2(123) + 2(13)]abc \\ & +\alpha\alpha\beta[(123) - (13) + (132) - (23) - 2e + 2(12)]abc\}. \end{aligned} \tag{1.67}$$

For a spin-independent operator $\mathcal{O}$ symmetric with respect to all permutations with

$$\mathcal{O}_{ij} \equiv \langle \Psi_i | \mathcal{O} | \Psi_j \rangle, \tag{1.68}$$

we find

$$\begin{aligned}
\mathcal{O}_{11} &= \frac{1}{12N_1^2}\{\langle abc|\mathcal{O}|[e+(12)-(132)-(13)][e+(12)-(123)-(13)]abc\rangle \\
&\quad +\langle abc|\mathcal{O}|[(123)+(23)-e-(12)][(132)+(23)-e-(12)]abc\rangle \\
&\quad +\langle abc|\mathcal{O}|[(132)+(13)-(123)-(23)][(123)+(13)-(132)-(23)]abc\rangle\},
\end{aligned} \tag{1.69}$$

$$\mathcal{O}_{11} = \frac{1}{12N_1^2}\langle abc|\mathcal{O}|[12e+12(12)-6(13)-6(23)-6(123)-6(132)]abc\rangle, \tag{1.70}$$

$$\begin{aligned}
\mathcal{O}_{22} &= \frac{1}{36N_2^2}\{\langle abc|\mathcal{O}|[e-(12)+(132)-(13)-2(123)+2(23)]\times \\
&\quad [e-(12)+(123)-(13)-2(132)+2(23)]abc\rangle \\
&\quad +\langle abc|\mathcal{O}|[(123)-(23)+e-(12)-2(132)+2(13)]\times \\
&\quad [(132)-(23)+e-(12)-2(123)+2(13)]abc\rangle \\
&\quad +\langle abc|\mathcal{O}|[(132)-(13)+(123)-(23)-2e+2(12)]\times \\
&\quad [(123)-(13)+(132)-(23)-2e+2(12)]abc\rangle\},
\end{aligned} \tag{1.71}$$

$$\mathcal{O}_{22} = \frac{1}{36N_2^2}\langle abc|\mathcal{O}|[36e-36(12)+18(13)+18(23)-18(123)-18(132)]abc\rangle, \tag{1.72}$$

$$\begin{aligned}
\mathcal{O}_{12} &= \frac{\sqrt{3}}{36N_1N_2}\{\langle abc|\mathcal{O}|[e+(12)-(132)-(13)]\times \\
&\quad [e-(12)+(123)-(13)-2(132)+2(23)]abc\rangle \\
&\quad +\langle abc|\mathcal{O}|[(123)+(23)-e-(12)]\times \\
&\quad [(132)-(23)+e-(12)-2(123)+2(13)]abc\rangle \\
&\quad +\langle abc|\mathcal{O}|[(132)+(13)-(123)-(23)]\times \\
&\quad [(123)-(13)+(132)-(23)-2e+2(12)]abc\rangle\},
\end{aligned} \tag{1.73}$$

$$\mathcal{O}_{12} = \frac{\sqrt{3}}{36 N_1 N_2} \langle abc|\mathcal{O}|[-18(13) + 18(23) + 18(123) - 18(132)]abc\rangle. \qquad (1.74)$$

By setting $\mathcal{O}$ to the identity operator $\mathcal{I}$, we find

$$N_1^2 = 1 + S_{ab}^2 - \tfrac{1}{2}S_{bc}^2 - \tfrac{1}{2}S_{ac}^2 - S_{ab}S_{bc}S_{ac}, \qquad (1.75)$$

$$N_2^2 = 1 - S_{ab}^2 + \tfrac{1}{2}S_{bc}^2 + \tfrac{1}{2}S_{ac}^2 - S_{ab}S_{bc}S_{ac}, \qquad (1.76)$$

$$S_{12} \equiv \langle \Psi_1|\Psi_2\rangle = -\sqrt{3}(S_{ac}^2 - S_{bc}^2)/(2N_1N_2). \qquad (1.77)$$

Using the H_3 Hamiltonian

$$\mathcal{H} = \sum_{i=1}^{3}\left(-\tfrac{1}{2}\nabla_i^2 - \tfrac{1}{r_{ai}} - \tfrac{1}{r_{bi}} - \tfrac{1}{r_{ci}}\right) + \tfrac{1}{r_{12}} + \tfrac{1}{r_{23}} + \tfrac{1}{r_{13}} + \tfrac{1}{r_{ab}} + \tfrac{1}{r_{bc}} + \tfrac{1}{r_{ac}}, \qquad (1.78)$$

the energy $\langle \Psi_1|\mathcal{H}|\Psi_1\rangle$ is found to be

$$\begin{aligned}
\mathcal{H}_{11} = \;& E_a + B_{aa} + E_b + A_{bb} + J_{ab} + \tfrac{1}{r_{ab}} \\
& + E_a + C_{aa} + E_c + A_{cc} + J_{ac} + \tfrac{1}{r_{ac}} \\
& + E_b + C_{bb} + E_c + B_{cc} + J_{bc} + \tfrac{1}{r_{bc}} \\
& - E_c - E_b - E_a \\
+ \{ \;& K_{ab} + 2S_{ab}(\Delta_{ab} + A_{ab} + B_{ab}) - S_{ab}^2(E_a + B_{aa} + E_b + A_{bb} + J_{ab}) \\
& - \tfrac{1}{2}[K_{ac} + 2S_{ac}(\Delta_{ac} + A_{ac} + C_{ac}) - S_{ac}^2(E_a + C_{aa} + E_c + A_{cc} + J_{ac})] \\
& - \tfrac{1}{2}[K_{bc} + 2S_{bc}(\Delta_{bc} + B_{bc} + C_{bc}) - S_{bc}^2(E_b + C_{bb} + E_c + B_{cc} + J_{bc})] \\
& - S_{ab}S_{ac}S_{bc}(-E_c - E_b - E_a) \\
& - S_{bc}S_{ac}(\Delta_{ab} + A_{ab} + B_{ab}) \\
& - S_{ab}S_{bc}(\Delta_{ac} + A_{ac} + C_{ac}) \\
& - S_{ab}S_{ac}(\Delta_{bc} + B_{bc} + C_{bc}) \\
& + 2S_{ab}[C_{ab} + (ab|cc)] - S_{ac}[B_{ac} + (ac|bb)] - S_{bc}[A_{bc} + (bc|aa)] \\
& + (\tfrac{1}{2}S_{bc}^2 + \tfrac{1}{2}S_{ac}^2)J_{ab} + (-S_{ab}^2 + \tfrac{1}{2}S_{bc}^2)J_{ac} + (-S_{ab}^2 + \tfrac{1}{2}S_{ac}^2)J_{bc} \\
& - S_{ab}^2(C_{aa} + C_{bb}) + \tfrac{1}{2}S_{ac}^2(B_{aa} + B_{cc}) + \tfrac{1}{2}S_{bc}^2(A_{bb} + A_{cc}) \\
& - S_{bc}S_{ac}C_{ab} - S_{ab}S_{bc}B_{ac} - S_{ab}S_{ac}A_{bc} \\
& - S_{ab}(ca|cb) - S_{ac}(ba|bc) - S_{bc}(ab|ac) \\
& + S_{ab}S_{ac}S_{bc}(J_{ab} + J_{ac} + J_{bc}) \\
& + S_{ab}S_{ac}S_{bc}(C_{aa} + C_{bb} + B_{aa} + B_{cc} + A_{bb} + A_{cc}) \qquad \}/ \\
\{ \;& 1 + S_{ab}^2 - \tfrac{1}{2}S_{ac}^2 - \tfrac{1}{2}S_{bc}^2 - S_{ab}S_{bc}S_{ac} \qquad \}.
\end{aligned} \qquad (1.79)$$

The terms in $\mathcal{H}_{11}$ have been carefully arranged so that substitution for H_2 E^{cl} and E^x terms is obvious and so that the magnitude of the terms decrease in the order written. The latter is done by grouping terms in the same order of e^{-w} per the H_2 results of the previous subsection. Thus, the five lines starting with $-S_{ab}S_{ac}S_{bc}(-E_c-E_b-E_a)$ are of order e^{-3w}, i.e., overlap cubed, the next four lines go as overlap to the fourth power, and the next two lines go as overlap to the fifth power. The energy $\langle\Psi_2|\mathcal{H}|\Psi_2\rangle$ is equivalently

$$
\begin{aligned}
\mathcal{H}_{22} = \;& E_a + B_{aa} + E_b + A_{bb} + J_{ab} + \tfrac{1}{r_{ab}} \\
& + E_a + C_{aa} + E_c + A_{cc} + J_{ac} + \tfrac{1}{r_{ac}} \\
& + E_b + C_{bb} + E_c + B_{cc} + J_{bc} + \tfrac{1}{r_{bc}} \\
& - E_c - E_b - E_a \\
+ \{ \;& -[K_{ab} + 2S_{ab}(\Delta_{ab} + A_{ab} + B_{ab}) - S_{ab}^2(E_a + B_{aa} + E_b + A_{bb} + J_{ab})] \\
& + \tfrac{1}{2}[K_{ac} + 2S_{ac}(\Delta_{ac} + A_{ac} + C_{ac}) - S_{ac}^2(E_a + C_{aa} + E_c + A_{cc} + J_{ac})] \\
& + \tfrac{1}{2}[K_{bc} + 2S_{bc}(\Delta_{bc} + B_{bc} + C_{bc}) - S_{bc}^2(E_b + C_{bb} + E_c + B_{cc} + J_{bc})] \\
& - S_{ab}S_{ac}S_{bc}(-E_c - E_b - E_a) \\
& - S_{bc}S_{ac}(\Delta_{ab} + A_{ab} + B_{ab}) \\
& - S_{ab}S_{bc}(\Delta_{ac} + A_{ac} + C_{ac}) \\
& - S_{ab}S_{ac}(\Delta_{bc} + B_{bc} + C_{bc}) \\
& - 2S_{ab}[C_{ab} + (ab|cc)] + S_{ac}[B_{ac} + (ac|bb)] + S_{bc}[A_{bc} + (bc|aa)] \\
& - (\tfrac{1}{2}S_{bc}^2 + \tfrac{1}{2}S_{ac}^2)J_{ab} + (S_{ab}^2 - \tfrac{1}{2}S_{bc}^2)J_{ac} + (S_{ab}^2 - \tfrac{1}{2}S_{ac}^2)J_{bc} \\
& + S_{ab}^2(C_{aa} + C_{bb}) - \tfrac{1}{2}S_{ac}^2(B_{aa} + B_{cc}) - \tfrac{1}{2}S_{bc}^2(A_{bb} + A_{cc}) \\
& - S_{bc}S_{ac}C_{ab} - S_{ab}S_{bc}B_{ac} - S_{ab}S_{ac}A_{bc} \\
& - S_{ab}(ca|cb) - S_{ac}(ba|bc) - S_{bc}(ab|ac) \\
& + S_{ab}S_{ac}S_{bc}(J_{ab} + J_{ac} + J_{bc}) \\
& + S_{ab}S_{ac}S_{bc}(C_{aa} + C_{bb} + B_{aa} + B_{cc} + A_{bb} + A_{cc}) \qquad \}/ \\
\{ \;& 1 - S_{ab}^2 + \tfrac{1}{2}S_{ac}^2 + \tfrac{1}{2}S_{bc}^2 - S_{ab}S_{bc}S_{ac} \qquad \}
\end{aligned}
\tag{1.80}
$$

The cross energy $\langle\Psi_1|\mathcal{H}|\Psi_2\rangle$ is required for the general case in the next section and is

$$\begin{aligned}\mathcal{H}_{12} = \tfrac{\sqrt{3}}{2}\{\ & -[K_{ac} + 2S_{ac}(\Delta_{ac} + A_{ac} + C_{ac}) - S_{ac}^2(E_a + C_{aa} + E_c + A_{cc} + J_{ac})] \\ & + K_{bc} + 2S_{bc}(\Delta_{bc} + B_{bc} + C_{bc}) - S_{bc}^2(E_b + C_{bb} + E_c + B_{cc} + J_{bc}) \\ & - (S_{ac}^2 - S_{bc}^2)[E_a + B_{aa} + E_b + A_{bb} + J_{ab} + \tfrac{1}{r_{ab}}] \\ & - (S_{ac}^2 - S_{bc}^2)[E_a + C_{aa} + E_c + A_{cc} + J_{ac} + \tfrac{1}{r_{ac}}] \\ & - (S_{ac}^2 - S_{bc}^2)[E_b + C_{bb} + E_c + B_{cc} + J_{bc} + \tfrac{1}{r_{bc}}] \\ & - (S_{ac}^2 - S_{bc}^2)(-E_c - E_b - E_a) \\ & - 2S_{ac}[B_{ac} - (ac|bb)] + 2S_{bc}[A_{bc} + (bc|aa)] \\ & + (S_{ac}^2 - S_{bc}^2)J_{ab} - S_{bc}^2 J_{ac} + S_{ac}^2 J_{bc} \\ & + S_{ac}^2(B_{aa} + B_{cc}) - S_{bc}^2(A_{bb} + A_{cc}) \qquad \}/ \\ \{\ & [1 + S_{ab}^2 - \tfrac{1}{2}S_{bc}^2 - \tfrac{1}{2}S_{ac}^2 - S_{ab}S_{bc}S_{ab}]^{\frac{1}{2}} \\ & [1 - S_{ab}^2 + \tfrac{1}{2}S_{bc}^2 + \tfrac{1}{2}S_{ac}^2 - S_{ab}S_{bc}S_{ab}]^{\frac{1}{2}} \qquad \} \end{aligned} \tag{1.81}$$

A.IV. H_3, General Case

Before writing an energy expression for a general spin coupling, one must remember that the wavefunctions Ψ_1 and Ψ_2 are not orthogonal. It is useful to define the function:

$$\overline{O}(S_{12}) = 2\left(\frac{1-\sqrt{1-S_{12}^2}}{S_{12}^2}\right) = \frac{2}{1+\sqrt{1-S_{12}^2}}, \tag{1.82}$$

which goes to one as S_{12} goes to zero and has the property

$$1 + \tfrac{1}{4}S_{12}^2\overline{O}^2 = \overline{O}. \tag{1.83}$$

Symmetrically orthogonalizing Ψ_1 and Ψ_2, we obtain

$$\overline{\Psi_1} = \overline{N}^{-1}(\Psi_1 - \tfrac{1}{2}S_{12}\overline{O}\Psi_2), \tag{1.84}$$

$$\overline{\Psi_2} = \overline{N}^{-1}(\Psi_2 - \tfrac{1}{2}S_{12}\overline{O}\Psi_1), \tag{1.85}$$

where

$$\overline{N}^2 = \overline{O}(1 - S_{12}^2). \tag{1.86}$$

Using an "overbar" to denote quantities related to these modified wavefunctions, we have

$$\begin{aligned}
\overline{\mathcal{H}_{11}} &= \overline{N}^{-2}[\mathcal{H}_{11} + (1-\overline{O})\mathcal{H}_{22} - S_{12}\overline{O}\mathcal{H}_{12}] \\
&= \overline{N}^{-2}[\mathcal{H}_{22} + \mathcal{H}_{11} - \overline{O}(\mathcal{H}_{22} - S_{12}\mathcal{H}_{12})], \qquad (1.87) \\
\overline{\mathcal{H}_{22}} &= \overline{N}^{-2}[\mathcal{H}_{22} + (1-\overline{O})\mathcal{H}_{11} - S_{12}\overline{O}\mathcal{H}_{12}] \\
&= \overline{N}^{-2}[\mathcal{H}_{22} + \mathcal{H}_{11} - \overline{O}(\mathcal{H}_{11} - S_{12}\mathcal{H}_{12})], \qquad (1.88) \\
\overline{\mathcal{H}_{12}} &= (1-S_{12}^2)^{-1}[\mathcal{H}_{12} - \tfrac{1}{2}S_{12}(\mathcal{H}_{11} + \mathcal{H}_{22})]. \qquad (1.89)
\end{aligned}$$

Finally, the energy for a general spin coupling

$$\chi = \chi_1 \sin\tfrac{1}{2}\gamma + \chi_2 \cos\tfrac{1}{2}\gamma \qquad (1.90)$$

is

$$\begin{aligned}
E &= \sin^2\tfrac{1}{2}\gamma\overline{\mathcal{H}_{11}} + \cos^2\tfrac{1}{2}\gamma\overline{\mathcal{H}_{22}} + 2\sin\tfrac{1}{2}\gamma\cos\tfrac{1}{2}\gamma\overline{\mathcal{H}_{12}} \\
&= \tfrac{1}{2}(1-\cos\gamma)\overline{\mathcal{H}_{11}} + \tfrac{1}{2}(1+\cos\gamma)\overline{\mathcal{H}_{22}} + \sin\gamma\overline{\mathcal{H}_{12}}, \qquad (1.91)
\end{aligned}$$

where, as with the case of H_2, the individual atomic energies are subtracted so that the zero of energy is three dissociated atoms.

For any functional form of the $\overline{\mathcal{H}_{ij}}$'s, the optimal spin coupling is found by setting the first derivative of E_L with respect to γ to zero. This yields

$$\tan\gamma = \frac{\overline{\mathcal{H}_{12}}}{\frac{1}{2}(\overline{\mathcal{H}_{22}} - \overline{\mathcal{H}_{11}})}. \qquad (1.92)$$

Use the relation $\cos\gamma = \pm(1+\tan^2\gamma)^{-\frac{1}{2}}$ and define

$$\overline{X} \equiv \sqrt{\tfrac{1}{4}(\overline{\mathcal{H}_{22}} - \overline{\mathcal{H}_{11}})^2 + \overline{\mathcal{H}_{12}}^2} \qquad (1.93)$$

to yield

$$\cos\gamma = -\tfrac{1}{2}(\overline{\mathcal{H}_{22}} - \overline{\mathcal{H}_{11}})/\overline{X}, \qquad (1.94)$$

where we have chosen the sign by noting that $\cos\gamma = 1$ when $\mathcal{H}_{12} = 0$ and $\mathcal{H}_{11} > \mathcal{H}_{22}$. This immediately leads to

$$\sin\gamma = -\overline{\mathcal{H}_{12}}/\overline{X}. \tag{1.95}$$

Substituting for γ, we get the energy for the optimal spin coupling:

$$E = \tfrac{1}{2}(\overline{\mathcal{H}_{22}} + \overline{\mathcal{H}_{11}}) - \overline{X}. \tag{1.96}$$

B. "Approximate" Valence Bond Descriptions

There are three main sets of approximations which are made in transforming the separate two-body and three-body equations of the previous appendix into useful equations. Common to all is the substitution of the two-body expressions of Subsection A.II into the $\mathcal{H}_{ij}$ terms of Subsection A.III followed by the derivation of a general energy expression per Subsection A.IV. The three sets of approximations are:

Subsection B.I: Retain overlap terms through third order.

Subsection B.II: Retain overlap terms through second order.

Subsection B.III: Neglect all overlap terms.

The last approximation is that of the London Equation. The generally good fit of the London Potential may lie in the fact that it encompasses a consistent approximation scheme. All terms except the classical and exchange energies can be considered to be a given order in overlap and thus set to zero under this approximation. We must assign an "overlap order" to all other terms for the first two approximation levels. This is done by the order of e^{-w} per the one- and two-center integrals of Subsection A.II and similarly for the three-center integrals of Subsection A.III. This

order is listed for the various terms of $\mathcal{H}_{11}$ following its equation in Subsection A.III. The approximation order must also be consistent with that used in finding the H_2 classical and exchange energies from the singlet and triplet energies. Thus, these definitions are different in the London case ($S^2 = 0$) than for the other two cases (S^2 retained). A general energy expression (the Generalized London Equation) is derived for each set of approximations.

B.I. Retain Overlap Through Third Order

First consider the H_2 classical and exchange energies. Retaining the S^2 terms in the classical energy and neglecting the S^4 term in the exchange energy, we have

$$E_{ab}^{cl} = \tfrac{1}{2}[(1 - S_{ab}^2)E_{ab}^T + (1 + S_{ab}^2)E_{ab}^S], \tag{1.97}$$

$$E_{ab}^{x} = \tfrac{1}{2}(E_{ab}^T - E_{ab}^S). \tag{1.98}$$

Next consider the H_3 $\mathcal{H}_{ij}$ energies of Subsection A.III. Of course the overlap squared and cubed terms of the normalization constants are retained. The terms fourth and fifth order in overlap such as $S_{bc}^2 J_{ab}$ and those that follow are set to zero. After substituting for the H_2 classical and exchange energies in terms of their defining integrals, we are left with just two types of third order terms. The first appears wherever there is an exchange term and is defined as

$$F_{ab} \equiv -2S_{ab}[C_{ab} + (ab|cc)]. \tag{1.99}$$

Since both of the terms in square brackets go as $S_{ac}S_{bc}$, this term overall contains each of the three overlaps. The second appears in both $\mathcal{H}_{11}$ and $\mathcal{H}_{22}$ and is

$$\begin{aligned} A \equiv\ & -S_{ab}S_{ac}S_{bc}(-E_c - E_b - E_a) \\ & - S_{bc}S_{ac}(\Delta_{ab} + A_{ab} + B_{ab}) \end{aligned} \tag{1.100}$$

$$- S_{ab}S_{bc}(\Delta_{ac} + A_{ac} + C_{ac})$$

$$- S_{ab}S_{ac}(\Delta_{bc} + B_{bc} + C_{bc}).$$

Each of these terms also goes as $S_{ab}S_{ac}S_{bc}$. Thus all "S^3" terms involve one S from each pair.

To facilitate the derivation of a general energy expression, we group terms using the following definitions:

$$C \equiv E_{ab}^{cl} + E_{ac}^{cl} + E_{bc}^{cl} - E_c - E_b - E_a, \tag{1.101}$$

$$E_{ab}^{y} \equiv E_{ab}^{x} + F_{ab}, \tag{1.102}$$

$$Y \equiv E_{ab}^{y} - \tfrac{1}{2}E_{ac}^{y} - \tfrac{1}{2}E_{bc}^{y}, \tag{1.103}$$

$$Z \equiv \tfrac{\sqrt{3}}{2}(E_{ac}^{y} - E_{bc}^{y}), \tag{1.104}$$

$$\mathcal{S} \equiv S_{ab}^2 - \tfrac{1}{2}S_{ac}^2 - \tfrac{1}{2}S_{bc}^2, \tag{1.105}$$

$$\mathcal{T} \equiv -\tfrac{\sqrt{3}}{2}(S_{ac}^2 - S_{bc}^2), \tag{1.106}$$

$$\mathcal{O} \equiv S_{ab}S_{ac}S_{bc}. \tag{1.107}$$

This allows us to write the $\mathcal{H}_{ij}$ and normalization terms as

$$\mathcal{H}_{11} = C + [-Y - A]/N_1^2, \tag{1.108}$$

$$\mathcal{H}_{22} = C + [Y - A]/N_2^2, \tag{1.109}$$

$$\mathcal{H}_{12} = S_{12}C + Z/N_1N_2, \tag{1.110}$$

$$S_{12} = \mathcal{T}/N_1N_2, \tag{1.111}$$

$$N_1^2 = 1 + \mathcal{S} - \mathcal{O}, \tag{1.112}$$

$$N_2^2 = 1 - \mathcal{S} - \mathcal{O}, \tag{1.113}$$

by inspection of the equations in Subsection A.III. After correction for wavefunction overlap per Subsection A.IV, we have

$$\overline{N^2\mathcal{H}_{11}} = [2 - \overline{O}(1 + S_{12}^2)]C + \{-Y[\overline{O} - \mathcal{S}(2 - \overline{O}) - \mathcal{O}\overline{O}] \tag{1.114}$$

$$- A[(2-\overline{O}) - \mathcal{S}\overline{O} - \mathcal{O}(2-\overline{O})] - \mathcal{T}Z\overline{O}\}/N_1^2N_2^2, \quad (1.115)$$

$$\overline{N^2\mathcal{H}_{22}} = [2-\overline{O}(1+S_{12}^2)]C + \{Y[\overline{O} + \mathcal{S}(2-\overline{O}) - \mathcal{O}\overline{O}] \quad (1.116)$$

$$- A[(2-\overline{O}) + \mathcal{S}\overline{O} - \mathcal{O}(2-\overline{O})] - \mathcal{T}Z\overline{O}\}/N_1^2N_2^2, \quad (1.117)$$

$$\overline{\mathcal{H}_{12}} = \{Z/N_1N_2 - S_{12}[\mathcal{S}Y - (1-\mathcal{O})A]/N_1^2N_2^2\}/(1-S_{12}^2). \quad (1.118)$$

Up to now we have made no approximations besides not including the fourth and fifth order integrals of the $\mathcal{H}_{ij}$. We now neglect terms fourth and higher order in overlap by setting:

$$S_{ab}^4 = 0, \quad (1.119)$$

$$\overline{O} = 1, \quad (1.120)$$

$$\overline{N} = 1, \quad (1.121)$$

$$S_{12}^2 = 0, \quad (1.122)$$

$$N_1N_2 = 1-\mathcal{O}, \quad (1.123)$$

$$N_1^2N_2^2 = (1-\mathcal{O})^2. \quad (1.124)$$

Note that there will be some cancellation of $1-\mathcal{O}$ from both the numerator and denominator so that it is best to approximate the exact $N_1^2N_2^2 = (1-\mathcal{O})^2 - \mathcal{S}^2$ as above versus $1-2\mathcal{O}$. Thus, the energy terms corrected for wavefunction overlap and including through third order in orbital overlap are

$$\overline{\mathcal{H}_{11}} = C + [(-Y-A)(1-\mathcal{S}-\mathcal{O}) - \mathcal{T}Z]/[1-\mathcal{O}]^2, \quad (1.125)$$

$$\overline{\mathcal{H}_{22}} = C + [(Y-A)(1+\mathcal{S}-\mathcal{O}) - \mathcal{T}Z]/[1-\mathcal{O}]^2, \quad (1.126)$$

$$\overline{\mathcal{H}_{12}} = Z/[1-\mathcal{O}], \quad (1.127)$$

$$\tfrac{1}{2}(\overline{\mathcal{H}_{11}} + \overline{\mathcal{H}_{22}}) = C + [\mathcal{S}Y - \mathcal{T}Z - (1-\mathcal{O})A]/[1-\mathcal{O}]^2, \quad (1.128)$$

$$\tfrac{1}{2}(\overline{\mathcal{H}_{22}} - \overline{\mathcal{H}_{11}}) = Y/[1-\mathcal{O}], \quad (1.129)$$

$$\overline{X}^2 = [Y^2+Z^2]/[1-\mathcal{O}]^2. \quad (1.130)$$

This leads to a general energy of

$$\begin{aligned} E &= E_{ab}^{cl} + E_{ac}^{cl} + E_{bc}^{cl} - E_c - E_b - E_a \\ &+ [(S_{ab}^2 - \tfrac{1}{2}S_{ac}^2 - \tfrac{1}{2}S_{bc}^2)E_{ab}^y \\ &+ (-\tfrac{1}{2}S_{ab}^2 + S_{ac}^2 - \tfrac{1}{2}S_{bc}^2)E_{ac}^y \\ &+ (-\tfrac{1}{2}S_{ab}^2 - \tfrac{1}{2}S_{ac}^2 + S_{bc}^2)E_{ac}^y]/[1 - S_{ab}S_{ac}S_{bc}]^2 \\ &- A/[1 - S_{ab}S_{ac}S_{bc}] \\ &- [E_{ab}^{y\,2} + E_{ac}^{y\,2} + E_{bc}^{y\,2} - E_{ab}^y E_{ac}^y - E_{ac}^y E_{bc}^y - E_{ab}^y E_{bc}^y]^{\frac{1}{2}}/[1 - S_{ab}S_{ac}S_{bc}]. \end{aligned} \tag{1.131}$$

B.II. Retain Overlap Through Second Order

We now consider also setting cubic terms in overlap to zero. The H_2 classical and exchange energies are the same as in the previous subsection. Next consider the $\overline{\mathcal{H}_{ij}}$ terms. Removing $\mathcal{O}$, it being an S^3 term, reduces $\frac{1}{2}(\overline{\mathcal{H}_{22}} - \overline{\mathcal{H}_{11}})$ and $\overline{X}$ to their London Equation forms. This leads to an important conclusion, that in this model, the spin coupling is exactly the same as the London spin coupling. The difference between the London and new $\frac{1}{2}(\overline{\mathcal{H}_{11}} + \overline{\mathcal{H}_{22}})$ value may be regarded as an overlap correction to the classical energy. The general energy equation is easily obtained from above by removing the $S_{ab}S_{ac}S_{bc}$ and A terms and replacing E_{ab}^y with E_{ab}^x. This yields

$$E = E_{ab}^{cl} + E_{ac}^{cl} + E_{bc}^{cl} - E_c - E_b - E_a \tag{1.132}$$

$$+ (S_{ab}^2 - \tfrac{1}{2}S_{ac}^2 - \tfrac{1}{2}S_{bc}^2)E_{ab}^x \tag{1.133}$$

$$+ (-\tfrac{1}{2}S_{ab}^2 + S_{ac}^2 - \tfrac{1}{2}S_{bc}^2)E_{ac}^x \tag{1.134}$$

$$+ (-\tfrac{1}{2}S_{ab}^2 - \tfrac{1}{2}S_{ac}^2 + S_{bc}^2)E_{ac}^x \tag{1.135}$$

$$- [E_{ab}^{x\,2} + E_{ac}^{x\,2} + E_{bc}^{x\,2} - E_{ab}^x E_{ac}^x - E_{ac}^x E_{bc}^x - E_{ab}^x E_{bc}^x]^{\frac{1}{2}}. \tag{1.136}$$

It is straight forward to substitute for the actual singlet and triplet energies at this point to obtain the energy expression in terms of the London Equation energy E_L:

$$E = E_L - \tfrac{1}{4}(S_{ac}^2 + S_{bc}^2)(E_{ab}^T - E_{ab}^S) - \tfrac{1}{4}(S_{ab}^2 + S_{bc}^2)(E_{ac}^T - E_{ac}^S) - \tfrac{1}{4}(S_{ab}^2 + S_{ac}^2)(E_{bc}^T - E_{bc}^S). \quad (1.137)$$

Note that the lowest order overlap corrections are only on other pairs, e.g., there is no S_{ab}^2 term multiplying $E_{ab}^{T,S}$. The importance of correcting the $\mathcal{H}_{ij}$ for wavefunction overlap to obtain the $\overline{\mathcal{H}_{ij}}$ is illustrated by this not being true if one uses the $\mathcal{H}_{ij}$ directly in determining the general potential.

B.III. Neglect All Overlap and 3-Body Terms

The terms resulting from setting all overlap terms to zero are easily written down based on the previous subsection. With wavefunction overlap now zero, i.e., $S_{12} = 0$, we have

$$\overline{\mathcal{H}_{11}} = \mathcal{H}_{11} = E_{ab}^S + \tfrac{3}{4}E_{ac}^T + \tfrac{1}{4}E_{ac}^S + \tfrac{3}{4}E_{bc}^T + \tfrac{1}{4}E_{bc}^S - E_c - E_b - E_a, \quad (1.138)$$

$$\overline{\mathcal{H}_{22}} = \mathcal{H}_{22} = E_{ab}^T + \tfrac{1}{4}E_{ac}^T + \tfrac{3}{4}E_{ac}^S + \tfrac{1}{4}E_{bc}^T + \tfrac{3}{4}E_{bc}^S - E_c - E_b - E_a, \quad (1.139)$$

$$\overline{\mathcal{H}_{12}} = \mathcal{H}_{12} = \tfrac{\sqrt{3}}{4}(E_{ac}^T - E_{ac}^S - E_{bc}^T + E_{bc}^S). \quad (1.140)$$

With the substitutions of pair 3 for ab, pair 2 for ac, and pair 1 for bc, we have derived the equations for $\mathcal{H}_{ij}$ used in deriving the London Potential.

Using the equations at the end of Appendix A, we find

$$\tan\gamma = \frac{\sqrt{3}(E_2^x - E_1^x)}{2E_3^x - E_2^x - E_1^x}, \quad (1.141)$$

$$\cos\gamma = \frac{-(E_3^x - \frac{1}{2}E_2^x - \frac{1}{2}E_1^x)}{(E_3^{x2} + E_2^{x2} + E_1^{x2} - E_3^xE_2^x - E_3^xE_1^x - E_2^xE_1^x)^{\frac{1}{2}}}, \quad (1.142)$$

$$\sin\gamma = \frac{-\sqrt{3}(\frac{1}{2}E_2^x - \frac{1}{2}E_1^x)}{(E_3^{x2} + E_2^{x2} + E_1^{x2} - E_3^xE_2^x - E_3^xE_1^x - E_2^xE_1^x)^{\frac{1}{2}}}, \quad (1.143)$$

γ	$E_L + E_c + E_b + E_a$
0	$E_3^T + \frac{1}{4}E_2^T + \frac{3}{4}E_2^S + \frac{1}{4}E_1^T + \frac{3}{4}E_1^S$
$\frac{\pi}{3}$	$E_1^S + \frac{3}{4}E_3^T + \frac{1}{4}E_3^S + \frac{3}{4}E_2^T + \frac{1}{4}E_2^S$
$\frac{2\pi}{3}$	$E_2^T + \frac{1}{4}E_1^T + \frac{3}{4}E_1^S + \frac{1}{4}E_3^T + \frac{3}{4}E_3^S$
π	$E_3^S + \frac{3}{4}E_2^T + \frac{1}{4}E_2^S + \frac{3}{4}E_1^T + \frac{1}{4}E_1^S$
$-\frac{2\pi}{3}$	$E_1^T + \frac{1}{4}E_3^T + \frac{3}{4}E_3^S + \frac{1}{4}E_2^T + \frac{3}{4}E_2^S$
$-\frac{\pi}{3}$	$E_2^S + \frac{3}{4}E_1^T + \frac{1}{4}E_1^S + \frac{3}{4}E_3^T + \frac{1}{4}E_3^S$

Table 1.6. Special Cases of Energy for Various γ's.

where $E_i^x = \frac{1}{2}(E_i^T - E_i^S)$ is independent of overlap. All of the special cases of one pair being purely singlet or triplet occur when γ is a multiple of $\frac{\pi}{3}$. These are shown in Table 1.6. Thus, we can consider γ as a reaction coordinate that would range from $-\frac{\pi}{3}$ to 0 to $\frac{\pi}{3}$ in the reaction

$$H_a \quad + \quad H_c - H_b \rightleftharpoons H_a \text{---} H_c \text{---} H_b \rightleftharpoons H_a - H_c \quad + \quad H_b.$$

Substituting the above equations for $\mathcal{H}_{ij}$, $\cos\gamma$, and $\sin\gamma$ in Equation (1.96) leads to the London Equation per Equation (1.12). For calculation of the functions $f^T(\gamma_i)$ and $f^S(\gamma_i)$ of Equation (1.15), it is prudent to find them without prejudice with respect to a given pair using the equation:

$$\cos\gamma_i = \frac{-(E_i^x - \frac{1}{2}E_j^x - \frac{1}{2}E_k^x)}{(E_3^{x2} + E_2^{x2} + E_1^{x2} - E_3^x E_2^x - E_3^x E_1^x - E_2^x E_1^x)^{\frac{1}{2}}}, \tag{1.144}$$

where i, j, and k are all distinct.

One final note on dissociation limits should be made. Remember that $E_a \rightarrow E_a^0$ only as both atoms b and c are far removed from atom a. If all three atoms dissociate, the pairwise energies become sums of atomic energies, e.g., $E_{ab}^{T,S} = E_a^0 + E_b^0$, and all atomic energies reach their isolated value, e.g., $E_a = E_a^0$. Thus the total energy is just $E = E_c^0 + E_b^0 + E_a^0$. As atom c, say, becomes separated from pair ab, the energy of pair ac does *not* reach the isolated atom limit, but instead $E_{ac}^{T,S} = E_a + E_c^0$ where

$E_a > E_a^0$ (remember that both are negative quantities). Thus, the total energy becomes:

$$E = E_{ab}^S + (E_a + E_c^0) + (E_b + E_c^0) - (E_c^0 + E_b + E_a) = E_{ab}^S, \tag{1.145}$$

the correct one pair limit.

That's all, folks.

References

[1] S. Sato, Bull. Chem. Soc. Jap. **28**, 450 (1955); J. Chem. Phys. **23**, 592 (1955); *ibid.* **23**, 2465 (1955).

[2] J.-C. Nieh and J. J. Valenti, Phys. Rev. Lett. **60**, 519 (1988); J. Chem. Phys. **92**, 1083 (1990).

[3] R. E. Continetti, B. A. Balko, Y. T. Lee, J. Chem. Phys. **93**, 5719 (1990).

[4] M .Mladenovic, M. Zhao, D. G. Truhlar, D. W. Schwenke, Y. Sun, and D. J. Kouri, Chem. Phys. Lett. **146**, 358 (1988); J. Phys. Chem. **92**, 7035 (1988).

[5] J. Z. H. Zhang and W. H. Miller, Chem. Phys. Lett. **153**, 465 (1988).

[6] D. E. Manolopoulos and R. E. Wyatt, Chem. Phys. Lett. **159**, 123 (1989).

[7] C. W. Bauschlicher, Jr., S. R. Langhoff, and H. Partridge, Chem. Phys. Lett. **170**, 345 (1990).

[8] E. Wurzberg and P. L. Houston, J. Chem. Phys. **72**, 4811 (1980); R. F. Heidner, J. F. Bott, C. E. Gardner, and J. E. Melzer, J. Chem. Phys. **72**, 4815 (1980); D. E. Neumark, A. M. Wodtke, G. N. Robinson, C. C. Hayden, and Y. T. Lee, Phys. Rev. Lett. **53**, 226 (1984); D. E. Neumark, A. M. Wodtke, G. N. Robinson, C. C. Hayden, and Y. T. Lee, J. Chem. Phys. **82**, 3045 (1985); P. S. Stevens, W. H. Brune, and J. G. Anderson, J. Phys. Chem. **93**, 4068 (1989); A. Weaver, R. B. Metz, S. E. Bradforth, and D. M. Neumark, J. Chem. Phys. **93**, 5352 (1990).

[9] H. F. Schaefer III, J. Phys. Chem. **89**, 5336 (1985); R. Steckler and D. G. Truhlar, J. Chem. Phys. **82**, 5499 (1985); D. W. Schwenke, R. Steckler, F. B.

Brown, and D. G. Truhlar, J. Chem. Phys. **84**, 5706 (1986); D. W. Schwenke, R. Steckler, F. B. Brown, and D. G. Truhlar, J. Chem. Phys. **86**, 2443 (1987); C. W. Bauschlicher, Jr., S. P. Walch, S. R. Langhoff, P. R. Taylor, and R. L. Jaffe, J. Chem. Phys. **88**, 1743 (1988); J. Z. H. Zhang and W. H. Miller, J. Chem. Phys. **88**, 4549 (1988); G. E. Scuseria and H. F. Schaefer III, J. Chem. Phys. **88**, 7024 (1988); C. W. Bauschlicher, Jr., S. R. Langhoff, T. J. Lee, and P. R. Taylor, J. Chem. Phys. **90**, 4296 (1989).

[10] For a history of H_3 potential energy surfaces, see D. G. Truhlar and R. E. Wyatt, Adv. Chem. Phys. **36**, 141 (1977).

[11] For a history of H_3 dynamics modeling, see D. G. Truhlar and R. E. Wyatt, Ann. Rev. Phys. Chem. **27**, 1 (1976).

[12] For a review of earlier work on exchange reactions including halogens, see C. A. Parr and D. G. Truhlar, J. Phys. Chem. **75**, 1844 (1971).

[13] E. Wigner, Ztschr. f. Phys. **40**, 492 (1926).

[14] W. Heisenberg, Ztschr. f. Phys. **41**, 239 (1927).

[15] F. London, "Probleme der Modernen Physik-Sommerfeld Festschrift" (S. Herzel, Leipzig, 1928).

[16] F. London, Ztschr. f. Elektrochem. **35**, 552 (1929).

[17] For an early review including complete derivation of the London Equation, see J. C. Slater, Phys. Rev. **38**, 1109 (1931).

[18] For a derivation of the London Equation as a special case of the four-electron problem, see S. Glasstone, K. J. Laidler, and H. Eyring, "The Theory of Rate Processes" (McGraw-Hill Book Company, Inc., New York, 1941).

[19] W. Heitler and F. London, Ztschr. f. Phys. **44**, 244 (1927).

[20] Y. Sugiura, Ztschr. f. Phys. **45**, 484 (1927).

[21] H. Eyring and M. Polanyi, Ztschr. f. phys. Chem. **B12**, 279 (1931).

[22] M. Polanyi, "Atomic Reactions" (Williams & Northgate Ltd., London, 1932).

[23] P. M. Morse, Phys. Rev. **34**, 57 (1929).

[24] H. Eyring, Naturwiss. **18**, 914 (1930). See also the preceding article, H. Eyring and M. Polanyi, *ibid.* **18**, 913 (1930).

[25] H. Eyring, J. Am. Chem. Soc. **53**, 2537 (1931).

[26] R. E. Weston, Jr., J. Chem. Phys. **31**, 892 (1959).

[27] J. K. Cashion and D. R. Hershbach, J. Chem. Phys. **40**, 2358 (1964).

[28] B. Liu, J. Chem. Phys. **58**, 1925 (1973).

[29] P. Siegbahn and B. Liu, J. Chem. Phys. **68**, 2457 (1978).

[30] D. G. Truhlar and C. J. Horowitz, J. Chem. Phys. **68**, 2466 (1978); *ibid.* **71**, 1514(E) (1979).

[31] B. Liu, J. Chem. Phys. **80**, 581 (1984).

[32] M. R. A. Blomberg and B. Liu, J. Chem. Phys. **82**, 1050 (1985).

[33] D. M. Ceperley and B. J. Adler, J. Chem. Phys. **81**, 5833 (1984).

[34] R. N. Barnett, P. J. Reynolds, and W. A. Lester, Jr., J. Chem. Phys. **82**, 2700 (1985).

[35] A. J. C. Varandas, F. B. Brown, C. A. Mead, and D. G. Truhlar, J. Chem. Phys. **86**, 6258 (1987).

[36] W. Kolos and L. Wolniewicz, J. Chem. Phys. **41** 3663 (1964); *ibid.* **43**, 2429 (1965); *ibid.* **49**, 404 (1968); Chem. Phys. Lett. **24**, 457 (1974); J. Molec. Spectosc. **54**, 303 (1975).

[37] F. W. Bobrowicz adn W. A. Goddard III in "Methods of Electronic Structure Theory," edited by H. F. Schaefer III (Plenum Press, New York, 1977).

[38] C. B. Musgrave, R. E. Donnelly, and W. A. Goddard III, unpublished results.

[39] C. B. Musgrave, J. K. Perry, R. C. Merkle, and W. A. Goddard III, submitted to Nanotech.

[40] J. C. Slater, "Quantum Theory of Molecules and Solids" Vol. 1 (McGraw-Hill Book Company, Inc., New York, 1963). Page 53 contains two "typos": add e^{-w} after $\alpha^2(1 + w - \frac{1}{3}w^2)$ on $-\alpha^2(K + S)$ line and replace $\frac{1}{3}w^2$ with $\frac{1}{3}w^3$ on αK line. Also note that Rydberg (not hartree) energy units are used.

Chapter 2

Optimal Meshes for Molecular Properties and Visualization

I. Introduction

While the previous chapter developed a model for potential energy surfaces which can be used to simulate reactions at polymer and crystalline surfaces, there are other areas of modeling large systems which need improvement. In order to determine surface related phenomena such as solvation energy for polymers, one requires the surface area. Current techniques are often computationally intensive and do not aid in distinguishing between, say, hydrophobic and hydrophilic areas or internal and external surface areas. Even with high speed graphic workstations, the generic rendering techniques used with molecular displays are not fast enough for "real-time" rotation of large molecules. Another seemingly unrelated field is the numerical integration of functions on spherical surfaces. This is required in many fields including chemistry. All of these applications share the need of distributing points on a sphere. Often either this is done without much consideration or, especially for the application of numerical integration, special grids of points are developed with unique properties but which have little application elsewhere. Here we introduce a general point mesh design which allows the generation of grids of unlimited size, which is able to adapt to several important point symmetries, and which has built-in

degrees of freedom which can be optimized for any of the above applications.

We make the distinction between a grid as merely a collection of points over the surface or volume under consideration and a mesh which also includes the concept of connectivity. A grid (with the possible addition of a point weighting system which then will be called a quadrature scheme) is sufficient for numerical integration while a mesh is necessary for rendering. The criteria for evaluation a grid or mesh varies depending on the application. For the purpose of the graphical rendering of a molecular display, the best mesh is the one that most quickly produces the best looking image. For the purpose of function evaluation, e.g., finding the surface area of a molecule, a grid with an even distribution of points over the surface is desirable. For the purpose of numerical integration, the exact integration of some number of basis functions may be desirable. In Section II we present a general mesh which is adaptable to each of these requirements and discuss spatial criteria for the evaluation of point distribution. In Section III we tune this mesh toward applications involving spherical harmonics and consider the particular application of integration on the unit sphere. We then discuss several applications involving molecular surfaces including visualization, accessible surfaces, and surface area.

II. Spherical Meshes

When a spherical grid of points is desired, often a lattice of longitude and latitude lines is implemented and no further thought is given to it. Actually, this turns out to be a poor choice as far as both rendering efficiency and point distribution. Our goal is to find a mesh of points for the unit sphere that is appropriate for all of the applications outlined in Section I. Since our requirements vary according to the application, we actually desire a set of meshes containing an orderly increase in

the number of points, from which we can choose the desired size. We use the term "sphere level" to distinguish among this set with a larger sphere level indicating a larger number of points. Instead of having a fixed set of positions for each mesh level, we desire a certain number of degrees of freedom so that the mesh can be fine tuned according to the application. Generally, this needs to be done only once for each type of application. We enforce two restrictions that guide our design of a mesh. First, we require that the connectivity of the mesh be such that connecting neighbor points leads to a triangular mesh. This leads to greater efficiency in computer algorithms when hardware support of triangular meshes means fewer point transformations. Second, we require that we can enforce a high degree of symmetry such as octahedral. This reduces the number of degrees of freedom that need be optimized for a given class of applications, allows us to adapt the mesh to a given set of basis functions that conform to a particular symmetry, and enforces a certain degree of uniformity. More general desirable properties include an even point distribution, an indexing system easily implemented in rendering applications, and a small set of data required to implement a given level of the mesh. The mesh we present is able to conform to all of these requirements. The rest of this section considers:

Mesh Connectivity. How the number and labeling of points on a sphere are determined and how they connect to form polygons.

Point Assignments. How coordinates are assigned to the points.

Point Distribution Criteria. How the uniformity of the distribution of a mesh of points is evaluated.

II.A. Mesh Connectivity

In distributing points on a sphere, it is convenient to use a three-tuple indexing system that is related to the (x, y, z) axes of a coordinate system. The set of allowed indices (i, j, k) and the set of allowed edges between two such values must be defined in order to define a mesh. Since it is trivial to scale coordinates by a given radius, we assume we are working on a unit sphere for the rest of this section. For a level L sphere of the traditional longitude and latitude lines approach (denoted as method LL form now on), first a preferred direction, z, is chosen and the associated index k is allowed to range from $-L$ to L. For each value of k except $\pm L$, allowed i and j values are those that fulfill $|i| + |j| = L$. This basically means that a circle of $4L$ points is drawn for each non-polar k value. For the poles $k = \pm L$, we let $i = j = 0$. The edges representing longitude lines are formed by connecting all pairs of legal points (i, j, k) and $(i, j, k+1)$ for $k = -L, \ldots, L-1$; the edges representing latitude lines are formed by connecting all pairs of points (i_1, j_1, k) and (i_2, j_2, k) such that $|i_1 - i_2| = 1$ and $|j_1 - j_2| = 1$ for $k = -L+1, \ldots, L-1$. Once the points and edges have been defined, the polygons by default have been defined, and in this case they consist of quadrilaterals between latitude lines k and $k+1$ for $k = -L+1, \ldots, L-2$ and triangles between latitude line $k = -L+1$ and south pole $k = -L$ and latitude line $k = L-1$ and north pole $k = L$. Method LL thus has $8L^2 - 4L + 2$ points, $16L^2 - 4L$ edges, and $8L^2$ polygons of which $8L$ are triangles and the rest are quadrilaterals. It has symmetry $D_{(4L)h}$ (actually, one could construct LL method spheres with symmetry $D_{(N)h}$ using N points per latitude line for any $N > 1$). This method has a higher density of points near the poles and a lower density near the equator. There are various ways of improving the point distribution of this method such as reducing the number of longitude lines used near

(i,j,k) Range			Point Label
$i=0$	$j=0$	$k=-L$	1
$i\geq 0$	$j>0$	$k\leq 0$	$2+2(L-k-1)(L+k)+i$
$i>0$	$j\leq 0$	$k\leq 0$	$2+2(L-k-1)(L+k)+(L+k)-j$
$i\leq 0$	$j<0$	$k\leq 0$	$2+2(L-k-1)(L+k)+2(L+k)-i$
$i<0$	$j\geq 0$	$k\leq 0$	$2+2(L-k-1)(L+k)+3(L+k)+j$
$i\geq 0$	$j>0$	$k>0$	$2+4L^2-2(L-k)(L-k+1)+i$
$i>0$	$j\leq 0$	$k>0$	$2+4L^2-2(L-k)(L-k+1)+(L-k)-j$
$i\leq 0$	$j<0$	$k>0$	$2+4L^2-2(L-k)(L-k+1)+2(L-k)-i$
$i<0$	$j\geq 0$	$k>0$	$2+4L^2-2(L-k)(L-k+1)+3(L-k)+j$
$i=0$	$j=0$	$k=L$	$2+4L^2$

Table 2.1. Single Index Point Labels from (i,j,k) Indices.

the poles of the sphere. Since a triangular mesh is preferable to a quadrilateral mesh due to the triangle mesh support of some graphics hardware, we comment on only one such method that uses latitude lines as the basis for a triangular grid. In fact we consider only one additional mesh which is the basis for various methods of point assignments, and we present that mesh next.

A level L sphere is defined as the collection of triangles made from nearest neighbor points with indices (i,j,k) such that $|i|+|j|+|k|=L$. Edges are formed from any pair of points such that the absolute value of one index is increased by one and the absolute value of a second index is decreased by one, leaving the third index unchanged. There are $2+4L^2$ points comprising $8L^2$ triangles and $12L^2$ edges. Either D_{4h} or O_h symmetry is easily enforced by enforcing relations among points with symmetric (i,j,k) indices. It is useful to reduce the three indices (i,j,k) to a single index labeling each point. A method for doing this is described in Table 2.1. When the polygons are rendered in increasing point order using this method, they are rendered from back to front ($k=-L$ to L), as necessary for transparency.

II.B. Point Assignments

We now consider the task of giving actual coordinate positions to the points of the meshes defined above. We almost had to go out of our way to describe the points and connectivity of method LL without giving the coordinates since they are so intertwined, but we present them now:

$$\begin{aligned} z &= \sin\left(\frac{k\pi}{2L}\right), \\ y &= \cos\left(\frac{k\pi}{2L}\right)\sin\left(\frac{j\pi}{2L}\right), \\ x &= \cos\left(\frac{k\pi}{2L}\right)\sin\left(\frac{i\pi}{2L}\right). \end{aligned} \tag{2.1}$$

Of course other methods of assigning the z coordinate which increase the spacing of latitude lines near the poles are possible, but we instead devote our time to more promising methods. We consider three different spatial methods of assigning coordinates to the triangular mesh described above. The first is the most closely related to method LL in that a preferred direction, z, is chosen and k is used as the basis of latitude lines between which triangles are drawn. This is denoted as method LT from now on and has D_{4h} symmetry. The other two methods are formed under the constraints of octahedral (O_h) symmetry. This helps to ensure uniform point concentration over the sphere and reduces the number of degrees of freedom that need be defined or optimized. Method IJK simply uses the (i, j, k) indices to define the points while method ARC is based on equally spacing points on equidistant arcs of the sphere. Point assignment methods optimized for a given application are discussed in the section dealing with that application since the criteria for an ideal grid varies according to the application.

Perhaps the easiest way to assign coordinates to these points is method LT. Here, k is used to fix the z-coordinate (latitude) and $4(L - |k|)$ equally spaced points are

put on each latitude line. The points on a unit sphere are thus:

$$\begin{aligned} z &= \sin\left(\frac{k\pi}{2L}\right), \\ y &= \cos\left(\frac{k\pi}{2L}\right)\sin\left[\frac{j\pi}{2(L-k)}\right], \\ x &= \cos\left(\frac{k\pi}{2L}\right)\sin\left[\frac{i\pi}{2(L-k)}\right]. \end{aligned} \tag{2.2}$$

Compare the mesh of methods of LL and LT in Figure 2.1, columns (a) and (b), respectively.

Perhaps the easiest way to assign points with O_h symmetry is to let:

$$(i,j,k) = \frac{(i,j,k)}{\sqrt{i^2+j^2+k^2}}. \tag{2.3}$$

The left-hand side of Equation (2.3) is a point index, and the right-hand side is a 3-tuple of coordinate values. The resulting mesh is shown in Figure 2.1, column (c). While this method is simple in design and simple to implement, it leads to a higher point concentration at the x, y, z-axes than at the centers of each octant. It is useful, however, as an alternative to method ARC as a starting point for optimization of our mesh for a particular purpose. Before presenting method ARC, we consider the general features of our mesh with O_h symmetry enforced.

Since the indices (i,j,k) must reflect octahedral symmetry, only the points with $0 \le i \le j \le k$, i.e.,[2]

$$i = 0,1,2,\ldots,\frac{L}{3}, \quad j = i, i+1, i+2, \ldots, \frac{L-i}{2}, \quad k = L-i-j$$

need be defined. Thus the number of independent points is:

$$1+\frac{L}{3}+\frac{L}{2}\left(\frac{L}{6}+1\right)+\frac{L+1}{2}\left(\frac{\frac{L}{3}+1}{2}\right)-\frac{\frac{L}{3}\left(\frac{L}{3}+1\right)}{2}-\frac{\frac{L}{6}\left(\frac{L}{6}+1\right)}{2}-\frac{\left(\frac{\frac{L}{3}+1}{2}\right)\left(\frac{\frac{L}{3}+1}{2}+1\right)}{2}.$$

[2]In this and other equations in this paper involving integer arithmetic only, truncation during division is assumed. Thus certain formulas, e.g., the number of independent points, are not easily reduced.

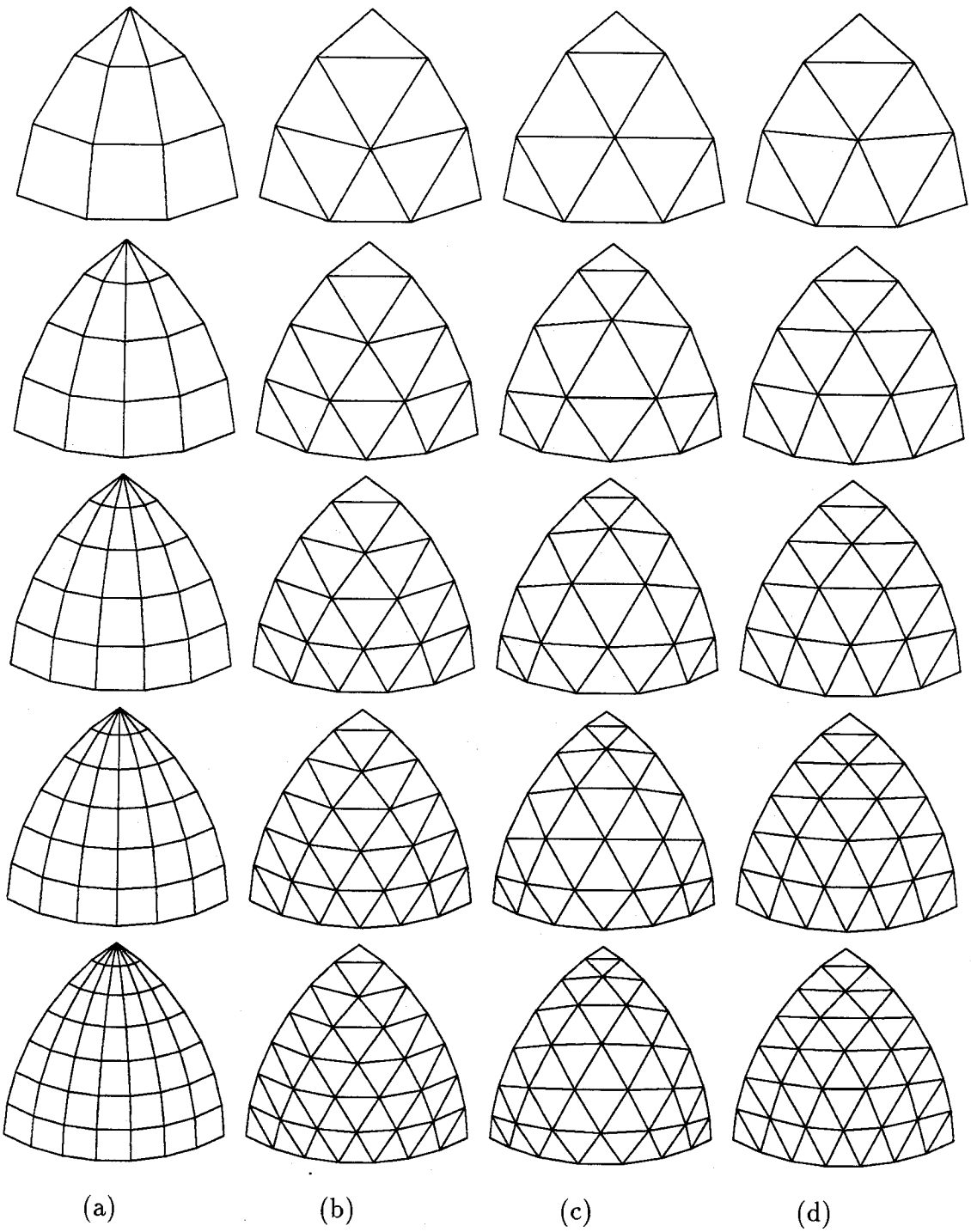


Figure 2.1. Mesh Types. One octant of meshes (a) LL, (b) LT, (c) IJK, and (d) ARC are shown for levels 3 to 7.

Type	(i,j,k) Values		d_n	(x,y,z)	Indep. Coord.
A	$i=0$	$j=0$	6	$(0,0,z)$	0
B	$i=0$	$j=1,\ldots,\frac{L-1}{2}$	24	$(0,y,z)$	1
C	$i=0$	$j=k=\frac{L}{2}$	12	$(0,z,z)$	0
D	$i=1,\ldots,\frac{L-1}{3}$	$j=i$	24	(y,y,z)	1
E	$i=1,\ldots,\frac{L-1}{3}$	$j=i+1,\ldots,\frac{L-i-1}{2}$	48	(x,y,z)	2
F	$i=1,\ldots,\frac{L-1}{3}$	$j=k=\frac{L-i}{2}$	24	(x,z,z)	1
G	$i=\frac{L}{3}$	$j=k=\frac{L}{3}$	8	(z,z,z)	0

Table 2.2. Point Types and Degeneracies d_n for General O_h Mesh.

These independent points can be broken down into seven different types which are listed in Table 2.2, where d_n is defined as the degeneracy of the point. The coordinates of the base point in the next to last column of this table obey $0 \leq x \leq y \leq z \leq 1$. The specific allowed ranges for the independent coordinates for each point type are given in Table 2.3. Relaxing the condition that $x \leq y \leq z$ reduces the distinct types by one since D and F become equivalent. The most general spherical grid with O_h symmetry is a combination of a certain number of each of these point types with distinct assignments of the unspecified x, y, z values. What our method yields is a logical progression of determining the types used for a given sphere level that distributes them evenly and the implied restriction that the (x,y,z) coordinates should be ordered as per their corresponding (i,j,k) values. For example, level 1 has only a type A point, level 2 has a type A point and a type C point, level 3 has one point each of types A, B, and G, etc. Note that since the points must lie on the unit sphere, there are 0, 1, or 2 independent coordinates per point depending on its symmetry. The independent polygons are those in the all positive octant containing any of the points with $i \leq j \leq k$. Those that point up, i.e., have only one point with the larger of the two k values, have (i,j,k) indices at that upper point of

Type	Indices	Point	Range(s)
A	$0 = i = j < k = L$	(0,0,1)	
B	$0 = i < j < k$	$(0, \sqrt{1-z^2}, z)$	$\sqrt{\frac{1}{2}} < z < 1$
C	$0 = i < j = k$	$(0, \sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$	
D	$0 < i = j < k$	$(\sqrt{\frac{1}{2}(1-z^2)}, \sqrt{\frac{1}{2}(1-z^2)}, z)$	$\sqrt{\frac{1}{3}} < z < 1$
E	$0 < i < j < k$	$(\sqrt{1-y^2-z^2}, y, z)$	$0 < y < \sqrt{\frac{1}{2}}$ [a] $\sqrt{\frac{1}{3}} < z < 1$
F	$0 < i < j = k$	$(\sqrt{1-2z^2}, z, z)$	$\sqrt{\frac{1}{3}} < z < \sqrt{\frac{1}{2}}$
G	$0 < i = j = k$	$(\sqrt{\frac{1}{3}}, \sqrt{\frac{1}{3}}, \sqrt{\frac{1}{3}})$	

Table 2.3. Independent Coordinates.

[a] For optimization purposes, $t = y(1-z^2)^{-\frac{1}{2}}$ with $0 < t < 1$ is used to ensure the point is on the unit sphere.

$i = 0, \ldots, \frac{L-1}{3}$, $j = i, \ldots, \frac{L-i-1}{2}$. Those that point down have indices at the lower point of $i = 0, \ldots, \frac{L-1}{3}$, $j = i, \ldots, \frac{L-i}{2}$. That means that there are

$$L\left[\frac{L-1}{3}+1\right] - 3\left[\frac{\left(\frac{L-1}{3}\right)\left(\frac{L-1}{3}+1\right)}{2}\right]$$

independent polygons in sphere level L. See Table 2.4 for a complete list of the features of each level up to twenty.

In method ARC, we let the points $(0, j, L-j)$ be evenly spaced on the great arc segment (0,0,1) to (0,1,0), i.e.,

$$(0, j, L-j) = \left(0, \sin\left(\frac{j\pi}{2L}\right), \cos\left(\frac{j\pi}{2L}\right)\right), \quad j = 0, 1, 2, \ldots, L. \tag{2.4}$$

Similarly, let the points $(i, j, L-i-j)$, $j = i, i+1, i+2, \ldots, L-2i$, be evenly spaced on the fixed-x arc segment (a_i, a_i, b_i) to (a_i, b_i, a_i) for $i = 1, 2, 3, \ldots, \frac{L}{3}$. Of course, the point $(\frac{L}{3}, \frac{L}{3}, \frac{L}{3})$, if it exists, must be $(\sqrt{\frac{1}{3}}, \sqrt{\frac{1}{3}}, \sqrt{\frac{1}{3}})$ by symmetry. By letting the centers of the arcs be equally spaced between the points $(\sqrt{\frac{1}{3}}, \sqrt{\frac{1}{3}}, \sqrt{\frac{1}{3}})$ and $(0, \sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}})$, the constants a_i and b_i and thus all the points are defined. The

Level	No. polys.	No. pnts.	Ind. polys.	Ind. pnts.	Ind. coord.	Type of Points
1	8	6	1	1	0	A
2	32	18	2	2	0	A+C
3	72	38	3	3	1	A+B+G
4	128	66	5	4	2	A+B+C+D
5	200	102	7	5	4	A+2B+D+F
6	288	146	9	7	5	A+2B+C+D+E+G
7	392	198	12	8	8	A+3B+2D+E+F
8	512	258	15	10	10	A+3B+C+2D+2E+F
9	648	326	18	12	13	A+4B+2D+3E+F+G
10	800	402	22	14	16	A+4B+C+3D+4E+F
11	968	486	26	16	20	A+5B+3D+5E+2F
12	1152	578	30	19	23	A+5B+C+3D+7E+F+G
13	1352	678	35	21	28	A+6B+4D+8E+2F
14	1568	786	40	24	32	A+6B+C+4D+10E+2F
15	1800	902	45	27	37	A+7B+4D+2E+12F+G
16	2048	1026	51	30	42	A+7B+C+5D+14E+2F
17	2312	1158	57	33	48	A+8B+5D+16E+3F
18	2592	1298	63	37	53	A+8B+C+5D+19E+2F+G
19	2888	1446	70	40	60	A+9B+6D+21E+3F
20	3200	1602	77	44	66	A+9B+C+6D+24E+3F

Table 2.4. Characteristics of Sphere Levels.

angles between the arcs are:

$$\beta_i = \frac{3i}{L}\cos^{-1}\left(\sqrt{\frac{2}{3}}\right), \quad i = 1, 2, 3, \ldots, \frac{L}{3}. \tag{2.5}$$

The coordinates of the points $(i, i, L-2i) = (a_i, a_i, b_i)$ are then given by $a_i = \sin\beta_i$ and $b_i = \sqrt{1-2a_i{}^2}$. Equally spacing the points $(i, j, L-i-j)$ on the arc segment (a_i, a_i, b_i) to (a_i, b_i, a_i) with x fixed at a_i, we find:

$$(i, j, L-i-j) = (a_i, \cos\beta_i \sin\gamma_{i,j}, \cos\beta_i \cos\gamma_{i,j}), \quad j = i, i+1, i+2, \ldots, L-2i \tag{2.6}$$

for $i = 1, 2, 3, \ldots, \frac{L}{3}$, where

$$\gamma_{i,j} = \left(\frac{j-i}{L-3i}\right)\left(\frac{\pi}{2} - 2\gamma_{i,i}\right) + \gamma_{i,i} \tag{2.7}$$

and $\sin\gamma_{i,i} = tan\beta_i$. The resulting mesh is shown in Figure 2.1, column (d). We mentioned earlier that we desired a uniform distribution of points. Three spatial attributes of a mesh that may help to define an even point distribution are its variation in edge length, polygon area, and area nearest each point. In the next subsection we illustrate methods of determining the standard deviation of these features for grids already presented plus present the grids resulting from minimizing the standard deviation of each mesh attribute.

II.C. Point Distribution Criteria

Applications such as numerical integration have very specific criteria for evaluating a quadrature, but for many other applications the criteria are more vague or too far removed from the grid use to allow grid optimization for the given problem. In these cases, geometric considerations can be used to establish a grid with uniformly distributed points which are more likely to produce desirable results with fewer points.

Before we can minimize the deviation in point distribution over the unit sphere or use it to compare our methods, we must associate a value with it. We now consider what we mean by an even point distribution. If we calculate a property at given points on a sphere for use in interpolation across the entire sphere, the length of a polygon edge indicates how far an interpolated point can be from a calculated point. The size of a polygon can be used to judge how wide an area its vertices must represent. For the purpose of weighting points for function evaluation, we must consider the area associated with each point. In general we are interested in the standard deviation of these properties:

$$\sigma_L^2 = \frac{1}{N_L} \sum_n d_n (a_n - \overline{a_L})^2, \tag{2.8}$$

where σ_L is the standard deviation of a_n in a level L sphere, N_L is the number of points, polygons, or edges on the sphere, n is summed over all independent (non-equivalent with respect to symmetry) points, d_n is the degeneracy of point n, a_n is the property concerned, and $\overline{a_L}$ is the average of a_n over the sphere. To aid comparison among levels, the fractional standard deviation, $\sigma_L/\overline{a_L}$, is actually used in figures. Using degeneracy instead of summing over all points is advantageous when σ_L is to be minimized. We next show how to find σ_L for the three properties edge length, polygon area, and point area (or "nearest area," i.e., the area nearest each point) for method LL, method LT, and any method based on the above triangular mesh under the confines of octahedral symmetry.

The length s of the polygon edge with vertices $\vec{P}_i$ and $\vec{P}_j$ on the unit sphere is found from the law of cosines:

$$s^2 = 2(1 - \vec{P}_i \cdot \vec{P}_j). \tag{2.9}$$

It is thus one of the easiest polygon attributes to calculate and optimize. For each level of the octahedral mesh defined above, the degrees of freedom were varied to

minimize the standard deviation σ_L of the edge length. The fractional standard deviations are shown in Figure 2.2 for various mesh schemes and for levels up to twenty. Meshes optimized below for other geometric properties also are included for comparison, and the oscillation of the point or nearest area method is discussed when it is presented. Note the vast improvement of the other methods over LL, and the near optimal performance of method ARC which was designed to be a good starting point for various mesh optimizations.

The distribution of the areas of individual polygons is the next subject we address. The area we use is that of the sphere segment corresponding to the polygon instead of the polygon itself. This insures that the sum of the areas is 4π and that $\overline{a_L} = 4\pi/N_L$, i.e., the surface area of the sphere and does not vary. The calculation of the areas of spherical polygons is more complicated than the calculation of polygon edges and as such is presented in Appendix A. Again, we were able to find minimal standard deviations of polygon area for each mesh level from one to twenty as shown in Figure 2.3. Again, the LL method is worst, and the ARC method is next best to the optimized method. As the level increases, the greater number of degrees of freedom does allow the optimized mesh to make significant gains over any other method.

In many applications, values are associated with or calculated at each point on the spherical grid. Thus, the idea of an even distribution of points themselves is perhaps more important than even distributions of properties related to polygons or edges. However, there is no property such as an edge length or polygon area associated with points which is as obvious a choice for points. The range of influence of a point is probably best defined by mapping a two-dimensional grid of Wigner-Seitz cells onto the sphere's surface. A Wigner-Seitz cell [1] is that volume or area which is nearest a given point than any other and is alternatively known as Dirichlet regions

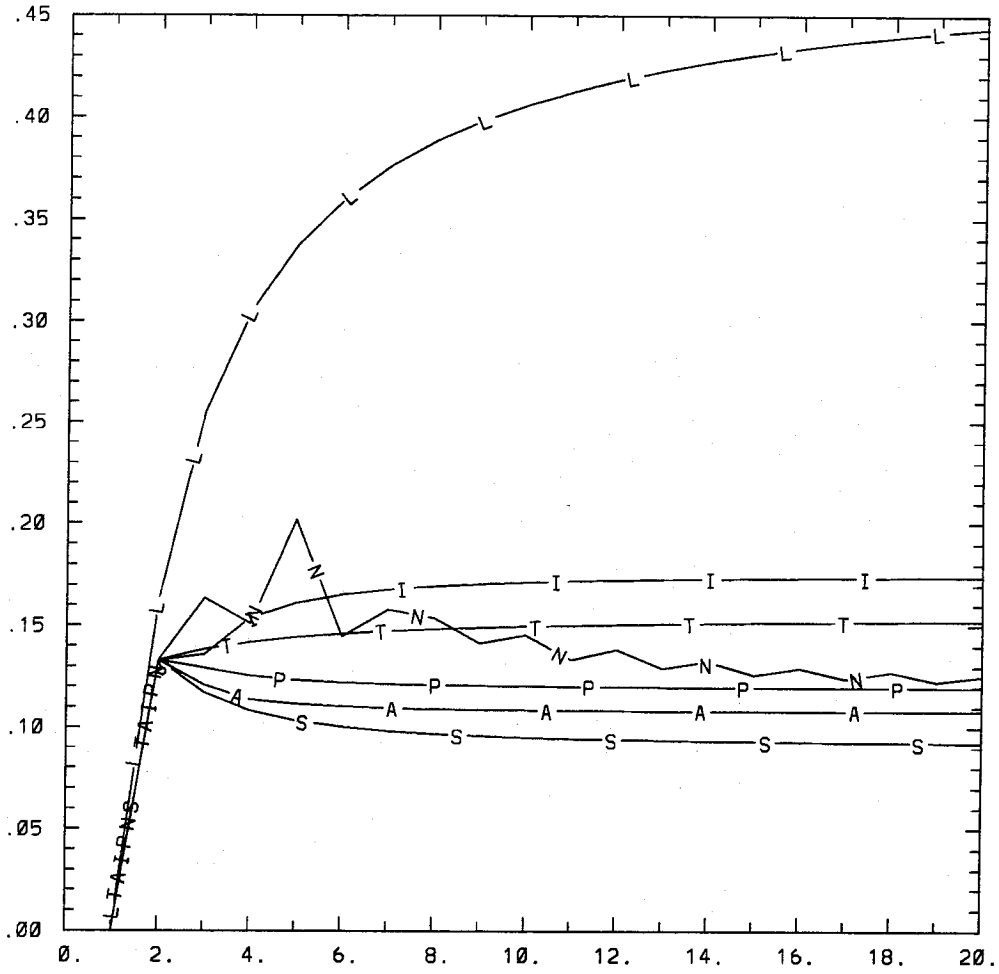


Figure 2.2. Edge Length Fractional Standard Deviation. L is for method LL, T is for method LT, I is for method IJK, A is for method ARC, S is optimized for minimal edge length (side) deviation, P is optimized for minimal polygon area deviation, and N is optimized for minimal point (a.k.a. nearest) area deviation.

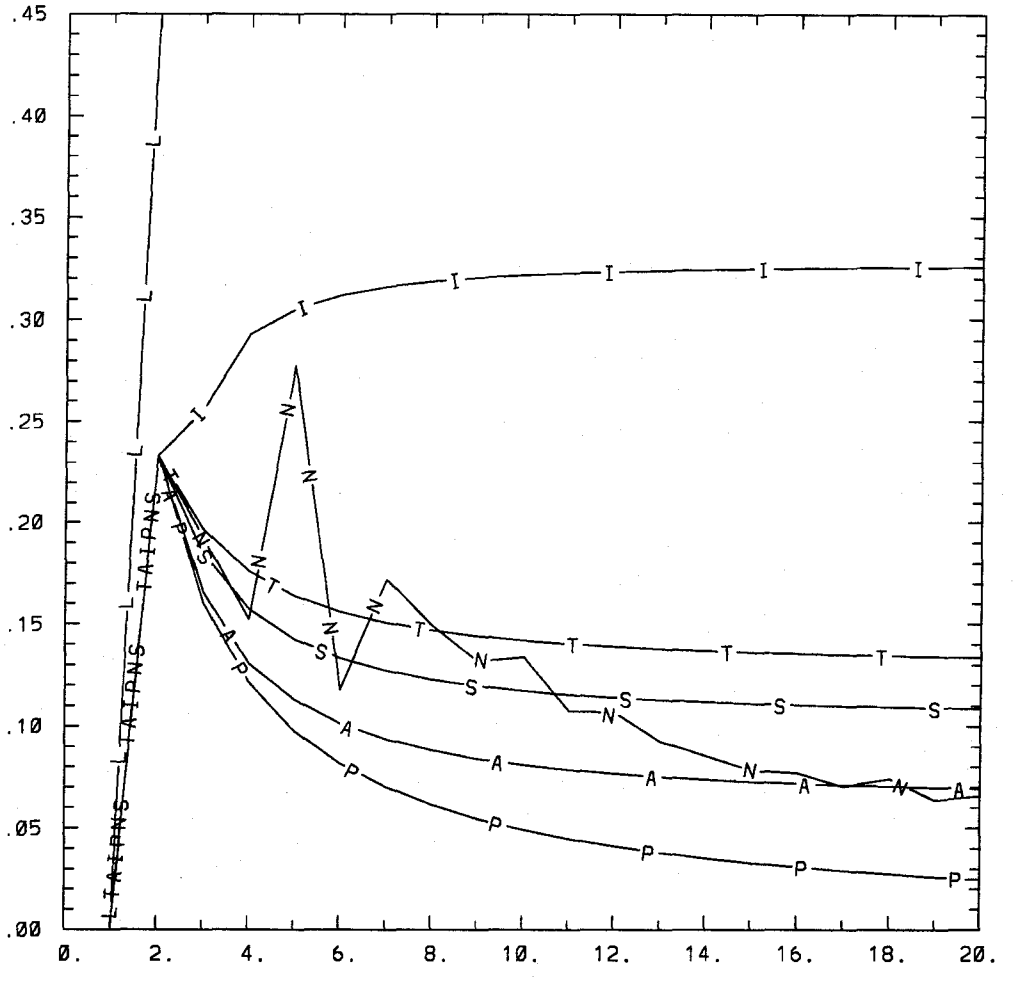


Figure 2.3. Polygon Area Fractional Standard Deviation. L is for method LL, T is for method LT, I is for method IJK, A is for method ARC, S is optimized for minimal edge length (side) deviation, P is optimized for minimal polygon area deviation, and N is optimized for minimal point (a.k.a. nearest) area deviation.

[2], Voronoi polyhedra [3], or domains [4]. In this application, we call this area either the point area or the nearest area (to each point). Again, the actual area used is the surface area of the sphere segment comprising the Wigner-Seitz cell instead of the area of the polygons on the surface. Thus the sum of the areas is still 4π. The use of symmetry aids the computation of nearest areas, so two definitions are useful. First, d_n is the degeneracy of a point on the sphere, i.e., the number of points with which it is symmetrically equivalent. Second, w_{mn} is the weight of polygon m at point n. There are often a number w_{mn} of symmetrically equivalent polygons meeting at a given point n. The area nearest point n can then be found by finding the area is closest to n for only those polygons m which are nonequivalent and multiplying that area by w_{mn} before summing. In order to handle the problem of optimizing deviation in point areas efficiently, it is best to make use of the features of the individual meshes. Thus various mesh methods are discussed in order, making use of the methods developed in Appendix A for finding the (Wigner-Seitz) area nearest a given point, while the resulting area deviations are reported as a group.

For method LL there are $L+1$ independent points with indices $(0, L, k)$ for $k = 0, \ldots, L$ since each point with the same $|k|$ is equivalent. They have degeneracy $d_n = 8L$ except for $k = 0$ and $k = L$ which have $d_n = 4L$ and $d_n = 2$, respectively. There are L independent polygons: the $L-1$ quadrilaterals between k and $k+1$ for $k = 0, \ldots, L-2$ and between $i = 0$ and $i = 1$, and the triangle $(0,0,L) - (1, L-1, 1) - (0, L, L-1)$. Each polygon includes two independent points with weights $w_{mn} = 2$ except for the points $(0,0,L)$ and $(0,L,0)$ which have $w_{mn} = 4L$ and $w_{mn} = 4$, respectively. See Figure 2.4 for the d_n and w_{mn} for $L = 4$ for all mesh methods.

For each quadrilateral between k and $k+1$, name the vertices as $\vec{P}_1 = (0, L, k)$, $\vec{P}_2 = (0, L, k+1)$, $\vec{P}_3 = (1, L-1, k+1)$, and $\vec{P}_4 = (1, L-1, k)$. Consider the area

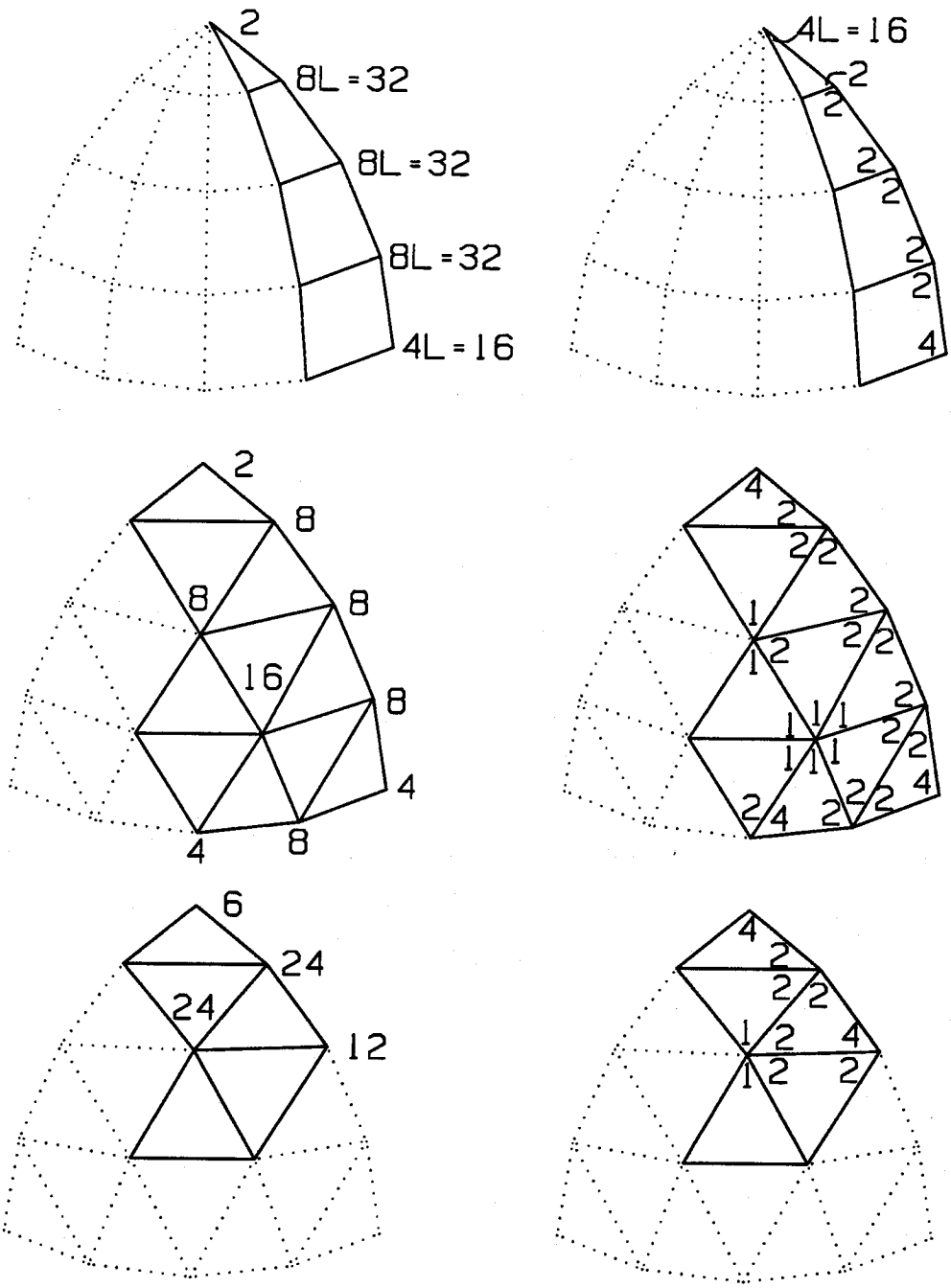


Figure 2.4. Degeneracies and Weights of Mesh Points. The d_n are on the left and the w_{mn} are on the right for methods LL (top), LT (middle), and IJK and ARC (bottom). Independent polygons are drawn with solid lines.

(i,j,k) Values		d_n
$k=0$	$i=0$	4
$k=0$	$i=1,\ldots,\frac{L-1}{2}$	8
$k=0$	$i=j=\frac{L}{2}$	4
$k=1,\ldots,L-1$	$i=0$	8
$k=1,\ldots,L-1$	$i=1,\ldots,\frac{L-k-1}{2}$	16
$k=1,\ldots,L-1$	$i=j=\frac{L-k}{2}$	8
$k=L$	$i=j=0$	2

Table 2.5. Degeneracies d_n for Method LT.

a_{mn} which this quadrilateral contributes to point $\vec{P}_1$. It is the area bounded by a quadrilateral with vertices $\vec{P}_1$, $\vec{P}_{\overline{12}}$, $\vec{C}$, and $\vec{P}_{\overline{41}}$, where $\vec{C}$ is equidistant from the four original points and $\vec{P}_{\overline{ij}}$, is the midpoint of the arc $(\vec{P}_i - \vec{P}_j)$; the area a_{mn} for point $\vec{P}_2$ is bounded by $\vec{P}_2$, $\vec{P}_{\overline{23}}$, $\vec{C}$, and $\vec{P}_{\overline{12}}$. See Appendix A for the methods used to determine $\vec{C}$ and a_{mn}. The general case of finding that part of the area of a triangle closest to one particular vertex is also given in Appendix A as is needed for the independent triangle between $k = L-1$ and $k = L$.

For method LT there are $\left(\frac{L+2}{2}\right)\left(\frac{L+3}{2}\right)$ independent points with (i,j,k) indices $k = 0,\ldots,L,\ \ i = 0,\ldots,\frac{L-k}{2}$. Their degeneracies are given in Table 2.5. There are $[L(L+1)]/2$ independent polygons, the triangles in the all positive octant containing any of the points with $i \leq j$.

For any method based on our generic mesh with octahedral symmetry, the methods of finding nearest areas are equivalent. Point degeneracies and other features are discussed in Subsections II.A and II.B where the method is defined. Appendix A again provides the methods of actually finding the area. Thus, we have all the information required for finding standard deviations of point areas for methods IJK and ARC.

An important distinction can be made for finding minimum standard deviations

for nearest areas versus those for edge lengths and polygon areas. While the greater number of independent edges or independent polygons is always greater than the number of degrees of freedom in our octahedral mesh, Table 2.4 shows that the number of independent coordinates equals and then exceeds the number of independent points as the level L increases. Thus, there is the hope of finding meshes with zero standard deviation of point area for $L \geq 7$. For $L \geq 9$, extra degrees of freedom exist, thus the optimal mesh is not unique. This means that we may be able to find meshes for which each point has the same nearest area for $L \geq 7$. A plot of the fractional standard deviation of the area nearest to each point versus level L is given in Figure 2.5. For level 8, no mesh with $\sigma_8 = 0$ was found as ten equations ($q_n - \overline{a_L} = 0$ for each independent point) in ten unknowns (the independent coordinates of the mesh) is not guaranteed of having a solution. Also note that as L increases, the point nearest the center cycles from a point type G to D to F as defined in Table 2.3. This determines the amount of freedom near the center of the octant and causes the oscillation of the variation of the other properties as seen above.

This section considered various geometric methods of determining a uniform point distribution and presented grids with the narrowest possible distribution of each mesh property. The method ARC is seen to have small variation in all three of the properties considered. A narrow distribution of Wigner-Seitz areas is perhaps the most useful property, and for higher L grids with all equal nearest areas are given. We consider the use of nearest areas for a weighting scheme of numerical quadratures in the next section.

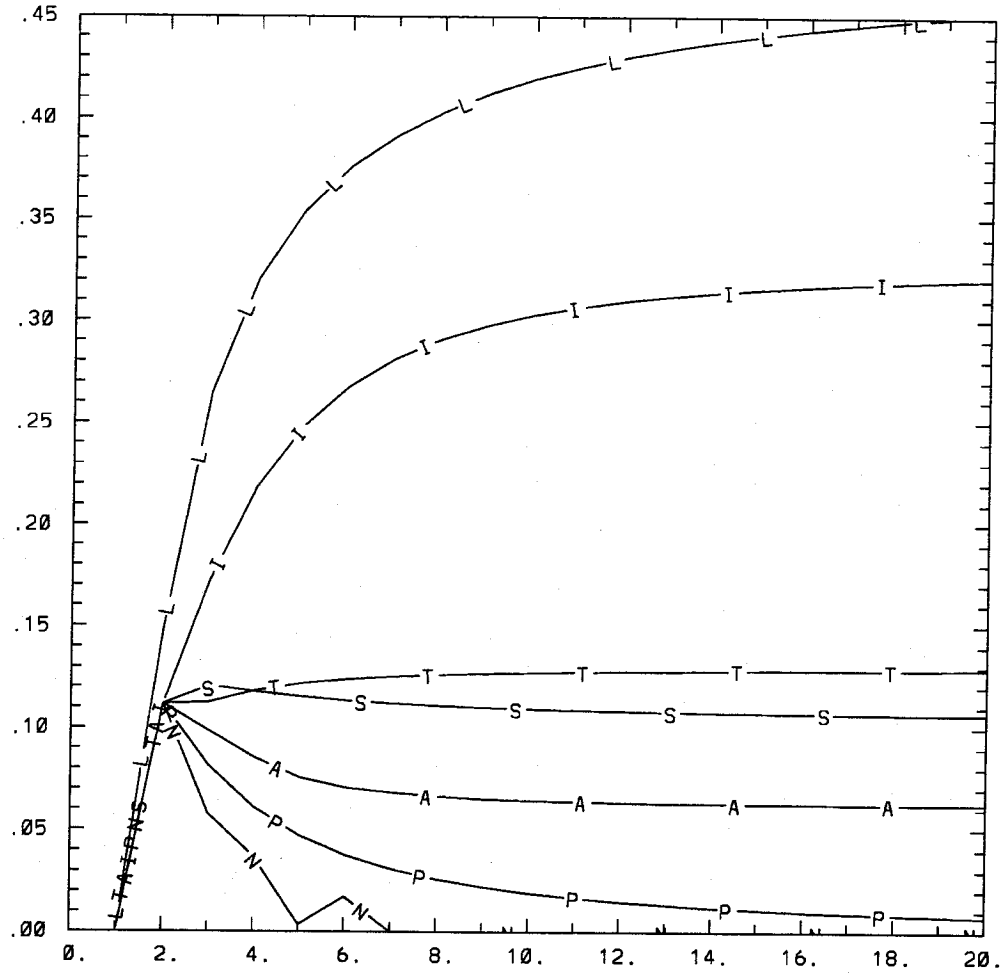


Figure 2.5. Point Area Fractional Standard Deviation. L is for method LL, T is for method LT, I is for method IJK, A is for method ARC, S is optimized for minimal edge length (side) deviation, P is optimized for minimal polygon area deviation, and N is optimized for minimal point (a.k.a. nearest) area deviation.

III. Function Evaluation on a Spherical Surface

The numerical evaluation of functions on the surface of a sphere arises in many fields. A set of points with associated weights are required. Often the points and/or weights are chosen to handle the numerical evaluation exactly for a small set of problems. By far the most important example is numerical integration. Here, the quadratures, i.e., points plus weights, are chosen to exactly integrate a given set of functions which usually are part of a basis set. The hope is that when the functions to be integrated are expanded in terms of the basis set that higher order terms are increasingly small. However, a second major concern is to use the fewest number of points to promote computation speed. Thus, the number of basis functions which can be exactly integrated must be limited. Often quadratures are thus developed in which both the points and weights are allowed to vary, thus integrating the greatest number of points with the fewest points. These are generally called Gauss quadratures. A second common method is to weight all points equally in the hope that this produces greater accuracy for the part of a function which cannot be represented by the basis set. These are called Chebyshev quadratures. It is our belief that neither of these methods are optimal. Fixing the weights to be proportional to the area nearest each point as defined in the previous section leads to a new quadrature scheme. It makes use of variation of the point positions in order to exactly integrate a given set of functions and thus requires the same number of points as the Chebyshev method. However, by assigning the point weight logically, i.e., by the area which the point represents, integration of functions not in the basis set is expected to be much improved. Spherical harmonics are the most likely choice of a basis for spherical integration grids. The real spherical harmonics are presented and briefly discussed in the first subsection. They are displayed using polar plots in yet another application

of spherical grids. Their relation with octahedral symmetry is then discussed to set up the following treatment of spherical quadratures using grids with octahedral symmetry.

III.A. Real Spherical Harmonics

The classical (real) spherical harmonics are given by [5]:

$$\begin{aligned} Y_{l0}(\theta,\phi) &= \sqrt{\frac{2l+1}{4\pi}}\, P_l(\cos\theta), \\ Y_{lm}^{(+)}(\theta,\phi) &= \sqrt{\frac{2l+1}{2\pi}}\sqrt{\frac{(l-m)!}{(l+m)!}}\, P_l^m(\cos\theta)\cos m\phi, \\ Y_{lm}^{(-)}(\theta,\phi) &= \sqrt{\frac{2l+1}{2\pi}}\sqrt{\frac{(l-m)!}{(l+m)!}}\, P_l^m(\cos\theta)\sin m\phi, \end{aligned} \tag{2.10}$$

where in this case, $m = 0, 1, \ldots, l$. To simplify our tables, we instead use the notation:

$$Z_{lm} \equiv \begin{cases} Y_{lm}^{(+)}, & m > 0 \\ Y_{lm}, & m = 0 \\ Y_{l\,-m}^{(-)}, & m < 0 \end{cases}, \tag{2.11}$$

where now $m = -l, -l+1, \ldots, l$. These satisfy the normalization condition:

$$\iint Z_{lm} Z_{l'm'} d\Omega = \delta_{ll'}\delta_{mm'}, \tag{2.12}$$

where $d\Omega = \sin\theta d\theta d\phi$. The quantum mechanical (complex) spherical harmonics (using the Condon-Shortley phase convention) are [5, 6]:

$$Y_{lm}(\theta,\phi) = (-1)^{\frac{|m|+m}{2}} \left(\frac{2l+1}{4\pi}\right)^{\frac{1}{2}} \left(\frac{(l-|m|)!}{(l+|m|)!}\right)^{\frac{1}{2}} P_l^{|m|}(\cos\theta) e^{im\phi}. \tag{2.13}$$

These two sets of functions are related by:

$$Y_{lm} = \begin{cases} (-1)^m \sqrt{\frac{1}{2}}\,(Z_{lm} + iZ_{l\,-m}), & m > 0 \\ Z_{lm}, & m = 0 \\ \sqrt{\frac{1}{2}}\,(Z_{l\,-m} - iZ_{lm}), & m < 0 \end{cases} \tag{2.14}$$

and

$$Z_{lm} = \begin{cases} \sqrt{\frac{1}{2}}[(-1)^m Y_{lm} + Y_{l\,-m}], & m > 0 \\ Y_{lm}, & m = 0 \\ i\sqrt{\frac{1}{2}}[(-1)^{m+1} Y_{l\,-m} + Y_{lm}], & m < 0 \end{cases} . \tag{2.15}$$

The Z_{lm}'s up through $l = 4$ are given in Table 2.6, which is an extension of Table 1.2 of Reference [6].

The real spherical harmonics are best plotted as polar plots where

$$r = |Z_{lm}(\theta, \phi)|. \tag{2.16}$$

The points $\vec{P}$ on this surface are thus

$$\vec{P} = |Z_{lm}|(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) = |Z_{lm}|(x, y, z), \tag{2.17}$$

where, in this paragraph, we assume (x, y, z) is a point on the unit sphere, i.e., $x^2 + y^2 + z^2 = 1$. The unnormalized normals $\vec{N}$ to this surface are given by

$$\vec{N} = (\frac{\partial \vec{P}}{\partial \theta}) \times (\frac{\partial \vec{P}}{\partial \phi}), \tag{2.18}$$

with

$$\begin{aligned} \frac{\partial \vec{P}}{\partial \theta} &= \frac{Z_{lm}}{|Z_{lm}|}\frac{\partial Z_{lm}}{\partial \theta}(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \\ &\quad + |Z_{lm}|(\cos\theta\cos\phi, \cos\theta\sin\phi, -\sin\theta), \end{aligned} \tag{2.19}$$

$$\begin{aligned} \frac{\partial \vec{P}}{\partial \phi} &= \frac{Z_{lm}}{|Z_{lm}|}\frac{\partial Z_{lm}}{\partial \phi}(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \\ &\quad + |Z_{lm}|(-\sin\theta\sin\phi, \sin\theta\cos\phi, 0). \end{aligned} \tag{2.20}$$

In terms of the Cartesian coordinates x, y, and z, these are

$$\frac{\partial \vec{P}}{\partial \theta} = \frac{Z_{lm}}{|Z_{lm}|}\frac{\partial Z_{lm}}{\partial \theta}(x, y, z) + |Z_{lm}|(z\frac{x}{\rho}, z\frac{y}{\rho}, -\rho), \tag{2.21}$$

$$\frac{\partial \vec{P}}{\partial \phi} = \frac{Z_{lm}}{|Z_{lm}|}\frac{\partial Z_{lm}}{\partial \phi}(x, y, z) + |Z_{lm}|(-y, x, 0), \tag{2.22}$$

l	m	$\sqrt{\frac{4\pi}{2l+1}}\, Z_{lm}$	$\sqrt{\frac{4\pi}{2l+1}}\, r^l Z_{lm}$
0	0	1	1
1	0	$\cos\theta$	z
1	1	$\sin\theta\cos\phi$	x
1	−1	$\sin\theta\sin\phi$	y
2	0	$\frac{1}{2}(3\cos^2\theta - 1)$	$\frac{1}{2}(3z^2 - r^2)$
2	1	$\sqrt{3}\ \cos\theta\sin\theta\cos\phi$	$\sqrt{3}\ zx$
2	−1	$\sqrt{3}\ \cos\theta\sin\theta\sin\phi$	$\sqrt{3}\ zy$
2	2	$\frac{1}{2}\sqrt{3}\ \sin^2\theta(\cos^2\phi - \sin^2\phi)$	$\frac{1}{2}\sqrt{3}\ (x^2 - y^2)$
2	−2	$\sqrt{3}\ \sin^2\theta\cos\phi\sin\phi$	$\sqrt{3}\ xy$
3	0	$\frac{1}{2}(5\cos^2\theta - 3)\cos\theta$	$\frac{1}{2}(5z^2 - 3r^2)z$
3	1	$\frac{1}{2}\sqrt{\frac{3}{2}}\ (5\cos^2\theta - 1)\sin\theta\cos\phi$	$\frac{1}{2}\sqrt{\frac{3}{2}}\ (5z^2 - r^2)x$
3	−1	$\frac{1}{2}\sqrt{\frac{3}{2}}\ (5\cos^2\theta - 1)\sin\theta\sin\phi$	$\frac{1}{2}\sqrt{\frac{3}{2}}\ (5z^2 - r^2)y$
3	2	$\frac{1}{2}\sqrt{15}\ \cos\theta\sin^2\theta(\cos^2\phi - \sin^2\phi)$	$\frac{1}{2}\sqrt{15}\ z(x^2 - y^2)$
3	−2	$\frac{1}{2}\sqrt{15}\ \cos\theta\sin^2\theta\cos\phi\sin\phi$	$\sqrt{15}\ zxy$
3	3	$\frac{1}{2}\sqrt{\frac{5}{2}}\ \sin^3\theta\cos\phi(\cos^2\phi - 3\sin^2\phi)$	$\frac{1}{2}\sqrt{\frac{5}{2}}\ x(x^2 - 3y^2)$
3	−3	$\frac{1}{2}\sqrt{\frac{5}{2}}\ \sin^3\theta\sin\phi(3\cos^2\phi - \sin^2\phi)$	$\frac{1}{2}\sqrt{\frac{5}{2}}\ y(3x^2 - y^2)$
4	0	$\frac{1}{8}(35\cos^4\theta - 30\cos^2\theta + 3)$	$\frac{1}{8}(35z^4 - 30z^2r^2 + 3r^4)$
4	1	$\frac{1}{2}\sqrt{\frac{5}{2}}\ (7\cos^2\theta - 3)\cos\theta\sin\theta\cos\phi$	$\frac{1}{2}\sqrt{\frac{5}{2}}\ (7z^2 - 3r^2)zx$
4	−1	$\frac{1}{2}\sqrt{\frac{5}{2}}\ (7\cos^2\theta - 3)\cos\theta\sin\theta\sin\phi$	$\frac{1}{2}\sqrt{\frac{5}{2}}\ (7z^2 - 3r^2)zy$
4	2	$\frac{1}{4}\sqrt{5}\ (7\cos^2\theta - 1)\sin^2\theta(\cos^2\phi - \sin^2\phi)$	$\frac{1}{4}\sqrt{5}\ (7z^2 - r^2)(x^2 - y^2)$
4	−2	$\frac{1}{2}\sqrt{5}\ (7\cos^2\theta - 1)\sin^2\theta\cos\phi\sin\phi$	$\frac{1}{2}\sqrt{5}\ (7z^2 - r^2)xy$
4	3	$\frac{1}{2}\sqrt{\frac{35}{2}}\ \cos\theta\sin^3\theta\cos\phi(\cos^2\phi - 3\sin^2\phi)$	$\frac{1}{2}\sqrt{\frac{35}{2}}\ zx(x^2 - 3y^2)$
4	−3	$\frac{1}{2}\sqrt{\frac{35}{2}}\ \cos\theta\sin^3\theta\sin\phi(3\cos^2\phi - \sin^2\phi)$	$\frac{1}{2}\sqrt{\frac{35}{2}}\ zy(3x^2 - y^2)$
4	4	$\frac{1}{8}\sqrt{35}\ \sin^4\theta(\cos^4\phi - 6\cos^2\phi\sin^2\phi + \sin^4\phi)$	$\frac{1}{8}\sqrt{35}\ (x^4 - 6x^2y^2 + y^4)$
4	−4	$\frac{1}{2}\sqrt{35}\ \sin^4\theta\cos\phi\sin\phi(\cos^2\phi - \sin^2\phi)$	$\frac{1}{2}\sqrt{35}\ xy(x^2 - y^2)$

Table 2.6. Real Spherical Harmonics.

l	m	$\sqrt{\frac{4\pi}{2l+1}}\,\frac{\partial Z_{lm}}{\partial\theta}$	$\sqrt{\frac{4\pi}{2l+1}}\,r^l\frac{\partial Z_{lm}}{\partial\theta}$
0	0	0	0
1	0	$-\sin\theta$	$-\rho$
1	1	$\cos\theta\cos\phi$	zx/ρ
1	-1	$\cos\theta\sin\phi$	zy/ρ
2	0	$-3\cos\theta\sin\theta$	$-3z\rho$
2	1	$\sqrt{3}\,(2\cos^2\theta-1)\cos\phi$	$\sqrt{3}\,(2z^2-r^2)x/\rho$
2	-1	$\sqrt{3}\,(2\cos^2\theta-1)\sin\phi$	$\sqrt{3}\,(2z^2-r^2)y/\rho$
2	2	$\sqrt{3}\,\cos\theta\sin\theta(\cos^2\phi-\sin^2\phi)$	$\sqrt{3}\,z(x^2-y^2)/\rho$
2	-2	$2\sqrt{3}\,\cos\theta\sin\theta\cos\phi\sin\phi$	$2\sqrt{3}\,zxy/\rho$
3	0	$\frac{1}{2}(3-15\cos^2\theta)\sin\theta$	$\frac{1}{2}(3r^2-15z^2)\rho$
3	1	$\frac{1}{2}\sqrt{\frac{3}{2}}\,(5\cos^2\theta-10\sin^2\theta-1)\cos\theta\cos\phi$	$\frac{1}{2}\sqrt{\frac{3}{2}}\,(5z^2-10\rho^2-r^2)zx/\rho$
3	-1	$\frac{1}{2}\sqrt{\frac{3}{2}}\,(5\cos^2\theta-10\sin^2\theta-1)\cos\theta\sin\phi$	$\frac{1}{2}\sqrt{\frac{3}{2}}\,(5z^2-10\rho^2-r^2)zy/\rho$
3	2	$\frac{1}{2}\sqrt{15}\,(2\cos^2\theta-\sin^2\theta)\sin\theta(\cos^2\phi-\sin^2\phi)$	$\frac{1}{2}\sqrt{15}\,(2z^2-\rho^2)(x^2-y^2)/\rho$
3	-2	$\sqrt{15}\,(2\cos^2\theta-\sin^2\theta)\sin\theta\sin\phi\cos\phi$	$\sqrt{15}\,(2z^2-\rho^2)xy/\rho$
3	3	$\frac{3}{2}\sqrt{\frac{5}{2}}\,\cos\theta\sin^2\theta(\cos^2\phi-3\sin^2\phi)\cos\phi$	$\frac{3}{2}\sqrt{\frac{5}{2}}\,z(x^2-3y^2)x/\rho$
3	-3	$\frac{3}{2}\sqrt{\frac{5}{2}}\,\cos\theta\sin^2\theta(3\cos^2\phi-\sin^2\phi)\sin\phi$	$\frac{3}{2}\sqrt{\frac{5}{2}}\,z(3x^2-y^2)y/\rho$

Table 2.7. Derivatives of Real Spherical Harmonics w.r.t. θ.

where $\rho \equiv \sqrt{x^2+y^2}$. If $\rho = 0$ ($x = 0,\ y = 0,\ z = \pm 1$), we set $\frac{x}{\rho} = \cos\phi$ to z and $\frac{y}{\rho} = \sin\phi$ to 0. We may also plot $r = Z_{lm}^2$ as a polar plot with:

$$\vec{P} = Z_{lm}^2(x,y,z), \tag{2.23}$$

$$\frac{\partial\vec{P}}{\partial\theta} = 2Z_{lm}\frac{\partial Z_{lm}}{\partial\theta}(x,y,z) + Z_{lm}^2(z\frac{x}{\rho}, z\frac{y}{\rho}, -\rho), \tag{2.24}$$

$$\frac{\partial\vec{P}}{\partial\phi} = 2Z_{lm}\frac{\partial Z_{lm}}{\partial\phi}(x,y,z) + Z_{lm}^2(-y,x,0), \tag{2.25}$$

See Tables 2.7 and 2.8 for the partial derivatives of the Z_{lm}'s with respect to θ and ϕ, respectively, up through $l = 3$. The polar plots are easily made using a spherical mesh such as ARC by scaling the unit sphere points by Z_{lm} or Z_{lm}^2 and

l	m	$\sqrt{\frac{4\pi}{2l+1}}\,\frac{\partial Z_{lm}}{\partial \phi}$	$\sqrt{\frac{4\pi}{2l+1}}\,r^l\frac{\partial Z_{lm}}{\partial \phi}$
0	0	0	0
1	0	0	0
1	1	$-\sin\theta\sin\phi$	$-y$
1	-1	$\sin\theta\cos\phi$	x
2	0	0	0
2	1	$-\sqrt{3}\ \cos\theta\sin\theta\sin\phi$	$-\sqrt{3}\ zy$
2	-1	$\sqrt{3}\ \cos\theta\sin\theta\cos\phi$	$\sqrt{3}\ zx$
2	2	$-2\sqrt{3}\ \sin^2\theta\cos\phi\sin\phi$	$-2\sqrt{3}\ xy$
2	-2	$\sqrt{3}\ \sin^2\theta(\cos^2\phi-\sin^2\phi)$	$\sqrt{3}\ (x^2-y^2)$
3	0	0	0
3	1	$-\frac{1}{2}\sqrt{\frac{3}{2}}\ (5\cos^2\theta-1)\sin\theta\sin\phi$	$-\frac{1}{2}\sqrt{\frac{3}{2}}\ (5z^2-r^2)y$
3	-1	$\frac{1}{2}\sqrt{\frac{3}{2}}\ (5\cos^2\theta-1)\sin\theta\cos\phi$	$\frac{1}{2}\sqrt{\frac{3}{2}}\ (5z^2-r^2)x$
3	2	$-2\sqrt{15}\ \cos\theta\sin^2\theta\cos\phi\sin\phi$	$-2\sqrt{15}\ zxy$
3	-2	$\sqrt{15}\ \cos\theta\sin^2\theta(\cos^2\phi-\sin^2\phi)$	$\sqrt{15}\ z(x^2-y^2)$
3	3	$\frac{3}{2}\sqrt{\frac{5}{2}}\ \sin^3\theta\sin\phi(\sin^2\phi-3\cos^2\phi)$	$\frac{3}{2}\sqrt{\frac{5}{2}}\ y(y^2-3x^2)$
3	-3	$\frac{3}{2}\sqrt{\frac{5}{2}}\ \sin^3\theta\cos\phi(\cos^2\phi-3\sin^2\phi)$	$\frac{3}{2}\sqrt{\frac{5}{2}}\ x(x^2-3y^2)$

Table 2.8. Derivatives of Real Spherical Harmonics w.r.t. ϕ.

finding the associated unit normals for use in Gouraud shading. Using the same mesh connectivity as for the sphere, the plots in Figures 2.6 and 2.7 were made for Z_{lm} and Z^2_{lm}, respectively. In the first figure, the positive lobes are colored green while the negative lobes are colored red. The relation of the real spherical harmonics with octahedral symmetry is considered next.

III.B. Z_{lm}'s and Octahedral Symmetry

Since we enforce octahedral with inversion (O_h) symmetry on the spherical mesh presented in Section II, it is worth considering the sets of Z_{lm}'s that form bases for the representations of O_h. We derive them next for the pure rotation group O since extension to O_h is trivial given that O_h is the direct product group $O \times C_i$, where C_i is the inversion group [7].

We first need to find the (reducible for $l > 1$) representation of O for which the set of $(2l+1)$ Z_{lm}'s of same l form a basis. A rotation of angle α about the z-axis (C^z_α) has the following effect:

$$C^z_\alpha Z_{lm} = \begin{cases} Z_{lm}\cos m\alpha + Z_{l\,-m}\sin m\alpha, & m > 0 \\ Z_{lm}, & m = 0 \\ Z_{lm}\cos m\alpha - Z_{l\,-m}\sin m\alpha, & m < 0 \end{cases} . \tag{2.26}$$

We use the convention of rotating the function instead of the coordinate system. The character of C^z_α is thus

$$\chi(C^z_\alpha) = \sum_{m=-l}^{l} \cos m\alpha = \frac{\sin(l+\frac{1}{2})\alpha}{\sin\frac{1}{2}\alpha}. \tag{2.27}$$

Actually, this is the character of a rotation about any axis since a new set of $(2l+1)$ functions orientated about the new axis instead of the z-axis is an orthonormal transform of the original set. This enables us to derive the representations given in Table 2.9.

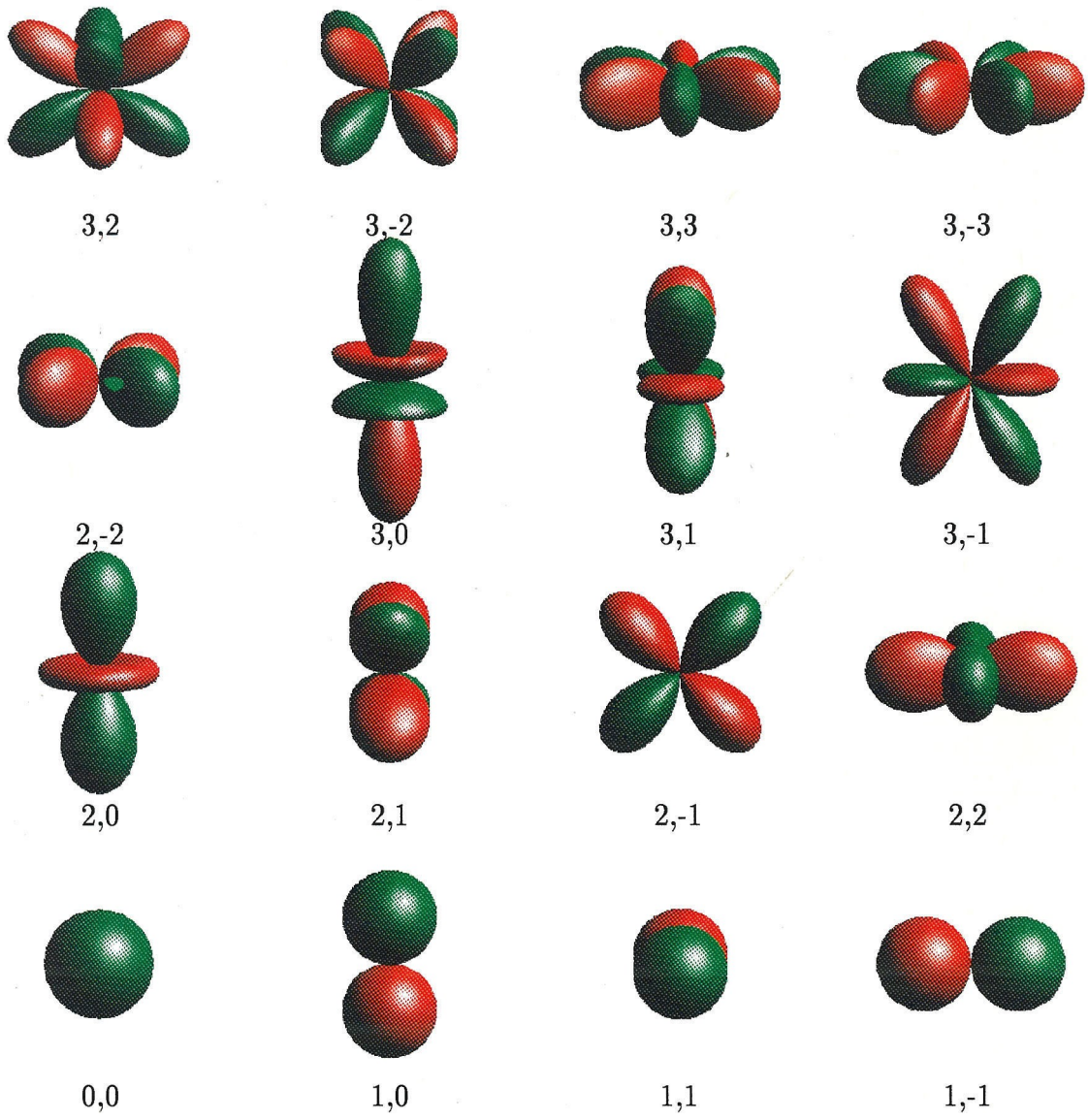


Figure 2.6. Polar Plots of Z_{lm}'s. Numbers below each figure are l, m.

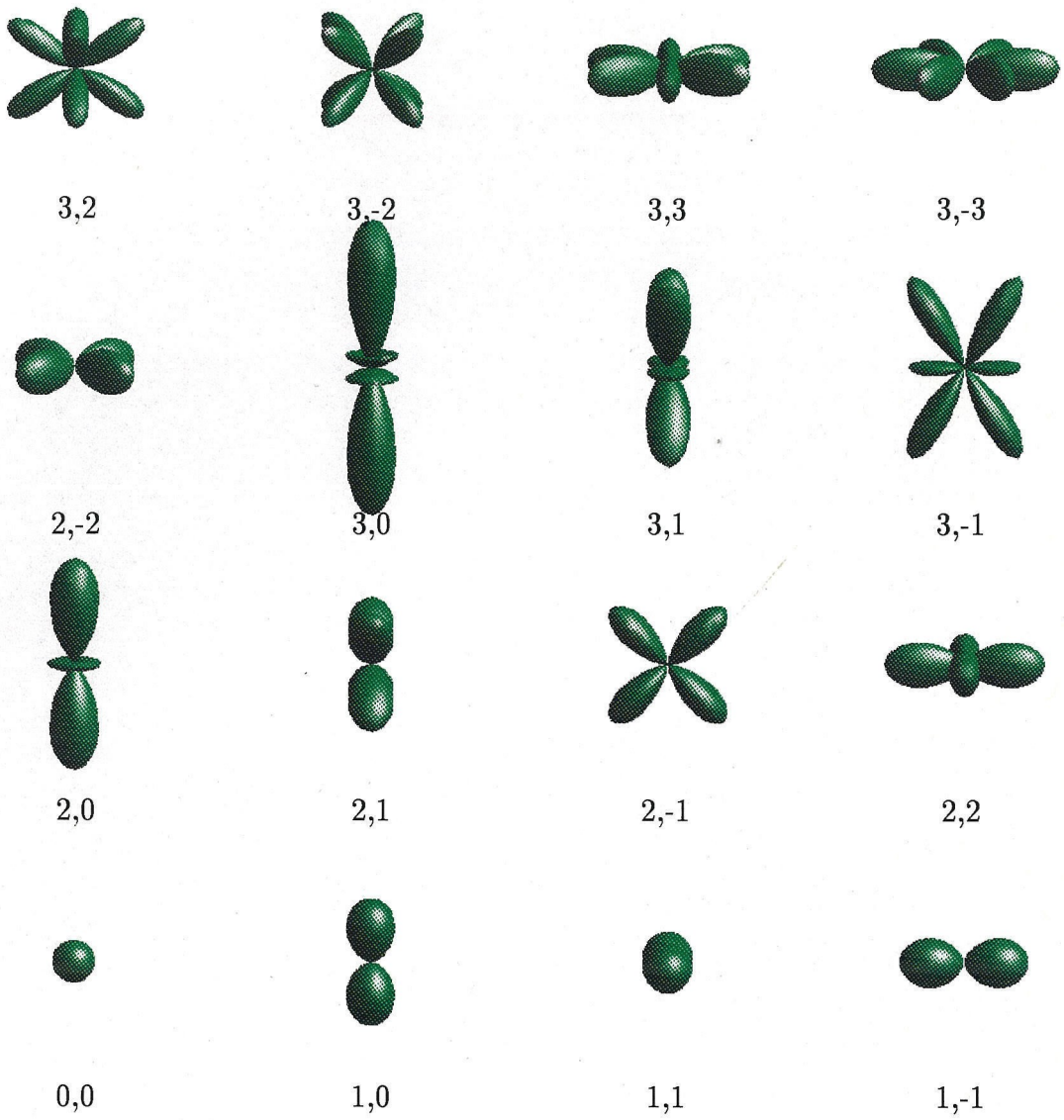


Figure 2.7. Polar Plots of Z_{lm}^2's. Numbers below each figure are l, m.

l	E	$8C_3$	$3C_2$	$6C_2'$	$6C_4$	Representation
0	1	1	1	1	1	A_1
1	3	0	-1	-1	1	T_1
2	5	-1	1	1	-1	$E+T2$
3	7	1	-1	-1	-1	$A_2+T_1+T_2$
4	9	0	1	1	1	$A_1+E+T_1+T_2$

Table 2.9. Irreducible Representations of O of $(2l+1)$ Z_{lm}'s with Same l.

In order to find the particular combinations of Z_{lm}'s that form a basis for the irreducible representations of Table 2.9, we consider the effects on the Z_{lm}'s by C_4^z (a rotation of $\frac{\pi}{2}$ counterclockwise about the z axis) and $C_3^{(111)}$ (a rotation of $\frac{2\pi}{3}$ about the axis from the origin through the point (1,1,1)) which are generators of the group O. We first must define the standard bases for the irreducible representations of O. We denote the basis for the totally symmetric representation A_1 as $|A_1a_1\rangle$. Due to the isomorphism of O with S_4, the symmetric permutation group on four symbols, the character of A_2 is 1 for a rotation in O of even parity and -1 for a rotation of odd parity with respect to interchange of the vertices of an octahedron. We denote the basis of A_2 as $|A_2a_2\rangle$, noting that $C_4^z a_2 = -a_2$ and $C_3^{(111)}a_2 = a_2$. For the basis of T_1, we use $|T_1x\rangle \propto x$, $|T_1y\rangle \propto y$, and $|T_1z\rangle \propto z$. The effect of our generators on this basis is $C_4^z(x,y,z) = (y,-x,z)$ and $C_3^{(111)}(x,y,z) = (y,z,x)$. Using the fact that $T_2 = A_2 \times T_1$ and denoting its basis as $|T_2\xi\rangle$, $|T_2\eta\rangle$, and $|T_2\zeta\rangle$, we have $C_4^z(\xi,\eta,\zeta) = (-\eta,\xi,-\zeta)$ and $C_3^{(111)}(\xi,\eta,\zeta) = (\eta,\zeta,\xi)$. Note that (ξ,η,ζ) transforms as (yz,xz,xy). Lastly, let the basis for E be $|E\theta\rangle \propto 2z^2-x^2-y^2$ and $|E\epsilon\rangle \propto \sqrt{3}(x^2-y^2)$. Making use of the effect of our group generators on (x,y,z), we determine $C_4^z(\theta,\epsilon) = (\theta,-\epsilon)$ and $C_3^{(111)}(\theta,\epsilon) = (-\frac{1}{2}\theta+\frac{\sqrt{3}}{2}\epsilon, -\frac{\sqrt{3}}{2}\theta-\frac{1}{2}\epsilon)$. These results are tabulated in Table 2.10. Once deriving the effects of C_4^z and $C_3^{(111)}$ on the

O		C_4^z	$C_3^{(111)}$
A_1	a_1	a_1	a_1
A_2	a_2	$-a_2$	a_2
E	θ	θ	$-\frac{1}{2}\theta + \frac{\sqrt{3}}{2}\epsilon$
	ϵ	$-\epsilon$	$-\frac{\sqrt{3}}{2}\theta - \frac{1}{2}\epsilon$
T_1	x	y	y
	y	$-x$	z
	z	z	x
T_2	ξ	$-\eta$	η
	η	ξ	ζ
	ζ	$-\zeta$	ξ

Table 2.10. Standard Bases for O and Effect of Group Generators.

Z_{lm}'s (see Table 2.11), comparing with the effects on our standard basis functions leads us to the desired bases.

By comparing Tables 2.10 and 2.11 or looking at the functions themselves, the bases for $l = 0, 1, 2$ are easily found. For $l = 3$, noting that $Z_{3\,-2}$ transforms as a_2 tells us that it is the basis for A_2. Noting that $C_4^z Z_{30} = Z_{30}$, it must be $|T_1 z\rangle$ since z is the only function of either T_1 or T_2 to transform into itself under C_4^z. It immediately follows that:

$$\begin{aligned} |T_1 x\rangle &= C_3^{(111)}|T_1 z\rangle &&= -\tfrac{1}{2}\sqrt{\tfrac{3}{2}}Z_{31} + \tfrac{1}{2}\sqrt{\tfrac{5}{2}}Z_{33}, \\ |T_1 y\rangle &= C_4^z|T_1 x\rangle &&= -\tfrac{1}{2}\sqrt{\tfrac{3}{2}}Z_{3\,-1} - \tfrac{1}{2}\sqrt{\tfrac{5}{2}}Z_{3\,-3}. \end{aligned}$$

Similarly noting that $C_4^z Z_{32} = -Z_{32}$, we have $|T_2\zeta\rangle = Z_{32}$ and thus:

$$\begin{aligned} |T_2\xi\rangle &= C_3^{(111)}|T_2\zeta\rangle &&= -\tfrac{1}{2}\sqrt{\tfrac{5}{2}}Z_{31} - \tfrac{1}{2}\sqrt{\tfrac{3}{2}}Z_{33}, \\ |T_2\eta\rangle &= C_4^z|T_2\xi\rangle &&= -\tfrac{1}{2}\sqrt{\tfrac{5}{2}}Z_{3\,-1} + \tfrac{1}{2}\sqrt{\tfrac{3}{2}}Z_{3\,-3}. \end{aligned}$$

The representation of $l = 4$ is $A_1 + E + T_1 + T_2$. Of the standard basis functions of these, only $|E\epsilon\rangle$ and $|T_2\zeta\rangle$ transform under C_4^z to minus themselves. This is true for Z_{42} and $Z_{4\,-2}$. To distinguish between them, we use the fact that $C_3^{(111)}|E\epsilon\rangle$ has

l	m	$C_4^z Z_{lm}$	$C_3^{(111)} Z_{lm}$
0	0	Z_{00}	Z_{00}
1	1	Z_{10}	Z_{11}
1	1	$Z_{1\,-1}$	$Z_{1\,-1}$
1	-1	$-Z_{11}$	Z_{10}
2	0	Z_{20}	$-\frac{1}{2}Z_{20}+\frac{1}{2}\sqrt{3}Z_{22}$
2	1	$Z_{2\,-1}$	$Z_{2\,-2}$
2	-1	$-Z_{21}$	Z_{21}
2	2	$-Z_{22}$	$-\frac{1}{2}\sqrt{3}Z_{20}-\frac{1}{2}Z_{22}$
2	-2	$-Z_{2\,-2}$	$Z_{2\,-1}$
3	0	Z_{30}	$-\frac{1}{2}\sqrt{\frac{3}{2}}Z_{31}+\frac{1}{2}\sqrt{\frac{5}{2}}Z_{33}$
3	1	$Z_{3\,-1}$	$-\frac{1}{4}Z_{3\,-1}+\frac{1}{4}\sqrt{15}Z_{3\,-3}$
3	-1	$-Z_{31}$	$-\frac{1}{2}\sqrt{\frac{3}{2}}Z_{30}+\frac{1}{2}\sqrt{\frac{5}{2}}Z_{32}$
3	2	$-Z_{32}$	$-\frac{1}{2}\sqrt{\frac{5}{2}}Z_{31}-\frac{1}{2}\sqrt{\frac{3}{2}}Z_{33}$
3	-2	$-Z_{3\,-2}$	$Z_{3\,-2}$
3	3	$-Z_{3\,-3}$	$-\frac{1}{4}\sqrt{15}Z_{3\,-1}-\frac{1}{4}Z_{3\,-3}$
3	-3	Z_{33}	$-\frac{1}{2}\sqrt{\frac{5}{2}}Z_{30}-\frac{1}{2}\sqrt{\frac{3}{2}}Z_{32}$
4	0	Z_{40}	$\frac{3}{8}Z_{40}-\frac{1}{4}\sqrt{5}Z_{42}+\frac{1}{8}\sqrt{35}Z_{44}$
4	1	$Z_{4\,-1}$	
4	-1	$-Z_{41}$	
4	2	$-Z_{42}$	$\frac{1}{4}\sqrt{5}Z_{40}-\frac{1}{2}Z_{42}-\frac{1}{4}\sqrt{7}Z_{44}$
4	-2	$-Z_{4\,-2}$	$-\frac{1}{2}\sqrt{\frac{1}{2}}Z_{4\,-1}+\frac{1}{2}\sqrt{\frac{7}{2}}Z_{4\,-3}$
4	3	$-Z_{4\,-3}$	
4	-3	Z_{43}	
4	4	Z_{44}	$\frac{1}{8}\sqrt{35}Z_{40}+\frac{1}{4}\sqrt{7}Z_{42}+\frac{1}{8}Z_{44}$
4	-4	$Z_{4\,-4}$	$-\frac{1}{2}\sqrt{\frac{7}{2}}Z_{4\,-1}-\frac{1}{2}\sqrt{\frac{1}{2}}Z_{4\,-3}$

Table 2.11. Effects of Group Generators on Z_{lm}'s. For $l=4$, only the results needed to determine the bases are listed.

a component of $-\frac{1}{2}|E\epsilon\rangle$ while $C_3^{(111)}|T_2\zeta\rangle$ has no component of $|T_2\zeta\rangle$. Thus we can assign $|E\epsilon\rangle = Z_{42}$ and $|T_2\zeta\rangle = Z_{4\,-2}$. This also leads to:

$$\begin{array}{rcccl} |E\theta\rangle & = & -2\sqrt{\frac{1}{3}}[C_3^{(111)}|E\epsilon\rangle + \frac{1}{2}|E\epsilon\rangle] & = & -\frac{1}{2}\sqrt{\frac{5}{3}}Z_{40} + \frac{1}{2}\sqrt{\frac{7}{3}}Z_{44}, \\ |T_2\xi\rangle & = & C_3^{(111)}|T_2\zeta\rangle & = & -\frac{1}{2}\sqrt{\frac{1}{2}}Z_{4\,-1} + \frac{1}{2}\sqrt{\frac{7}{2}}Z_{4\,-3}, \\ |T_2\eta\rangle & = & -C_4^z|T_2\xi\rangle & = & -\frac{1}{2}\sqrt{\frac{1}{2}}Z_{41} - \frac{1}{2}\sqrt{\frac{7}{2}}Z_{43}. \end{array}$$

We now note that $|A_1a_1\rangle$, $|T_1z\rangle$, and $|E\theta\rangle$ transform under C_4^z into themselves. Z_{40}, Z_{44}, and $Z_{4\,-4}$ have this property. Since $C_3^{(111)}Z_{4\,-4}$ contains terms not in the transforms of the other two, $|A_1a_1\rangle = aZ_{40} + bZ_{44}$. Since the Z_{42} term must cancel, we have $-\frac{1}{4}\sqrt{5}a + \frac{1}{4}\sqrt{7}b = 0$. Using $a^2 + b^2 = 1$ and the fact that a must be positive (since Z_{40} term of $C_3^{(111)}Z_{40}$ is positive), we get $|A_1a_1\rangle = \frac{1}{2}\sqrt{\frac{7}{3}}Z_{40} + \frac{1}{2}\sqrt{\frac{5}{3}}Z_{44}$. This leaves $|T_1z\rangle = Z_{4\,-4}$, and we find:

$$\begin{array}{rcccl} |T_1x\rangle & = & C_3^{(111)}|T_1z\rangle & = & -\frac{1}{2}\sqrt{\frac{7}{2}}Z_{4\,-1} - \frac{1}{2}\sqrt{\frac{1}{2}}Z_{4\,-3}, \\ |T_1y\rangle & = & C_4^z|T_1x\rangle & = & \frac{1}{2}\sqrt{\frac{7}{2}}Z_{4\,1} - \frac{1}{2}\sqrt{\frac{1}{2}}Z_{4\,3}. \end{array}$$

The basis functions are tabulated in Table 2.12 after extension to O_h.

We now consider the use of these basis functions to gain insight on the orthonormalization of the Z_{lm}'s. Let $|l\Gamma_i\beta_j\rangle$ be the j^{th} function of the basis for Γ_i using Z_{lm}'s of total angular momentum l, and let $I(l\Gamma_i\beta_j|l'\Gamma_{i'}\beta_{j'})$ be our numerical approximation to $\langle l\Gamma_i\beta_j|l'\Gamma_{i'}\beta_{j'}\rangle$. Basis functions with distinct symmetry are automatically orthogonal on our O_h grid as are any two functions of the same symmetry but with distinct coordinates of that symmetry. Thus, the orthogonality of $I(l\Gamma_i\beta_j|l'\Gamma_{i'}\beta_{j'})$ is ensured for $i \neq i'$ and/or $j \neq j'$. Further, the orthogonality of one particular $I(l\Gamma_i\beta_j|l'\Gamma_i\beta_j)$ ensures the orthogonality of all of them, e.g., setting $I(1T_{1u}z|3T_{1u}z) = 0$ in turn sets $I(1T_{1u}x|3T_{1u}x) = 0$ and $I(1T_{1u}y|3T_{1u}y) = 0$. Requiring the normality of one function of a basis set is sufficient to gain the normality of all functions in that basis. Thus, setting $I(1T_{1u}z|1T_{1u}z) = 1$ in turn sets $I(1T_{1u}x|1T_{1u}x) = 1$ and $I(1T_{1u}y|1T_{1u}y) = 1$. Thus, the conditions for exactly in-

l	Basis		
0	$\|A_{1g}a_1\rangle$	$=$	Z_{00}
1	$\|T_{1u}x\rangle$	$=$	Z_{11}
	$\|T_{1u}y\rangle$	$=$	$Z_{1\,-1}$
	$\|T_{1u}z\rangle$	$=$	Z_{10}
2	$\|E_g\theta\rangle$	$=$	Z_{20}
	$\|E_g\epsilon\rangle$	$=$	Z_{22}
	$\|T_{2g}\xi\rangle$	$=$	$Z_{2\,-1}$
	$\|T_{2g}\eta\rangle$	$=$	Z_{21}
	$\|T_{2g}\zeta\rangle$	$=$	$Z_{2\,-2}$
3	$\|A_{2u}a_2\rangle$	$=$	$Z_{3\,-2}$
	$\|T_{1u}x\rangle$	$=$	$-\frac{1}{2}\sqrt{\frac{3}{2}}Z_{31}+\frac{1}{2}\sqrt{\frac{5}{2}}Z_{33}$
	$\|T_{1u}y\rangle$	$=$	$-\frac{1}{2}\sqrt{\frac{3}{2}}Z_{3\,-1}-\frac{1}{2}\sqrt{\frac{5}{2}}Z_{3\,-3}$
	$\|T_{1u}z\rangle$	$=$	Z_{30}
	$\|T_{2u}\xi\rangle$	$=$	$-\frac{1}{2}\sqrt{\frac{5}{2}}Z_{31}-\frac{1}{2}\sqrt{\frac{3}{2}}Z_{33}$
	$\|T_{2u}\eta\rangle$	$=$	$-\frac{1}{2}\sqrt{\frac{5}{2}}Z_{3\,-1}+\frac{1}{2}\sqrt{\frac{3}{2}}Z_{3\,-3}$
	$\|T_{2u}\zeta\rangle$	$=$	Z_{32}
4	$\|A_{1g}a_1\rangle$	$=$	$\frac{1}{2}\sqrt{\frac{7}{3}}Z_{40}+\frac{1}{2}\sqrt{\frac{5}{3}}Z_{44}$
	$\|E_g\theta\rangle$	$=$	$-\frac{1}{2}\sqrt{\frac{5}{3}}Z_{40}+\frac{1}{2}\sqrt{\frac{7}{3}}Z_{44}$
	$\|E_g\epsilon\rangle$	$=$	Z_{42}
	$\|T_{1g}x\rangle$	$=$	$-\frac{1}{2}\sqrt{\frac{7}{2}}Z_{4\,-1}-\frac{1}{2}\sqrt{\frac{1}{2}}Z_{4\,-3}$
	$\|T_{1g}y\rangle$	$=$	$\frac{1}{2}\sqrt{\frac{7}{2}}Z_{4\,1}-\frac{1}{2}\sqrt{\frac{1}{2}}Z_{4\,3}$
	$\|T_{1g}z\rangle$	$=$	$Z_{4\,-4}$
	$\|T_{2g}\xi\rangle$	$=$	$-\frac{1}{2}\sqrt{\frac{1}{2}}Z_{4\,-1}+\frac{1}{2}\sqrt{\frac{7}{2}}Z_{4\,-3}$
	$\|T_{2g}\eta\rangle$	$=$	$-\frac{1}{2}\sqrt{\frac{1}{2}}Z_{41}-\frac{1}{2}\sqrt{\frac{7}{2}}Z_{43}$
	$\|T_{2g}\zeta\rangle$	$=$	$Z_{4\,-2}$

Table 2.12. Z_{lm} Bases for O_h.

l	Normality	Orthogonality
0	$I(0A_{1g}a_1\|0A_{1g}a_1) = 1$	
1	$I(1T_{1u}z\|1T_{1u}z) = 1$	
2	$I(2E_g\epsilon\|2E_g\epsilon) = 1$ $I(2T_{2g}\zeta\|2T_{2g}\zeta) = 1$	
3	$I(3A_{2u}a_2\|3A_{2u}a_2) = 1$ $I(3T_{1u}z\|3T_{1u}z) = 1$ $I(3T_{2u}\zeta\|3T_{2u}\zeta) = 1$	$I(1T_{1u}z\|3T_{1u}z) = 0$
4	$I(4A_{1g}a_1\|4A_{1g}a_1) = 1$ $I(4E_g\epsilon\|4E_g\epsilon) = 1$ $I(4T_{1g}z\|4T_{1g}z) = 1$ $I(4T_{2g}\zeta\|4T_{2g}\zeta) = 1$	$I(0A_{1g}a_1\|4A_{1g}a_1) = 0$ $I(2E_g\epsilon\|4E_g\epsilon) = 0$ $I(2T_{2g}\zeta\|4T_{2g}\zeta) = 0$

Table 2.13. Conditions for Exact Numerical Orthonormality of Z_{lm}'s.

tegrating the orthonormality of the spherical harmonics through a given l are only those listed in Table 2.13.

III.C. U_3 Integration

We now combine the results of the previous subsection with the general mesh presented in Section II to create new quadrature schemes. We then compare the new Gaussian and "nearest area" quadratures with some commonly used quadratures. We consider numerical integration over U_3, the surface of the unit sphere in 3-D Euclidean space, but the resulting spherical grids can be combined with a radial grid for application to volumes. In general, we consider the approximation [8]:

$$I(f) \equiv \iint f(x,y,z)d\Omega \approx \sum_{i=1}^{N} w_i f(x_i,y_i,z_i). \tag{2.28}$$

We limit ourselves to the use of positive weights w_i and to sample points (x_i, y_i, z_i) which lie on the unit sphere, i.e., $x_i^2 + y_i^2 + z_i^2 = 1$.

It seems desirable to have a quadrature scheme which exactly integrates

$$I_{lml'm'} \equiv \iint Z_{lm} Z_{l'm'} d\Omega = \delta_{ll'}\delta_{mm'} \tag{2.29}$$

for all $l = 0, \ldots, n$ and $m = -l, \ldots, l$ for some degree n. For the moment, let us consider an integration scheme based on only one of the independent points listed in Table 2.2. The number of points of this scheme is d_i, the degeneracy of the point, and the weights w_j for these points must all be equal. For a type E base point with coordinates (a, b, c), the integration points are $(\pm a, \pm b, \pm c)$, $(\pm a, \pm c, \pm b)$, $(\pm b, \pm a, \pm c)$, $(\pm b, \pm c, \pm a)$, $(\pm c, \pm a, \pm b)$, and $(\pm c, \pm b, \pm a)$. The orthonormality constraints $I(l\Gamma_i\beta_j|l'\Gamma_i\beta_j)$ are conveniently listed in terms of:

$$\sigma_1 \equiv 1, \tag{2.30}$$

$$\sigma_2 \equiv a^2b^2 + b^2c^2 + a^2c^2, \tag{2.31}$$

$$\sigma_3 \equiv a^2b^2c^2. \tag{2.32}$$

Applying the bases of Table 2.12 and some algebra, the values of the conditions of Table 2.13 are listed in Table 2.14. Thus, both of the normality conditions $I(0A_{1g}a_1|0A_{1g}a_1) = 1$ and $I(1T_{1u}z|1T_{1u}z) = 1$ have reduced to just $\sum_{j=1}^{d_i} w_j(\frac{1}{4\pi}) = 1$ or simply $w_j = \frac{4\pi}{d_i}$. Thus, any combination of the points in Table 2.2 will integrate the orthonormality conditions of the Z_{lm}'s through $l = 1$ if $\sum_j w_j = 4\pi$. For a given independent point (a_i, b_i, c_i), we define $I_i(l\Gamma\beta|l'\Gamma\beta)$ to be the right hand side of Table 2.14 with $a = a_i$, $b = b_i$, and $c = c_i$.

We now return to our general integration scheme of a combination of various point types. Letting $|l\Gamma\beta\rangle_j$ be $|l\Gamma\beta\rangle$ evaluated at point j and the sum over j be over all points in our grid, we can now reduce the expression

$$I(l\Gamma\beta|l'\Gamma\beta) = \sum_j w_j(\langle l\Gamma\beta|_j)(l'\Gamma\beta\rangle_j) \tag{2.33}$$

l	Normality	Orthogonality
0	$I(0A_{1g}a_1\|0A_{1g}a_1) = 1$	
1	$I(1T_{1u}z\|1T_{1u}z) = 1$	
2	$I(2E_g\epsilon\|2E_g\epsilon) = \frac{5}{2}(1-3\sigma_2)$ $I(2T_{2g}\zeta\|2T_{2g}\zeta) = 5\sigma_2$	
3	$I(3A_{2u}a_2\|3A_{2u}a_2) = 105\sigma_3$ $I(3T_{1u}z\|3T_{1u}z) = \frac{7}{12}(4-15\sigma_2+75\sigma_3)$ $I(3T_{2u}\zeta\|3T_{2u}\zeta) = \frac{35}{4}(\sigma_2-9\sigma_3)$	$I(1T_{1u}z\|3T_{1u}z) = \sqrt{\frac{7}{3}}(1-5\sigma_2)$
4	$I(4A_{1g}a_1\|4A_{1g}a_1) = \frac{21}{4}(1-10\sigma_2+25\sigma_2^2)$ $I(4E_g\epsilon\|4E_g\epsilon) = \frac{15}{8}(1-10\sigma_2-84\sigma_3+49\sigma_2^2)$ $I(4T_{1g}z\|4T_{1g}z) = \frac{105}{4}(\sigma_2+3\sigma_3-4\sigma_2^2)$ $I(4T_{2g}\zeta\|4T_{2g}\zeta) = \frac{15}{4}(\sigma_2+7\sigma_3)$	$I(0A_{1g}a_1\|4A_{1g}a_1) = \frac{\sqrt{21}}{2}(1-5\sigma_2)$ $I(2E_g\epsilon\|4E_g\epsilon) =$ $\frac{15\sqrt{3}}{8}(-2+13\sigma_2-63\sigma_3)$ $I(2T_{2g}\zeta\|4T_{2g}\zeta) = \frac{5\sqrt{3}}{2}(21\sigma_3-\sigma_2)$

Table 2.14. Values of Orthonormality Conditions for a Point (a, b, c).

to

$$I(l\Gamma\beta|l'\Gamma\beta) = \sum_i \frac{d_i w_i}{4\pi} I_i(l\Gamma\beta|l'\Gamma\beta), \tag{2.34}$$

where the sum over i is just over the independent points (a_i, b_i, c_i). The condition for orthonormalization of s and p functions then becomes $\sum_i d_i w_i = 4\pi$, i.e., the exact integration of σ_1. Including d functions also requires exact integration of σ_2, including f functions also requires σ_3, including g functions also requires σ_2^2, etc.

This somewhat surprising result is based on one fact and one lemma. Since the product of two spherical harmonics with l_1 and l_2 can be expanded in terms of a sum of spherical harmonics up to $l_1 + l_2$, the orthonormality conditions through l are fulfilled if the quadrature exactly integrates the spherical harmonics, i.e., $\iint Z_{lm} d\Omega$ is exact, through order $2l$. The lemma given by Lebedev [9] is that any polynomial invariant under octahedral symmetry can be expressed on the unit sphere in terms of a polynomial in σ_2 and σ_3. Although as a lemma this is not proven, we have shown support for this in the above derivations. Thus, we now consider the exact

integration of the spherical harmonics themselves through a given order $l \leq n$ to be the basis of a quadrature of degree n. Existing quadratures fulfilling this description include those of Lebedev [9-11] and Glatzmaier [12]. The Glatzmaier scheme has $(l+1)^2$ point grids based on latitude lines and uses the weights to solve the quadrature conditions and thus consists of Gauss quadratures. Lebedev presents both Gauss (unequal weights) and Chebyshev (equal weights) quadratures which have octahedral symmetry.

It is our working hypothesis that by using nearest point weights we can obtain quadratures which are more accurate than existing ones for spherical harmonics of higher degree than that of the quadrature. The number of conditions which must be satisfied for a degree n quadrature determines the number of independent coordinates and thus the mesh level we must use. For a degree n quadrature integrating all spherical harmonics with $l \leq n$, we have $n/2$ conditions (not including that the weights w_i sum to 4π) based on the above lemma and our supporting derivations. For Chebyshev or nearest-area quadratures this tells us the number of required independent coordinates, while for Gauss quadratures one plus this (so that the w_i sum to 4π) should be the number of independent points plus the number of independent coordinates. Table 2.4 gives the number of independent coordinates for various levels of our general octahedral mesh. A level 1 spherical mesh can integrate the spherical harmonics themselves through $l = 3$, i.e., f functions, and the orthonormality conditions (ONC) through $l = 1$, i.e, p functions. A level 2 mesh still has no independent coordinates, so we can do no better with a Chebyshev or nearest-area quadrature. A level 3 mesh has one independent coordinate so we can integrate the Z_{lm}'s through $l = 5$ and the ONC's through $l = 2$. A level 4 mesh has two independent coordinates and can integrate Z_{lm}'s through $l = 7$ and ONC's through $l = 3$. A level 5 mesh adds two independent coordinates for a total of four

and can thus integrate Z_{lm}'s through $l = 11$ and the ONC's through $l = 5$. We were able to obtain both Chebyshev and nearest-area quadratures for each of these levels. They are best presented in terms of integrating higher order spherical harmonics for which we next present a numerical measure.

Since exact integration of the spherical harmonics only depends on exact integration of polynomials in σ_2 and σ_3, we can limit our concern to the integrals

$$S_{nm} \equiv \iint \sigma_2^n \sigma_3^m d\Omega. \tag{2.35}$$

Our numerical approximation to these based on independent points i with weights w_i and degeneracies d_i is

$$I_{nm} \equiv \sum_i w_i d_i [\sigma_2^n \sigma_3^m]|_i. \tag{2.36}$$

A measure of the deviation s from exact integration of the spherical harmonics through $l_{\max}$ is given by

$$s^2 = \frac{1}{N_p} \sum_{4n+6m \le l_{\max}} \left(1 - \frac{I_{nm}}{S_{nm}}\right)^2. \tag{2.37}$$

Figure 2.8 compares our degree 7 (level 4) Chebyshev and nearest area quadratures with the degree 9 Gauss quadratures of Lebedev, while Figure 2.9 compares various degree 11 quadratures. These plots demonstrate the stability of using nearest area quadratures versus Chebyshev or Gauss quadratures. The degree 7 quadrature shows at least a twofold improvement in accuracy for $l = 11$, and stable calculation within an error of six percent for large l. Since we weight the integration of polynomials in σ_2 and σ_3 equally while orthonormalization produces more lower order polynomials (see Table 2.14), the accuracy of integrating the orthonormality conditions is actually greater. The accuracy is even higher when one considers all of the orthogonality conditions which are exactly zero by reasons of symmetry.

Numerical quadratures find application in almost all scientific fields, and quadratures involving spherical surfaces are prevalent in meteorology and chemistry among

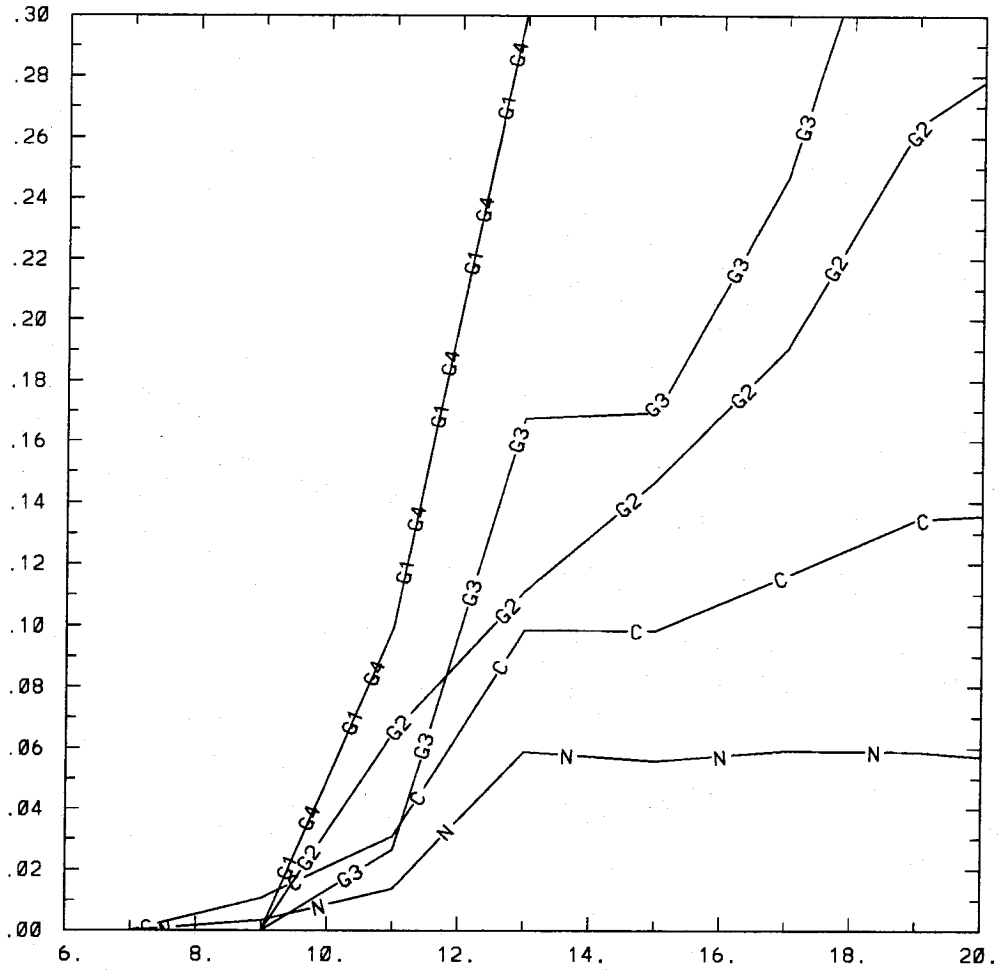


Figure 2.8. New Degree 7 Quadratures. N is a new degree 7 nearest area quadrature, C is a new degree 7 Chebyshev quadrature, and G1, G2, G3, and G4 are four degree 9 quadratures of Lebedev [9].

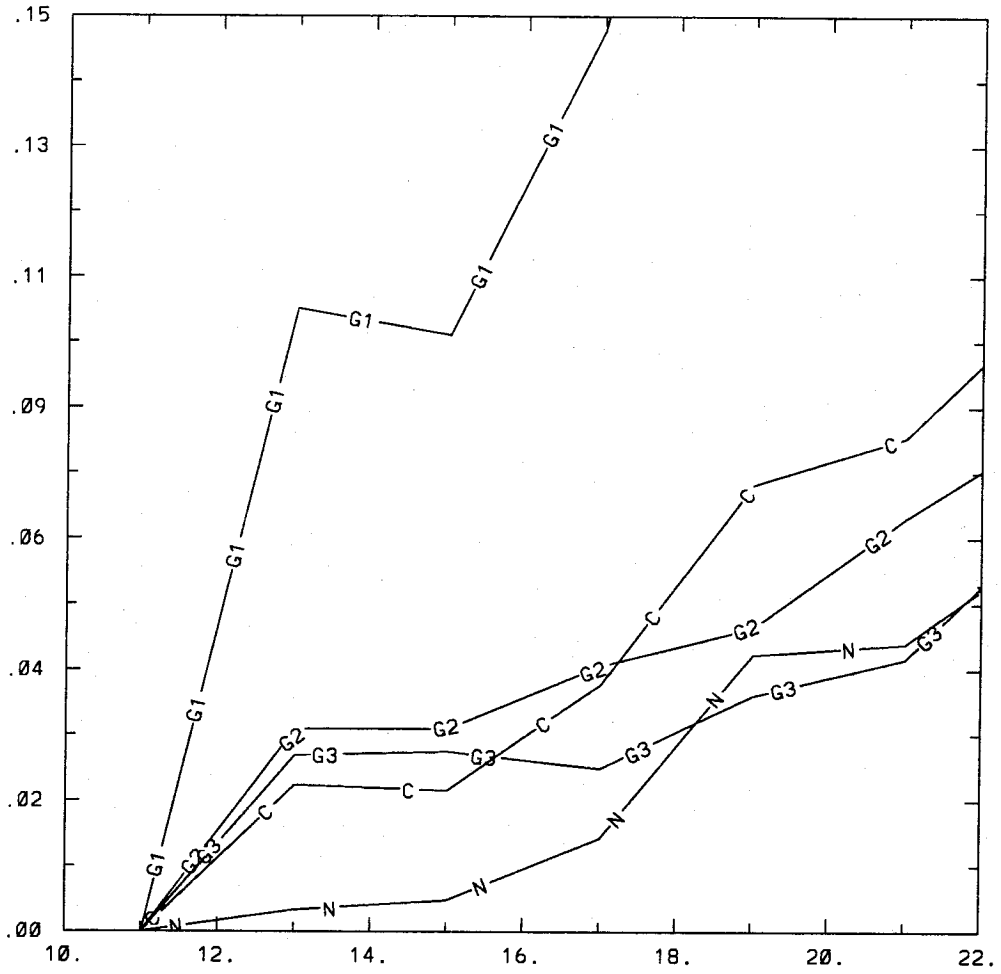


Figure 2.9. New Degree 11 Quadratures. N is a new degree 11 nearest area quadrature, C is a degree 11 Chebyshev quadrature, and G1, G2, and G3 are three degree 11 quadratures with all but the first from Lebedev [9].

other fields. Quadratures generally fall into one of two categories: Chebyshev with equally weighted points and Gauss with unequally weighted points. In constructing quadratures that exactly integrate a given number of polynomials, the weights are often used as independent variables to produce degree n quadratures using the fewest number of points. However, this is shown above to mean that higher degree polynomials are less accurately integrated than Chebyshev quadratures. We have introduced a new general quadrature scheme of basing the weights on the Wigner-Seitz or Voronoi volumes (areas in two dimensions). Basing a point's weight on the volume it represents results in greater accuracy for higher order polynomials. We have applied these ideas to spherical surfaces based on the grids presented in Section II. If it is definitely known that the expansion of the integrand in terms of spherical harmonics rapidly declines with larger l, then the Lebedev Gauss quadratures are probably the most efficient. However, if the expansion includes an unknown amount of higher l terms as is often the case, our nearest area quadratures are more efficient.

IV. Interactive Atomic Displays

Graphics workstations have allowed physical scientists to use computer graphics as a research tool instead of just a method of displaying their results. This is the fundamental basis of scientific visualization, a buzzword encompassing a large set of techniques that increase scientific understanding through graphical displays. Many techniques useful to the chemist are in use or under development, but it is the display of large molecules that is his foremost tool. To be useful as a research tool, molecular displays have to be efficient enough so that a molecule can be viewed from various angles and perspectives. Graphics hardware has advanced to the point where up to a million or so polygons can be rendered per second. However, reliance

on the graphics hardware to do all of the work greatly reduces efficiency and thus productivity. Also, more modest hardware is often the only available hardware, and then it becomes essential to have efficient molecular displays. Using a good understanding of computer graphics, it is possible to increase rendering performance as high as tenfold by not relying on the graphics hardware. Software coordinate transformations and software clipping can greatly decrease the amount of work done in the graphics pipeline. Because most molecular displays are based on spheres or sphere segments, the use of light models to precalculate colors removes lighting calculations from the graphics hardware and reduces them to a table lookup. The first two subsections present the methods and resulting savings of implementing these techniques in software. The use of color can aid our understanding of atomic interactions, but our software must be flexible enough to adapt to the needs of various molecular systems. The third subsection looks at how the complex reconstuction of semiconductor surfaces can be understood through the appropriate molecular display.

IV.A. Matrix Transformations and Clipping

"Real time" is an often used yet often vague buzzword concerning computing. For the purposes of this section, we define it as being about to see a new perspective of a molecule with at most a very short delay after entering some command (such as turning a dial or moving a mouse) to instigate it. Fast hardware frame buffers have taken much of the work of polygon rendering and hidden surface removal away from the programmer [13]. A significant amount of the work done in the graphics pipeline is the matrix transformation of vectors to allow for various orientations of the display. For atoms, however, the spherical mesh comprising the atomic sphere is a constant which is merely translated and scaled to account for the atom position

and size, respectively. While software running on the main processor is not nearly as fast as specialized graphics hardware for doing matrix transformations, there are great savings in doing the atomic transformation once instead of for each point in an atomic mesh. Once this is done, an additional savings is realized in that gross (atomic-level) clipping can be done in software saving the amount of fine (polygon-level) clipping required of the graphics hardware. The implementation of these techniques and the resulting savings are presented in this subsection.

A software implementation of matrix transformations requires a solid understanding of the order of transformations. As orthogonal projections are well suited for molecules and the most efficient in that surface normals need not be rescaled, we limit our discussion to orthogonal or nonperspective transformations. A separation is usually made between the viewing transformations related to viewport and eye positions and modeling transformations related to the rendered object itself. The least work is required of the renderer if it does only the orthogonal viewport transformation. Zooms and viewport translations are easily included in this. We place the burden of all sphere translations and scalings on the software. This is rather easy to implement. Further, we require that the software handle all rotations input by the user by the use of dials, mouse, or menus. As this requires various polling of devices and/or menu setup, it is more difficult. Our scheme polls the appropriate devices and calls low level graphics routines for rotation as would normally be done. However, before redrawing the screen, the modeling transformation matrix (MTM) is retrieved into software and then set to unity. Then by allowing the MTM to operate on atomic positions in software, the graphics hardware must only perform the viewing transformations and not the modeling transformations on all of the mesh points. To test these savings, a generation 7 β-alanine starburst dendrimer [14] containing 3814 atoms is rotated using several transformation schemes. Timings were done on

	4D/380VGX	4D/25TG
Hardware rotations and color	0.9062	7.6646
Software rotations and hardware color	0.4746	4.0960
Software rotations and color	0.4257	3.2499

Table 2.15. Rendering Times. Elapsed times per screen update in seconds are given for 3814 atoms using a level 3 mesh for two Silicon Graphics workstations. Hardware Gouraud shading is used for all. The first two lines include hardware lighting calculations, while the last uses precalculated software lighting calculations as discussed in Subsection IV.B.

two Silicon Graphics computers: a high end graphics workstation, the 4D/380VGX rated at one million unlit, unshaded, meshed triangles per second and 180 thousand unlit, Gouraud shaded, polygons per second and a low to mid range workstation, the 4D/25TG rated at 57 thousand triangles and 26 thousand polygons of the same type. The results are given in Table 2.15. For both machines, using hardware rotations is about 1.9 times slower than software rotations. The major part of these savings comes from not being required to render both halves of the sphere in the former instance. This gives evidence that the drawing speed is primarily rendering bound as should be the case for efficient displays.

Including transformations in software also allows other speed gains and the implementation of several new features. The additional speed gains are obtained mainly from clipping. Just as transforming atomic coordinates instead of mesh coordinates saves time, clipping entire atoms instead of the individual triangles in the mesh can be a great gain. The latter is seen only when a portion of the atoms fall beyond the six clipping planes of the viewport, but it is often the case that one wishes to study a smaller segment of a large molecule. A variety of software clipping has been implemented in our software. This includes no software clipping which lets the hardware renderer do the work, software clipping of atoms totally outside of the viewport which reduces the amount sent to the hardware renderer, and software clipping of

Hardware	Software	4D/380VGX	4D/25TG
MT[a], clipping, lighting		169.97	972.00
clipping, lighting	MT	90.32	511.57
lighting	MT, clipping	43.36	192.77
	MT, clipping, lighting	39.49	148.12
	near clipping[b]	28.40[c]	124.08
	near and far clipping[d]	24.57	96.56

Table 2.16. Pan Times. Elapsed times for 140 0.5Å steps through a 3814 atom molecule using a 20Å slab.

[a]MT stands for modeling transformations.

[b]MT, clipping, and lighting in software plus the clipping of atoms only partially within the near clipping plane.

[c]The last two 4D/380VGX times were done while the machine had a heavier load than above so times may be higher in comparison.

[d]MT, clipping, and lighting in software plus the clipping of atoms only partially within both near and far clipping planes.

atoms partially outside of the viewport (or specified clipping planes) which can make the view more understandable. A test case is given by panning from front to back on the above starburst dendrimer. This enables the channels or open spaces within the dendrimer to be better understood. The timing results are given in Table 2.16. This shows that doing hardware clipping is a factor of 2.1 and 2.7 times slower than software clipping for this test. Further gains are made by removing partial atoms from the picture which can actually clarify the picture. In all, up to a tenfold gain is possible by switching as much as possible into software, mainly by treating the atom as a unit instead of each polygon as a unit.

The best demonstration of the usefulness of panning through a molecule is to do so in real time or in a video. Figure 2.10 at least gives insight into the sterics of the center of a β-alanine starburst dendrimer by showing a slab of 20Å near the center. The various areas of congestion and of open volumes are easily seen. The

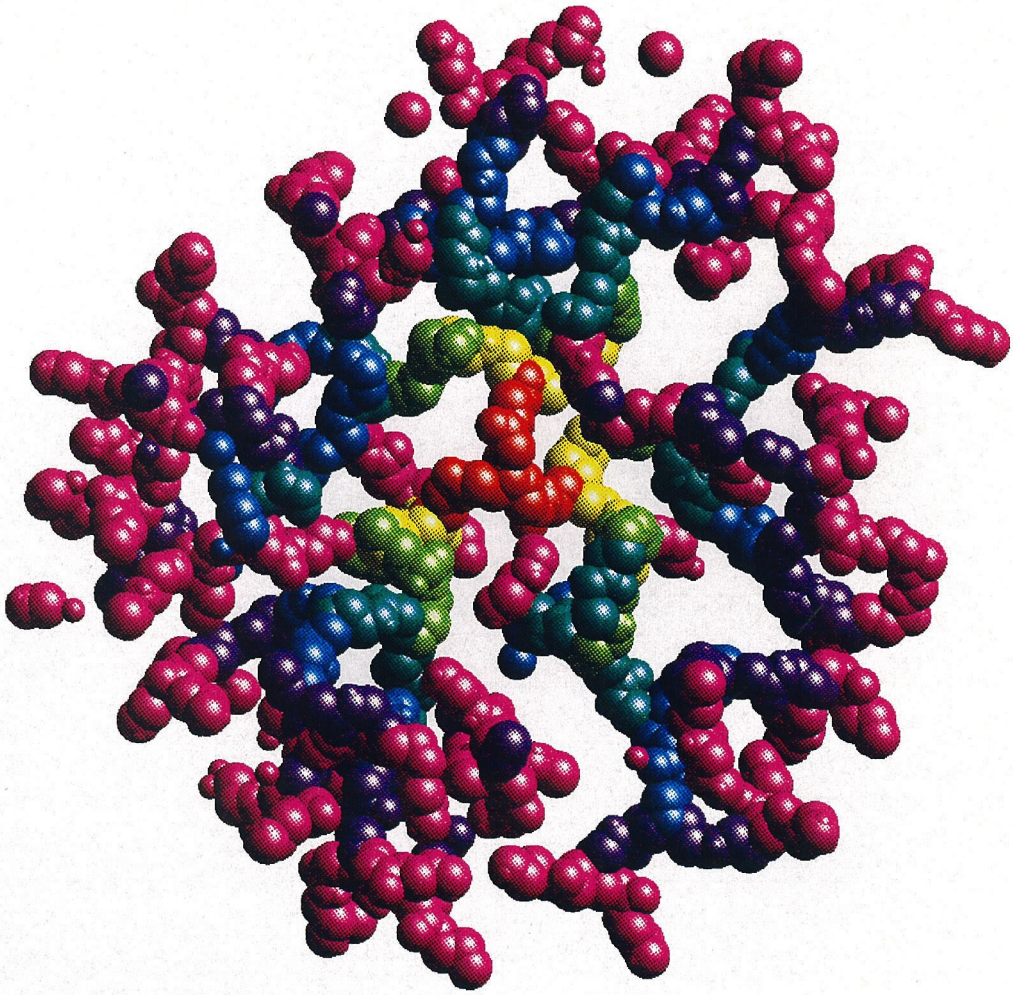


Figure 2.10. Starburst Dendrimer Slab. A 20Å slice of a generation 7, β-alanine starburst dendrimer is shown. Polymer units are colored from red to magenta as one progresses from the core through the higher generations.

use of color to distinguish the polymer units is an additional feature of our software which promotes understanding of the connectivity and changes in steric effects as the generation increases. Further use of color particular to a given application is discussed in Subsection IV.C after a discussion of using a software light model to precalculate colors in the next section.

IV.B. Light Models

Computer graphics models of light have grown increasingly complex. In ray-traced images, the color at a given point depends upon surface properties at that point and on other points which effect how much light falls on the first point. With such models, effects such as reflection, translucence, and shadows are obtainable. However, ray tracing is computationally intensive since many rays must be "traced" from their source through many reflections and transmissions. For the purpose of molecular displays, a much simpler model is possible. The use of polygon rendering techniques yield visual information about an object's color, shininess, texture, and shape while reducing color dependence to the given point only. A general such lighting model is

$$I = AD + \sum_i I_i[(\vec{N} \cdot \vec{L}_i)D + (\vec{R} \cdot \vec{L}_i)^{E_s}S]/r_i^2, \tag{2.38}$$

where I is the color intensity at the vertex of the polygon, A is the ambient light intensity, D is the diffuse color of the vertex, I_i is the intensity of light source i, S is the specular color at the vextex, E_s is the specular exponent of the material, and r_i is the distance from the vertex to light source i. Actually, there are three such equations, one each for red, green, and blue, with each of the color intensities having separate red, green, and blue components. The vector $\vec{N}$ is the surface normal at the vertex, $\vec{L}_i$ is the normalized direction from the vextex to light source i, and $\vec{R}$ is the normalized reflection through the normal of the eye vector which runs from the

vextex to the viewer. This model gives quite realistic "plastic" surfaces.

Two further approximations are useful in reducing the time of the lighting calculation. The first is the "sunlight" approximation which lets the dimensions of the object be small with respect to the distance from the object to the light source. This allows us to replace I_i/r_i^2 with the constant I_i in the above equation by considering light to fall with equal intensity on all surfaces. The second is the "telescope" approximation. This allows the eye vector to be considered a constant. Combined, these equations reduce the lighting calculation considerably.

So far, only standard graphics techniques have been discussed. If the modeling transformations are done in software as given in the previous subsection, the normals of a sphere's surface do not change with respect to the eye vector. Thus for a given color sphere of any size, the lighting calculation needs be done only once. In precomputing colors in such a manner, an additional 10% to 20% in elapsed time is saved as seen by comparing the last two lines of Table 2.15. At this point, the calculation is limited primarily by the polygon scan-line conversion which is the best we can expect to obtain.

With well under a second required to update a screen, we may consider the viewing of 4000 atoms to be marginally in "real time" on a high-end graphics workstation. This was only obtained by good understanding of graphics techniques and lighting models instead of relying on the graphics hardware, as fast as it is, to do all of the work. For low to mid range workstations, a thousand or less atoms now can be rotated in real time. This work is meant to give the chemist a better tool with which to work, and next we consider how to use such a tool with an effective use of color to better understand semiconductor surfaces.

IV.C. Reconstructed Silicon Surfaces

One field in which geometric considerations are important but difficult to understand is the reconstruction of semiconductor surfaces. Since Schlier and Farnsworth [15] first observed the 7x7 reconstruction of Si(111) in a LEED experiment, a large number of models were proposed and debated until a stacking fault model first developed by McRae [16] and then extended by others [17] came into general acceptance. The visualization of these models was often limited to circle drawings with larger circles extending farther from the page. The nature of the three-dimensional interactions were difficult to perceive. The use of color to distinguish atomic layers is most useful in this case. Instead of requiring the specification of separate atom types for each atomic layer, a more useful and general approach is to select color based on a coordinate perpendicular to the surface. In instituting such a method in our visualization software, we have additionally allowed certain atom types to retain a distinct color. Two pictures of this surface are presented. Figure 2.11 shows the entire surface at an angle. The stacking fault is easily seen as the channels of green crossing the surface. The lower layer atom with a dangling bond is just visible in the center and more easily seen along the front edge. Figure 2.12 shows a second view which makes use of entire atom clipping in software in order to better visualize the dimer-based stacking fault.

This section has addressed several issues of note in visualizing chemical systems. The overall theme is that it is best to make molecular display software a partner to the graphics hardware instead of relying on the latter. Much faster displays result from working with the atom as a unit as far as transformations and clipping. This also increases user control to allow specialized clipping and color schemes. We have implemented several color options which were applied to polymers and surfaces but

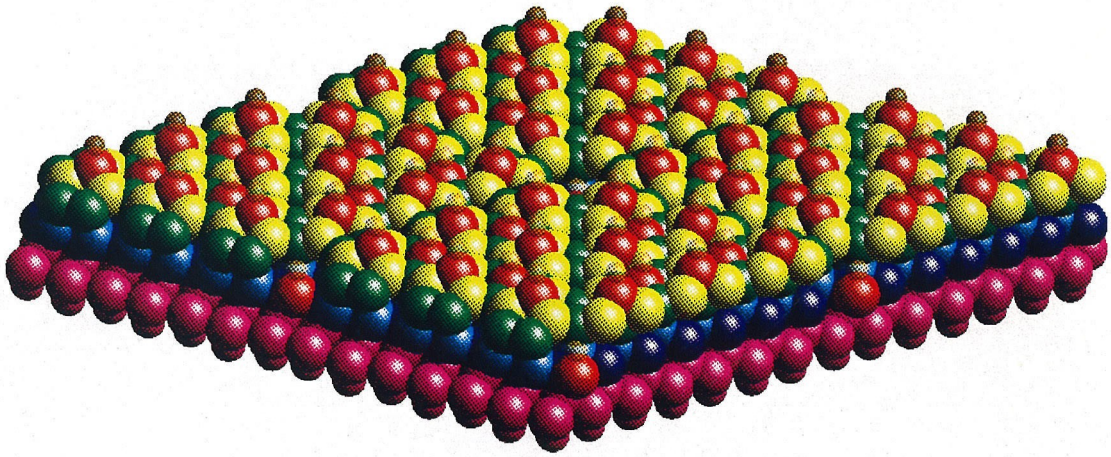


Figure 2.11. Si(111) 7x7 Surface. Green layer contains the dimer-based stacking fault. Red silicon atoms have dangling bonds represented by small, transparent orange spheres.

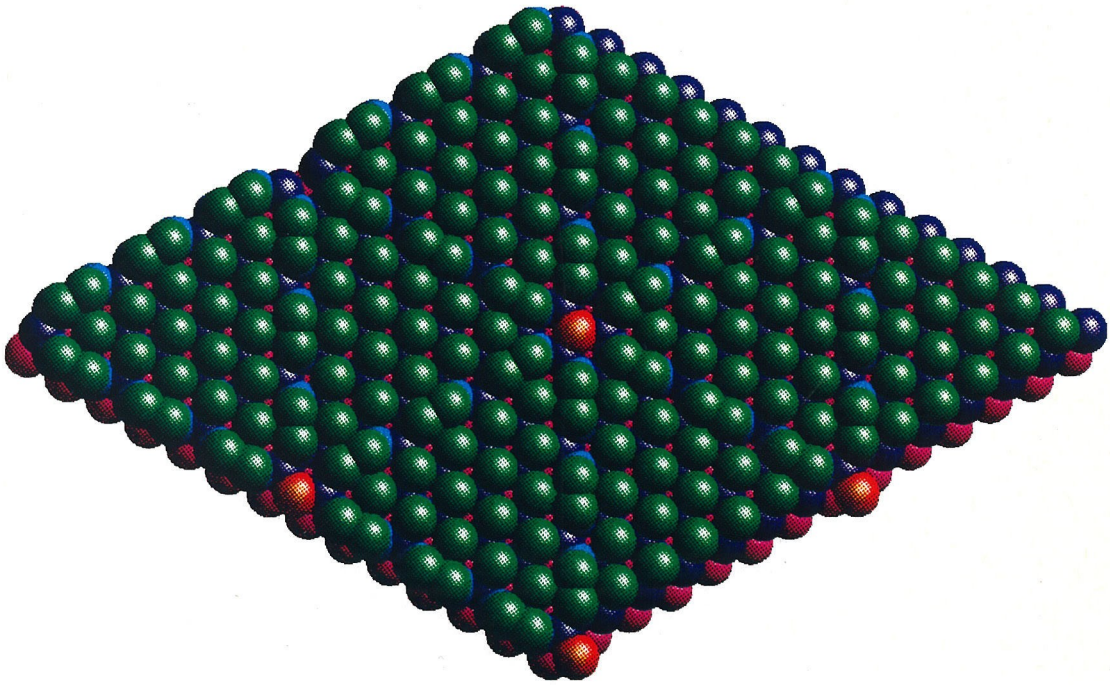


Figure 2.12. Si(111) 7x7 Stacking Fault. Dimers causing the incomplete upper layers are seen radiating from the center.

are flexible enough to handle many specialized needs. Overall, this work should reduce the time and effort required of the chemist to investigate physical properties of large molecules and crystals.

V. Molecular Surfaces

Various definitions of molecular surfaces have application in finding various physical and chemical properties of molecules. Examples of applications include analyzing the topology of large polymers [14], calculating electrostatic potentials [18], surface curvatures [19], or atomic overlaps [20] for analyzing possible protein-protein or protein-ligand interactions, calculating the solvation energies of proteins [21], and determining the pore structure of zeolite catalysts [22]. Examples of surface definitions (all broadly classed as molecular surfaces) include the van der Waals surface, the solvent-accessible surface of Lee and Richards [23], and the molecular surface of Richards [24]. Algorithmic implementation of the van der Waals and solvent-accessible surfaces are equivalent and two implementations are a longitude-lattidude based method [25] and one based on the pentakisdodecahedron [26]. Various representations of the Richards' molecular surface include grid [27], dot [28,29], cube [30,31], and analytical [32] methods.

While Section IV made use of atomic information to speed displaying molecules, here we make use of graphical clipping algorithms to speed calculating surface area and related properties. We deal principally with the accessible surface definition as it is the most adaptable to efficient calculation. Any analytical calculation which involves finding the intersections of all atoms is computationally too expensive. Methods based on polygons are also more complex than need be as the decision to keep a polygon involves looking at each of its vertices. We start with our general mesh of

Section II with the nearest (point) area weighting scheme of Subsection II.C. The basic problem is then to remove all of those points which fall within other atomic spheres. The surface area is then the sum of the nearest areas of the remaining points. We look at an efficient method of implementing this method in terms of clipping and apply methods of grouping surface areas in this section.

V.A. Clipping-Based Surface Area Calculation

We already discussed in the introduction that we believe that a point-based method involving the nearest Wigner-Seitz area of each point is likely to be the most efficient. There are still other decisions in designing an algorithm. The standard method of removing those mesh points of atom A which fall within atom B is to find the distance between each point of atom A and the center of atom B and compare this to the atom B's radius. Actually, comparing distances squared is, of course, more efficient. An alternative is to find the plane which defines the intersection of the two spheres. Then the points falling on the far side of the plane are clipped, i.e., removed from the active point list. This method is easily extended to find the contact part of the molecular surface of Richards. The details of the method are presented first, followed by comparison with a standard program for finding surface areas.

Finding the plane of intersection of two spheres is illustrated in Figure 2.13. The cut of atom i removed from overlap with atom j is thus defined by r_{cut} and a normal vector from r_i to r_j with

$$r_{\text{cut}} = \frac{r_i^2 + r_{ij}^2 - r_j^2}{2r_{ij}}. \tag{2.39}$$

A complete description of the van der Waal's or accessible surface is given by a list of each atom's position and radius plus the removed cuts in terms of the cut radius and normal. Both positive and negative r_{cut} values are possible with less than half

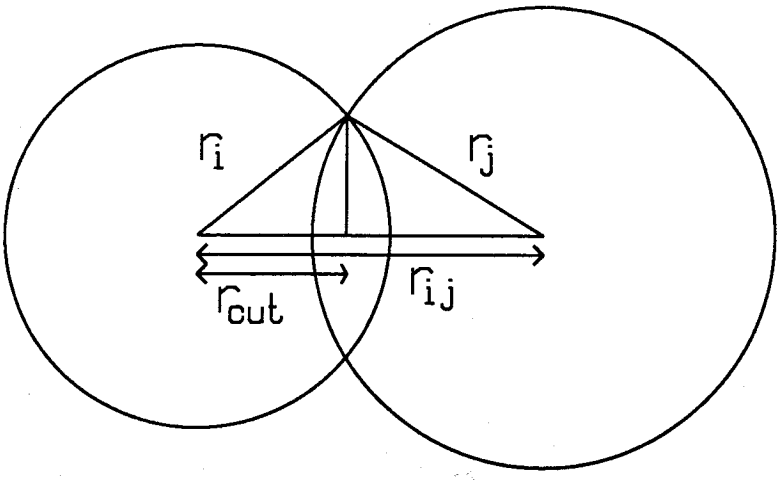


Figure 2.13. Intersection of Two Spheres.

of the sphere removed for $r_{\text{cut}} > 0$, exactly half for $r_{\text{cut}} = 0$, and more than half for $r_{\text{cut}} < 0$. It is possible to use this information to find the surface area analytically, however this would not be efficient due to the complexity of possible intersections and the floating point operations involved. Our grid algorithm is presented next.

For each atom, we must check against all other atoms to see if they overlap. It is important to first check whether atoms overlap at all before seeing if the points on one atom fall within another. This reduces the order of the problem from N^2M to N^2 for N atoms and M mesh points. The most computationally intensive work is then determining whether atoms overlap. A cube method where atoms are bucket sorted in three-dimensions can reduce the number of atom-atom checking but increases storage and programming, may or may not increase performance given the required sorting and looping, and will vary in performance with the characteristics of the system. However, there is another method suggested by clipping algorithms which can aid immensely. Before a distance calculation is made between a pair of atoms, it is possible to trivially reject many pairs by comparing distances in one dimension

Machine	GEPOL	Low Memory	High Memory
SGI 4D/380	87.8	10.9/14.1	7.6/11.2
FPS 522	178.8	30.6/37.7	18.8/25.9
Alliant FX/80-8	222.7	54.8/65.5	31.3/42.1

Table 2.17. Time to Find van der Waals Surface Area. Times are in seconds. GEPOL is Reference [33], and low and high memory are our programs. When two times are given, the first does not include setup which is done only once if surface area is repeatedly calculated during a simulation while second is the total time. GEPOL run had NDIV=2 for 96 points while our runs had $L = 5$ for 102 points.

only. For each of x, y, and z, the difference in one coordinate value is compared against the sum of the two atomic radii. Introducing this method adds virtually no program or storage overhead, and the program execution is cut by more than half for large systems.

We implemented two versions of this algorithm. One uses little more memory than that required for the atoms and one set of mesh points. The second can theorectically cut execution time by a factor of two but requires storing a set of logical variables the size of the mesh for each atom. The inner loop of the latter method loops over only those atoms not yet encountered in the outer loop and clips the points of both atoms of any pairs as they are found. The timing results of these two methods are compared with the GEPOL87 program of Pascual-Ahuir and E. Silla [33] in Table 2.17. It is seen that the speed increase of our method is almost an order of magnitude.

The accuracy of both of these calculations is better than $\pm 0.1\%$. Higher accuracy comes at the expense of many more points which is quite costly. The number of points for the GEPOL method as given in Reference [26] is 32, 96, 336, 1296, and 5136 for NDIV=1 to 5. Our mesh builds more gradually with $4L^2 + 2$ points for a given level L. However, more accurate calculations may be best obtained by extending the clipping analogy instead of adding more points. Since our method

already finds the appropriate clipping planes, it would be possible to clip the actual triangles which form the spherical mesh. This method of obtaining accurate surface areas may be implemented with a rather course mesh, however it may be necessary to base it on the low memory implementation to avoid difficulties in storing many clipped triangles.

Up to now it has been very costly to try to find surface areas at every time step of a simulation in order to add, for example, solvation energies. Our clipping based algorithm has demonstrated a remarkable improvement over certain current techniques. Since memory is often limited, one version finds all pairs with a given atom at one time, finds the clipping planes, removes (or clips) the appropriate points, and stores the remaining area of that atom. A second version finds each atom pair only once, thus saving time, but requires about eight megabytes of storage for a $L = 5$ (102 point), 4000 atom problem using LOGICAL*2 variables. As all areas are associated with an atom, it is trivial to collect areas separated by various atom types. To be most useful to the greatest range of problems, a variable method of assigning types must be implemented. Our software can group surface areas by atom type, polymer chain, coordinate value, or any other assignments available in an input file. An example of the usefulness of this grouping is given next.

V.B. Effective Grouping of Surface Areas

The subdivision of surface areas can give us much more insight into a system than a single total surface area. Various assignments are possible with that of atom type perhaps being the most useful. This can aid in finding the distinct solvation energies of hydrophobic and hydrophilic regions, for example. The usefulness of alternative schemes is best shown by example. We return to the β-alanine starburst dendrimers of Reference [14]. These polymers start from a center nitrogen and systematically

Generation	Displacement from Core						
	1	2	3	4	5	6	7
1	117.6						
2	96.3	119.3					
3	99.8	102.9	121.0				
4	99.5	98.8	99.2	121.7			
5	100.8	97.2	92.2	103.3	120.0		
6	104.6	102.0	101.7	101.0	100.7	123.3	
7	104.2	101.9	101.6	101.4	101.5	101.8	123.5

Table 2.18. van der Waals Surface Area per Monomer Unit. For each generation of β-alanine starburst dendrimers, the surface area is collected by the separation in monomers from the core. Areas are in $Å^2$.

replace external hydrogens with additional polymer units to build star-like radiating structures. It is desirable to assess the changes in the characteristics of these structures as the generation (the number of times the external hydrogens have been replaced) increases. In Reference [34], Naylor presents the change in surface area per added monomer unit as the generation is increased. Using standard input files, we are able to obtain a more direct measure by calculating the surface area per monomer unit for each level of added monomer units for every generation. The results for the van der Waals surface and for the solvent-accessible surface with a probe radius of 1.4Å (suitable for water) are given in Tables 2.18 and 2.19. Obviously, the most external monomer units have the greatest surface area per unit. The van der Waals surface area of the external units remain rather constant as the generation increases. The solvent-accessible surface area of the external units show a decrease beginning with generation four. This demonstrates that at van der Waals distances there is freedom of movement among the terminal monomers but that congestion begins to limit the ability to solvate the terminal units at generation four and above. The decrease in accessible area for the next-to-terminal monomers already begins at gen-

Generation	Displacement from Core						
	1	2	3	4	5	6	7
1	217.4						
2	132.6	219.8					
3	133.7	142.1	221.0				
4	133.9	127.0	125.7	210.0			
5	136.1	126.0	116.9	119.8	192.6		
6	137.8	119.3	117.8	104.0	107.5	185.6	
7	136.8	109.6	120.2	105.1	108.2	102.6	178.1

Table 2.19. Accessible Surface Area per Monomer Unit. For each generation of β-alanine starburst dendrimers, the surface area is collected by the separation in monomers from the core. Areas are in Å^2.

eration three as they must contend with interactions both on the core and surface sides of the dendrimer. Note that the accessible area of monomer units nearest the core remain rather stable as the core nitrogen restricts the degrees of freedom of these three units.

VI. Summary

This chapter has pursued the use of spherical meshes in several areas of research. A general mesh scheme was presented from which one can choose the required number of points, polygons, or degrees of freedom. The meshes were analyzed and optimized with regard to several important geometric quantities for efficient use in a large variety of problems. In particular the nearest area grid method was presented and applied to spherical harmonic displays, numerical integration, visualization of large molecules and crystals, and the calculation of molecular properties. This new grid methodology associates an area (or weight) with each point. The area is that which is closest to that point than any other. Where appropriate, intermediate applied

mathematical applications were explored which have use beyond the field of chemistry. The chemical applications are mainly oriented around chemical visualization and molecular surface properties, two important fields where both algorithmic and conceptual advances have been made. A good example of each is the better understanding of the reconstruction of semiconductor surfaces and the investigation of the topologies of starburst dendrimers. These techniques involving the surface properties of large systems are complementary to the modeling of reactive dynamics of the first chapter which can combine to create the more accurate simulation of large systems with much decreased effort.

A. Finding Areas on a Sphere

The methods of this appendix are used to find the surface area closer to one point than any other in a spherical mesh. Specifically, we derive equations to find that part of the area within a polygon that is closest to one particular vertex for two cases: first, a quadrilateral composed of segments of latitude and longitude lines and second, a general triangle on a sphere. In both cases we actually require the area of the sphere segment defined by the vertices of the polygon instead of the flat polygon itself. The area of a polygon of k sides on a unit sphere is [35]:

$$A = \sum_{i=1}^{k} \alpha_i - (k-2)\pi, \tag{2.40}$$

where α_i is the angle between two of the arcs of the polygon. Since we usually will be starting from the point coordinates instead of the angles, we show how to derive the angles from the coordinates next.

Consider three points, $\vec{P}_1$, $\vec{P}_2$, and $\vec{P}_3$, on a unit sphere. We wish to know the angle α between the two arcs $(\vec{P}_2 - \vec{P}_1)$ and $(\vec{P}_2 - \vec{P}_3)$. Note that this is not the angle $(\vec{P}_1 - \vec{P}_2 - \vec{P}_3)$ which is not tangent to the spheres surface. First we must find

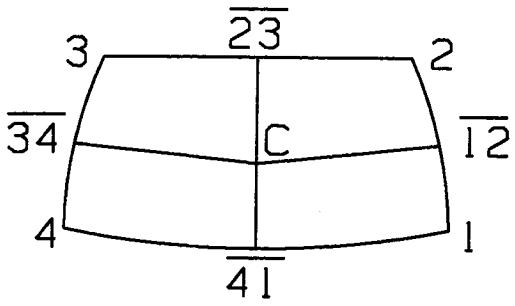


Figure 2.14. Quadrilateral on Sphere with "Center" and Arc Midpoints.

the normalized directions $\vec{S}_{21}$ and $\vec{S}_{23}$ which are tangent to the sphere's surface and point in the directions of the arcs $(\vec{P}_2 - \vec{P}_1)$ and $(\vec{P}_2 - \vec{P}_3)$, respectively, at $\vec{P}_2$. Since $\vec{S}_{21}$ is perpendicular to both $\vec{P}_2$ and $\vec{P}_2 \times \vec{P}_1$, we have

$$\vec{S}_{21} = \frac{(\vec{P}_2 \times \vec{P}_1) \times \vec{P}_2}{|\vec{P}_2 \times \vec{P}_1|} \tag{2.41}$$

and a similar expression for $\vec{S}_{23}$. The angle between the arcs is now found by

$$\cos\alpha = \vec{S}_{21} \cdot \vec{S}_{23} = \frac{\vec{P}_1 \cdot \vec{P}_3 - (\vec{P}_2 \cdot \vec{P}_1)(\vec{P}_2 \cdot \vec{P}_3)}{|\vec{P}_2 \times \vec{P}_1||\vec{P}_2 \times \vec{P}_3|}. \tag{2.42}$$

Now consider the case of a latitude-longitude quadrilateral. Let it have vertices $\vec{P}_1$ through $\vec{P}_4$, and let $\vec{P}_{\overline{ij}}$ be the center of the arc $(\vec{P}_i - \vec{P}_j)$ (see Figure 2.14 for a diagram). The point positions are given by:

$$\begin{aligned} \vec{P}_1 &= (0, \rho_1, z_1), \\ \vec{P}_2 &= (0, \rho_2, z_2), \\ \vec{P}_3 &= (\rho_2 \sin\delta, \rho_2 \cos\delta, z_2), \\ \vec{P}_4 &= (\rho_1 \sin\delta, \rho_1 \cos\delta, z_1), \end{aligned} \tag{2.43}$$

where z_1 and z_2 are the z-coordinates of the latitude lines, $\rho_i^2 + z_i^2 = 1$ for $i = 1, 2$, and $\delta = \frac{\pi}{2L}$ is the angle between two longitude lines. The arc midpoints are given by:

$$\vec{P}_{\overline{ij}} = \frac{\vec{P}_i + \vec{P}_j}{|\vec{P}_i + \vec{P}_j|}. \tag{2.44}$$

Let the point $\vec{C}$ be equidistant from all four of the original vertices. The area within the polygon closest to either $\vec{P}_1$ or $\vec{P}_4$ is the area contained by the quadrilateral $(\vec{P}_1 - \vec{P}_{\overline{12}} - \vec{C} - \vec{P}_{\overline{41}})$; the area closest to $\vec{P}_2$ or $\vec{P}_3$ is contained by the quadrilateral $(\vec{P}_2 - \vec{P}_{\overline{23}} - \vec{C} - \vec{P}_{\overline{12}})$. All that remains to do then is to find the point $\vec{C}$. We start with

$$\vec{C} = \left(\rho_c \sin\frac{\delta}{2}, \rho_c \cos\frac{\delta}{2}, z_c\right), \tag{2.45}$$

where $\rho_c^2 + z_c^2 = 1$. Since $\vec{P}_1 \times \vec{C} = \vec{P}_2 \times \vec{C}$, we can find the ratio of ρ_c to z_c by the equation:

$$\cos\left(\frac{\delta}{2}\right)(\rho_1 - \rho_2)\rho_c = (z_2 - z_1)z_c, \tag{2.46}$$

and then solve for them by normalizing. Now we have all the information required to find the a_{mn}'s for the quadrilaterals of method LL.

We now desire to find the area within a triangle with vertices $\vec{P}_1$, $\vec{P}_2$, and $\vec{P}_3$ which is closest to each vertex. Again let $\vec{P}_{\overline{ij}}$ be the midpoint of arc $(\vec{P}_i - \vec{P}_j)$ and $\vec{C}$ be the point equidistant from $\vec{P}_1$, $\vec{P}_2$, and $\vec{P}_3$. Also let $\vec{P}_{ij} = \vec{P}_j - \vec{P}_i$. We know that $\vec{C}$ is perpendicular to $\vec{P}_{12}$ and $\vec{P}_{23}$, thus:

$$\vec{C} = \frac{\vec{P}_{12} \times \vec{P}_{23}}{|\vec{P}_{12} \times \vec{P}_{23}|}. \tag{2.47}$$

There are two main cases to consider: one that $\vec{C}$ falls within the triangle and two that $\vec{C}$ falls outside the triangle. See Figure 2.15 for a diagram of each case. $\vec{C}$ will be within the triangle if all of the following apply:

$$\begin{aligned} \vec{P}_{12} \cdot \vec{P}_{23} &\geq 0, \\ \vec{P}_{23} \cdot \vec{P}_{31} &\geq 0, \\ \vec{P}_{31} \cdot \vec{P}_{12} &\geq 0. \end{aligned} \tag{2.48}$$

In that case the area within the triangle closest to point $\vec{C}$ is the area of quadrilateral $(\vec{P}_1 - \vec{P}_{\overline{12}} - \vec{C} - \vec{P}_{\overline{31}})$ and similarly for the other points. Since at most one of the

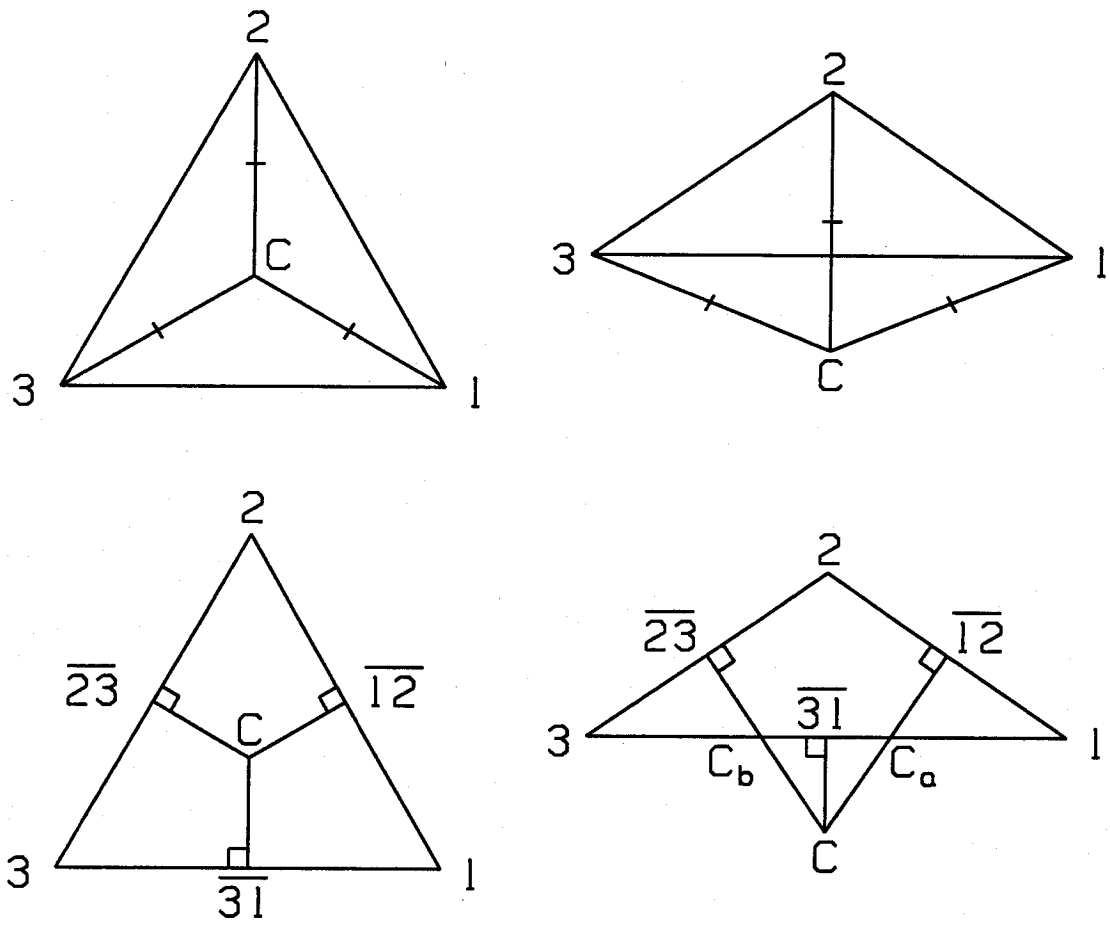


Figure 2.15. Triangle on Sphere with $\vec{C}$ Inside (left) and Outside (right) of Triangle.

relations in Equations (2.48) is false, we will only consider the case $\vec{P}_{12} \cdot \vec{P}_{23} < 0$ which is shown in Figure 2.15. Here the area of the triangle closest to $\vec{P}_1$ is bounded by the triangle $(\vec{P}_1 - \vec{P}_{\overline{12}} - \vec{C}_a)$, the area closest to $\vec{P}_3$ is bounded by the triangle $(\vec{P}_3 - \vec{C}_b - \vec{P}_{\overline{23}})$, and the area closest to $\vec{P}_2$ is bounded by the pentagon $(\vec{P}_2 - \vec{P}_{\overline{23}} - \vec{C}_b - \vec{C}_a - \vec{P}_{\overline{12}})$ (in practice, subtracting the area of the first two areas from the area of triangle $(\vec{P}_1 - \vec{P}_2 - \vec{P}_3)$ is easier). Thus, once we solve for $\vec{C}_a$ and $\vec{C}_b$, we can find the appropriate areas. Since these points lie on the arc $(\vec{P}_1 - \vec{P}_3)$, we can write

$$\vec{C}_a = \frac{\alpha\vec{P}_1 + (1-\alpha)\vec{P}_3}{|\alpha\vec{P}_1 + (1-\alpha)\vec{P}_3|}, \tag{2.49}$$

and

$$\vec{C}_b = \frac{\beta\vec{P}_3 + (1-\beta)\vec{P}_1}{|\beta\vec{P}_3 + (1-\beta)\vec{P}_1|}, \tag{2.50}$$

where α and β are constants to be determined. Since $\vec{C}_a \cdot \vec{P}_{12} = 0$ and $\vec{C}_b \cdot \vec{P}_{23} = 0$, we can determine that

$$\alpha = \frac{\vec{P}_3 \cdot \vec{P}_{12}}{\vec{P}_3 \cdot \vec{P}_{12} - \vec{P}_1 \cdot \vec{P}_{12}} \tag{2.51}$$

and

$$\beta = \frac{\vec{P}_1 \cdot \vec{P}_{23}}{\vec{P}_1 \cdot \vec{P}_{23} - \vec{P}_3 \cdot \vec{P}_{23}}. \tag{2.52}$$

References

[1] E. P. Wigner and F. Seitz, Phys. Rev. **43**, 804 (1933).

[2] G. L. Dirichlet, Z. Reine Angew. Math. **40**, 216 (1850).

[3] G. F. Voronoi, Z. Reine Angew. Math. **134**, 198 (1908).

[4] F. C. Frank and J. S. Kasper, Acta Crystallogr. **11** 184 (1958).

[5] Eugene Butkov, *Mathematical Physics* (Addison-Wesley Publishing Company, Reading, Massachusetts, 1968).

[6] Mitchel Weissbluth, *Atoms and Molecules* (Academic Press, Inc., New York, 1978).

[7] For a similar discussion based on the Y_{lm}'s instead of the Z_{lm}'s and using standard bases based on raising and lowering operators instead of Cartesian coordinates, see J. S. Griffith, *The Theory of Transition-Metal Ions* (Cambridge University Press, London, 1961).

[8] For a general coverage of numerical integration techniques, see A. H. Stroud, *Approximate Calculation of Multiple Integrals* (Prentice Hall, Inc., Englewood Cliffs, New Jersey, 1971).

[9] V. I. Lebedev, Zh. Vychisl. Mat. Mat. Fiz. **15**, 48 (1975). English translation in U.S.S.R. Comput. Math. & Math. Phys., **15**, No. 1, 44-51 (1975).

[10] V. I. Lebedev, Zh. Vychisl. Mat. Mat. Fiz. **16**, 293 (1976). English translation in U.S.S.R. Comput. Math. & Math. Phys., **16**, No. 2, 10-24 (1976).

[11] V. I. Lebedev, Sibirsk. Mat. Zh. **18**, 132 (1977). English translation in Siberian Math. **18**, No. 1, 99-107 (1977).

[12] G. A. Glatzmaier, J. Comput. Phys. **55**, 461 (1984). See also R. A. Friesner, J. Phys. Chem. **92**, 3091 (1988) for a more complete description.

[13] For a general coverage of computer graphics techniques, see J. D. Foley and A. Van Dam, *Fundamentals of Interactive Computer Graphics* (Addison-Wesley, Reading, Massachusetts, 1982).

[14] For a review of starburst dendrimers, see D. A. Tomalia, A. M. Naylor, W. A. Goddard, Angew. Chem. **29**, 117 (1990).

[15] R. E. Schlier and H. E. Farnsworth, J. Chem. Phys. **30**, 917 (1959).

[16] E. G. McRae, Phys. Rev. B **28**, 2305 (1983); P. A. Bennet, L. C. Feldman, Y. Kuk, E. G. McRae, and J. E. Rowe, Phys. Rev. B **28**, 3656 (1983).

[17] K. Takayanagi, Y. Tanishiro, M. Takahashi, and S. Takahashi, J. Vac. Sci. Technol. A **3**, 1502 (1985); K. Takayanagi, Y. Tanishiro, M. Takahashi, and S. Takahashi, Surf. Sci. **164**, 367 (1986).

[18] P. K. Weiner, R. Langridge, J. M. Blaney, R. Schaefer, and P. A. Kollman, Proc. Natl. Acad. Sci. USA **79**, 3754 (1982).

[19] M. L. Connolly, J. Molec. Graph. **4**, 3 (1986).

[20] I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge, and R. E. Ferrin, J. Mol. Biol. **161**, 269 (1982); R. L. DesJarlais, R. P. Sheridan, G. L. Seibel, J. S. Dixon, I. D. Kuntz, and R. Venkataraghavan, J. Med. Chem. **31**, 722 (1988).

[21] D. Eisenberg and A. D. McLachlan, Nature **319**, 199 (1986).

[22] B. C. Gates, J. R. Katzer, G. C. A. Schuit, *Chemistry of Catalytic Processes* (McGraw-Hill Book Company, New York, 1979).

[23] B. Lee and F. M. Richards, J. Mol. Biol. **55**, 379 (1971).

[24] F. M. Richards, Ann. Rep. Biophys. Bioeng. **6**, 151 (1977).

[25] S. Meirtus, E. Scrocco, and J. Tomasi, Chem. Phys. **55**, 117 (1981).

[26] J. L. Pascual-Ahuir and E. Silla, J. Comp. Chem. **8**, 778 (1987).

[27] J. Greer and B. L. Bush, Proc. Natl. Acad. Sci. USA **75**, 303 (1978).

[28] M. L. Connolly, Thesis, Univ. of Calif. at Berkeley (1981).

[29] L. H. Pearl and A. Honegger, J. Mol. Graph. **1**, 9 (1983).

[30] J. J. Müller, J. Appl. Cryst. **16**, 74 (1983).

[31] M. Y. Pavlov and B. A. Fedorov, Biopolymers **22**, 1507 (1983).

[32] M. L. Connolly, J. Appl. Cryst. **16**, 548 (1983).

[33] J. L. Pascual-Ahuir and E. Silla, Quantum Chemistry Program Exchange, Program No. 554.

[34] A. M. Naylor, Thesis, California Institute of Technology (1989).

[35] D. J. Stuik, *Lectures on Classical Differential Geometry* (Addison-Wesley Press, Inc., Cambridge, Massachusetts, 1950) pp. 153-159.