ABSTRACT

FLUID-PHASE THERMODYNAMICS FROM MOLECULAR-LEVEL PROPERTIES AND INTERACTIONS BASED IN QUANTUM THEORY

by
Steven G. Arturo

A methodology to predict the thermodynamics of macroscopic fluid systems from quantum chemistry and statistical thermodynamics has been developed. This work extends the group-contribution concepts most utilized in chemical engineering. Computational chemistry software is used to define the geometries and electron density profiles of target molecules. Atoms in Molecules theory and associated software packages are used to calculate rigorous properties of the functional groups within molecules of interest. These properties are incorporated into an intermolecular potential energy function which describes interactions between entire molecules as a set of interactions between functional groups. This information is applied to a lattice-fluid model with the capability to predict volumetric properties of pure fluids and vapor/liquid equilibrium properties of mixture systems. This work develops a bridge from chemistry at the molecular level to the statistical mechanics at the macroscopic system level.

The rigorous properties of functional groups lead to the application of first-principles mathematical models that qualitatively agree with volumetric properties of pure fluids and predict vapor/liquid equilibrium behavior for near-ambient mixtures of alkanes, alcohols and ethers. The theoretical and computational efforts developed in this work offer engineers the ability to determine molecular-level modeling parameters within engineering models without the use of experiment.

FLUID-PHASE THERMODYNAMICS FROM MOLECULAR-LEVEL
PROPERTIES AND INTERACTIONS BASED IN QUANTUM THEORY

by
Steven G. Arturo

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Chemical Engineering

Otto H. York Department of Chemical Engineering

May 2005

FLUID-PHASE THERMODYNAMICS FROM MOLECULAR-LEVEL PROPERTIES AND INTERACTIONS BASED IN QUANTUM THEORY

Steven G. Arturo

Dr. Dana E. Knox, Dissertation Advisor   Date
Professor of Chemical Engineering, NJIT

Dr. Robert B. Barat, Committee Member   Date
Professor of Chemical Engineering, NJIT

Dr. Joseph W. Bozzelli, Committee Member   Date
Distinguished Professor of Chemistry, NJIT

Dr. Michael C. Y. Huang, Committee Member   Date
Assistant Professor of Chemical Engineering, NJIT

Dr. Lev N. Krasnoperov, Committee Member   Date
Professor of Chemistry, NJIT

BIOGRAPHICAL SKETCH

Author: Steven G. Arturo

Degree: Doctor of Philosophy

Date: May, 2005

Undergraduate and Graduate Education

- Doctor of Philosophy in Chemical Engineering
  New Jersey Institute of Technology, Newark, NJ, 2005

- Bachelor of Science in Applied Mathematics
  New Jersey Institute of Technology, Newark, NJ, 2000

Major: Chemical Engineering

Presentations and Publications:

Arturo, S.G. and Knox, D.E., "Steps to a Bridge: Multiscale Modeling of the Thermodynamics of a Fluid," AIChE 2004 Annual Meeting, paper 46a, 2004.

Carrillo, M., Kim, H., Arturo, S.G. and Knox, D.E., "Prediction of Mixture Properties by Atomic-Contribution Methods," poster presented at the AIChE 2004 Annual Meeting in Austin, TX, November 7-12, 2004.

Arturo, S.G. and Knox, D.E., "Pure Species Thermodynamics using Lattice Fluid Theory with a Study into Critical Fluctuation Phenomena," poster presented a the AIChE 2004 Annual Meeting in Austin, TX, November 7-12, 2004.

Arturo, S.G. and Knox, D.E., "Interaction Model for Functional Groups Using Properties Found in Quantum Calculations," paper presented at the 224th ACS National Meeting in Boston, MA, August 18-22, 2002.

Knox, D.E. and Arturo, S.G., "Use of *ab initio* Calculations for Evaluation of Parameters in Solubility Models," poster presented at the IUPAC 10th International Symposium on Solubility Phenomena in Varna, Bulgaria, July 21-26, 2002.

to Mom, Dad
to Laura, Bobby, Danny, Raffela
to Jonathon, Jaimie, Daniel, Kimberly
to Grandparents, Uncles, Aunts, Cousins
to Cornelia, Alex, Irina, Andrew, Hellen, Sandy
to my dearest friends

to Emi
from the first day of this journey, to the last.

ACKNOWLEDGMENT

TABLE OF CONTENTS

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

TABLE OF CONTENTS
(Continued)

LIST OF TABLES

LIST OF FIGURES

Figure   **Page**

LIST OF SYMBOLS

A	molar Helmholtz energy
A_k	external surface area of group k
A^E	excess molar Helmholtz energy
A^t	total Helmholtz energy
a_j	occupation number of quantum state j
$\mathbf{a}$	vector describing distribution of energy states
B	second virial coefficient
B_{ii}	second virial coefficient of species i
B_{ij}	cross second virial coefficient
b_i	volume of molecule i
b_j	coefficient of series expansion within hydrogen atom wavefunction
C	third virial coefficient
D	fourth virial coefficient
E	energy
E_0	ground state energy
$E_0^{[i]}$	the ith derivative of the ground-state energy with respect to the perturbation parameter λ
E_S	interaction energy evolved from orbital overlap
E_{xc}	exchange-correlation energy
$\tilde{E}$	reduced energy eigenvalue of quantum system
$\bar{E}$	average energy
E^t	total energy
E_{HF}	Hartree-Fock energy
$E_{int}^{exp/6}$	exponential-6 intermolecular potential energy function

E_{int}^{HS}	hard-sphere intermolecular potential energy function
E_{int}^{Kihara}	Kihara intermolecular potential energy function
E_{int}^{LJ}	Lennard-Jones intermolecular potential energy function
E_{int}^{SW}	square well intermolecular potential energy function
E_{int}^{t}	total interaction energy
$E_{internal}^{t}$	total system energy attributable to internal molecular energy modes
E_{j}^{t}	total energy of quantum state j
E_{0}	ground state energy of a quantum system
$E_{0}^{t,(i)}$	the ith derivative of the total ground state energy with respect to the perturbation parameter λ
$\bar{E}^{t}$	average total energy
$\bar{E}_{q\mu}$	long-range charge-dipole interaction energy
$\bar{E}_{\mu\alpha}$	long-range dispersion interaction energy caused by a dipole
$\bar{E}_{\mu\mu}$	long-range dipole-dipole interaction energy
e	charge of one electron
F_{i}	general thermodynamic property for species i
$\bar{F}_{i}$	average atomic property of i
f_{j}	general contribution of group j to a thermophysical property
f_{ij}	Mayer f-function
$\hat{f}_{i}$	Fockian operator of molecule i
G	molar Gibbs energy
G_{i}	molar Gibbs energy of species i
G^{E}	excess molar Gibbs energy
G^{id}	molar Gibbs energy of an ideal solution
G^{ig}	molar ideal gas Gibbs energy
G^{R}	molar residual Gibbs energy

G^t	total Gibbs energy
G_i^t	total Gibbs energy of pure species i
$\left(G^t\right)^E$	total excess Gibbs energy
$\left(G^t\right)^{id}$	total Gibbs energy of an ideal solution
g_{ijk}^{GTF}	Gaussian-type function representing an orbital with integer exponents i, j, and k
g_{nlm}^{STO}	Slater-type orbital with quantum numbers n, l, and m
H	molar enthalpy

Greek Symbols

α	Lagrange undetermined multiplier
α_i	polarizability of i
β	Lagrange undetermined multiplier
Γ_k	the activity coefficient of surface segment/group k
Γ_{ij}	nonrandomness factor in lattice-fluid theory
γ_i	activity coefficient of species i
Δ	Gibbs ensemble partition function
δ	three-body interaction function
δ_{mn}	difference of second virial coefficients in mixture system
ε	magnitude of potential energy well
ε_i	energy eigenvalue for electron i in HF-SCF method
ε_{ij}	the interaction energy between i and j
ϕ_i	volume fraction of species i
φ	polar angle
κ	factor within exponent in generalized Guggenheim statistics
κ	isothermal compressibility

λ	perturbation parameter
μ	reduced mass within the hydrogen atom
μ_i	chemical potential of species i
μ_i	dipole moment of i
$\nu_{0,i}$	characteristic ground-state electronic frequency of i
$\nu_{i,k}$	number of group k within molecule i
Θ_k	surface area fraction for group k
θ	azimuthal angle
θ_i	surface area fraction for species i
θ_i'	surface area fraction for species i on fully occupied lattice
θ_{ij}	local surface area fraction of j molecules around i molecules
ρ	electron density
ρ_0	ground state electron density
Φ_i	factor containing non-idealities of species i in vapor phase
σ	effective radius of interacting body within empirical intermolecular potentials
τ_Q	terms within summation of canonical partition function
τ_G	terms within summation of Gibbs partition function
τ_{ij}	exponential of interchange energy
Ω	degeneracy, ways function
Ω_k	degeneracy of energy level k
Ξ	grand canonical ensemble partition function
Ψ	the wavefunction of a system
Ψ^*	complex conjugate of the wavefunction
Ψ_0	the ground-state wavefunction of a system
ψ	hydrogen atom wavefunction

ψ_{nlm}	hydrogen atom wavefunction
ϖ_i	factor containing degeneracy and lattice symmetry factors for molecule i
ξ	Gaussian-type function exponent
χ	perturbation parameter
ζ	Slater-type orbital exponent

Miscellaneous Symbols

$\mathfrak{A}$	total number of systems in the ensemble
$\mathfrak{E}$	total energy of the ensemble
∇	gradient operator
∇^2	Laplacian operator
∇_i^2	Laplacian operator with respect to coordinates of i
$(ij\|kl)$	two electron integral involving orbitals i, j, k and l

Superscripts

ath	athermal system
E	excess
ig	ideal gas system
id	ideal solution
t	total
R	residual
$-$	average
$\sim$	reduced
$\wedge$	operator
(i)	pure species i

$[i]$ i^{th} order

$\{i\}$ i^{th} iterate

Acronyms

COSMO Conductor-Like Screening Model

UNIFAC UNIQUAC Functional Group Activity Coefficient

SGA Spherical Gaussian Approximation

CHAPTER 1

INTRODUCTION

1.1 Objective

This work lays out an algorithm that achieves thermodynamic properties of fluids from rigorous properties at the molecular and functional group level. The fluid properties of interest include conditions and compositions for vapor/liquid equilibrium of binary mixtures, pure species volumetric behavior and heats of formation. The functional group properties calculated in this work include structural, electrostatic and energetic characteristics. The modeling results are geared toward application to systems of interest to chemical engineers, while the fundamental approach and the molecular-level concepts studied are geared toward models of interest to physical chemists.

The layout of this work is separable into four topics: chemical engineering science; computational chemistry; quantum chemistry of interactions between closed-shell molecules; and statistical mechanics of lattice-fluid systems. Chapters 2 and 3 review the fundamental expressions within classical and statistical thermodynamics to theoretically describe macroscopic and molecular systems. Chapters 4 and 5 review the statistical methods chemists and engineers have formulated for describing fluid systems, namely lattice-fluid models and the modifications made by engineers to describe real systems. Chapters 6 and 7 review the theory and methods behind quantum chemistry that predict energies and electron density profiles of molecules numerically, and Chapter 8 presents results of computations for molecules of interest to this work. Chapter 9 reviews the concept of the functional group in modeling molecular behavior, and Chapter 10

presents functional group property calculations using a rigorous method based in quantum chemical theory. Chapter 11 reviews classical and modern approaches to the problem of interacting closed-shell molecules. Chapter 12 presents novel models for interactions between molecules and functional groups utilizing calculations of Chapters 8 and 10 with the theoretical expressions of Chapter 11. Chapters 13 and 14 describe novel applications of both the molecular-level properties and interaction energies to lattice-fluid theory towards a first-principles description of the thermodynamics of pure and mixture systems. Chapters 1 through 7, 9 and 11 review prior work, while the remaining chapters document the efforts of the author.

1.2 Background Information

Engineers are responsible for large scale processes important to the health and well-being of people in today's society. The chemicals that engineers assist in manufacturing lead to a higher standard of living for people by producing food for a more stable food supply, detergents for a cleaner lifestyle and medicines for a longer and healthier life. Today's chemical engineer must devise safe, economical and efficient processes to bring large amounts of advanced products to those who can benefit.

The science of chemical thermodynamics solves a variety of problems for chemical engineers who oversee large-scale operations of fluid systems, the systems within which most of today's mass-produced chemicals are created. Knowledge of fluid properties is used to accurately build the reactors and plants where manufacturing is centered. The processes may also be optimized after models of heat and work effects on the system are derived from thermodynamics. All stages of the process, from the initial

streams to intermediate properties to output flows, are able to be incorporated into thermodynamic models of the process, thus allowing for a full characterization of the product at several stages.

In all levels of processing, properties of the fluid phase are essential. Engineers must know properties at various temperatures and pressures, two thermodynamic variables readily controlled and monitored. Engineers must be able to predict the response of fluids to changing heat and work effects, therefore properties of pure fluids and mixtures must be known for relevant temperatures and pressures. For mixture systems, which are the majority of systems in practice, properties must be known for varying compositions of species as well. These properties are achievable with knowledge of the volumetric behavior and heat capacity of the fluid, which are mathematical relations available to the engineer.

Of particular importance to engineers is the process of separating chemicals. Reactions that form the target compound rarely go to completion, resulting in a complex mixture of reactants, by-products, solvents and the desired species. The compound of interest usually needs to be separated from the remaining species, and sometimes this must be done on a large scale. The classic approach is to use a distillation process, which takes advantage of the different vapor pressures of the species at different conditions and compositions.

The solubility of a trace compound within solvents is another thermodynamic concept used in separations. Whether it be the solubility of a gaseous species within a liquid or a solid, or the solubility of a solid or liquid species within a gas, a range of

separation techniques exist by taking advantage of the affinity a compound has to a different phase or solvent.

In all systems studied by chemical engineers, the behavior of fluids at the various conditions is representable either through tables and graphs, or analytically using correlations or first-principles approaches. For the former, experiments have to be conducted to amass volumetric data for pure systems and phase behavior data for mixture systems at different conditions and compositions. A large number of experiments have to be conducted to get a fine representation of all the states of interest, and states between experimental results are then interpolated. To get such information, a great amount of time and money must be invested. As an alternative, a limited number of experiments may be used to correlate parameters within a mathematical expression. Analytical expressions result from such a treatment, although the correlation is usually only good within the range of conditions within the experiments. Analytical expressions with a predictive capability over a full range of conditions are possible if the fundamental driving forces of the system are defined. These mathematical expressions are usable by scientists and engineers to give more timely results of an unknown system than by methods that involve experiment. Having such expressions available makes engineers more efficient and effective at predicting the behaviors of fluids in a process.

CHAPTER 2

CLASSICAL THERMODYNAMICS

Equilibrium thermodynamics is the study of natural laws that describe the state of material at given conditions (e.g. temperature and pressure) and composition. The study of such laws commenced with the study of steam engines to determine how much energy can be extracted in the form of work and how much energy is wasted in the form of heat. Scientists continue striving to understand the underlying principles of energy and how it drives processes. Beginning as an applied science, the framework has matured into abstract mathematical concepts and now is used by chemical engineers and scientists to make all kinds of quantitative assessments, on systems from the molecular level to systems on the scale of astronomical bodies.

This chapter reviews concepts familiar to those who have learned undergraduate chemical engineering thermodynamics. The state function formalism is reviewed, and notation conventions are established; these may differ from those with which chemists are familiar. The quantification of nonidealities, namely through defining residual and excess properties, and the utilization of equations of state and activity coefficients within these expressions are shown. The gamma-phi formulation, a set of equations used by engineers to calculate vapor/liquid equilibrium conditions, and the algorithms using these equations are introduced. Those who wish for a more in depth discussion are referred to a standard textbook in the area (Smith, Van Ness, & Abbott, 1996).

2.1 The State Function Formalism

The state function formalism of equilibrium thermodynamics simplifies the description of systems so that engineers need only be concerned with the initial and final conditions of a process. The change in energy of the system is independent of the path the process takes. Different external effects change the path off the system. Such concepts fall within the study of kinetics and reactor design. Thermodynamics is used by engineers when they need to establish the initial and final conditions, any intermediate states, and limits of heat and work effects on the system. Such effects can then be applied to create the desired output.

The most important state function within all of thermodynamics is the Gibbs energy. By the Second Law of Thermodynamics, a pure species system that has achieved equilibrium has reached a state that minimizes the Gibbs energy at the given temperature and pressure. The Gibbs energy expresses a balance between the internal energy U of the system, the entropy S at the system temperature T, and the pressure-volume effect that the system encounters to exist with volume V at a thermodynamic pressure p (Levenspiel, 1996)

$$G = U - TS + pV \tag{2.1}$$

A state function of theoretical interest is the Helmholtz energy of the system, which is expressed here through the Gibbs energy

$$A = G - pV = U - TS \tag{2.2}$$

In processes, the enthalpy of the system is important to determine heat effects and phase changes of flow systems. This state function is given by

$$H = G + TS = U + pV \tag{2.3}$$

The Gibbs energy is most important to engineers because of the natural variables of the state function. Changes are wholly describable through changes in the pressure, temperature and composition of the system, conditions engineers control and measure using a thermometer, barometer and species-measuring device. In differential form, a change in the total Gibbs energy of a system with multiple species is given by

$$dG^t = V^t dp - S^t dT + \sum_i \left(\frac{\partial G^t}{\partial n_i} \right)_{T,p,n_j} dn_i \qquad (2.4)$$

where the extensive properties here are given as total quantities, denoted by the superscript t, n_i is the number of moles of species i, and where the partial derivative is taken while holding T, p, and $n_{j \neq i}$ constant. A more usable form of the above differential is given by the dimensionless form of the equation, found to be

$$d \left(\frac{G^t}{RT} \right) = \frac{V^t}{RT} dp - \frac{H^t}{RT^2} dT + \sum_i \left(\frac{\partial G^t / RT}{\partial n_i} \right)_{T,p,n_j} dn_i \qquad (2.5)$$

where R is the ideal gas constant. The total entropy term is replaced by an equivalent expression including total enthalpy, a quantity more accessible to engineers.

A differential change equation for the Helmholtz energy is similarly expressed, but with different natural variables

$$dA^t = -p dV^t - S^t dT + \sum_i \left(\frac{\partial A^t}{\partial n_i} \right)_{V^t,T,n_j} dn_i \qquad (2.6)$$

where the change in A^t is expressed in terms of changes in volume, temperature and species amounts. The differential expression for a change in enthalpy is not as practical, since the change in entropy of the system needs to be known

$$dH^t = V^t dp + T dS^t + \sum_i \left(\frac{\partial H^t}{\partial n_i} \right)_{p,S^t,n_j} dn_i \tag{2.7}$$

Equations (2.4) through (2.7) are called the fundamental property relations (FPR) for the respective energies.

In each of the above FPRs, the change in the energy with respect to the amount of species i is called the chemical potential of species i. The chemical potential μ_i is given by

$$\mu_i = \left(\frac{\partial G^t}{\partial n_i} \right)_{T,p,n_j} = \left(\frac{\partial A^t}{\partial n_i} \right)_{V^t,T,n_j} = \left(\frac{\partial H^t}{\partial n_i} \right)_{p,S^t,n_j} \tag{2.8}$$

This quantity measures the affinity of a given species to a given phase. If two phases are in contact and molecules are able to move between phases, a molecule of type i moves from the phase with higher chemical potential to the phase with lower chemical potential. This flux of molecules continues until the chemical potential of both phases are equal. Equality of μ_i for each species in all phases defines the equilibrium of the system.

2.2 Real versus Ideal Systems

The study of the change in energy of a real system begins with the study of change in an idealized system. Expressions developed for idealized systems are almost always simpler and are used as first-order approximations. Gas-phase systems are related to the ideal gas state, where molecules neither interact nor displace any volume. Real gases tend toward ideal gases at low pressure, high temperature, and large molar volumes. The nonidealities of real gases develop from the volumes taken up by molecules and the forces of attraction and repulsion between them. Liquid-mixture systems are similarly

related to ideal solutions. Properties of ideal solutions of a given composition are the weighted average of pure liquid properties added to an idealized entropic effect. Ideal solutions are good approximations to mixture systems with species whose molecules have similar sizes and interactions. The nonidealities arise when molecules with different structures and/or unlike polarities are mixed.

Engineers relate the deviation of system properties of a real gas from the ideal state by considering residual properties. Only the equation of state (EoS) of the fluid is needed to calculate these properties. To see this, consider the definition of the molar residual Gibbs energy of the system

$$G^{R}(T,p) = G(T,p) - G^{ig}(T,p) \tag{2.9}$$

where each system exists at the same temperature and pressure. The FPR in Equation (2.5) is used for the pure species ideal gas and residual property

$$d\left(\frac{G^{ig}}{RT}\right) = \frac{V^{ig}}{RT}dp - \frac{H^{ig}}{RT^{2}}dT \tag{2.10}$$

$$d\left(\frac{G^{R}}{RT}\right) = \frac{V^{R}}{RT}dp - \frac{H^{R}}{RT^{2}}dT \tag{2.11}$$

Finding the changes in the Gibbs energy is straightforward for an ideal gas with knowledge of the ideal gas heat capacity and the ideal gas EoS. For a constant-temperature process, the change in the residual Gibbs energy is calculated by integrating the differential change from the ideal gas pressure, $p = 0$, to the pressure of interest

$$\frac{G^{R}}{RT} = \int_{0}^{p}[Z(T,p)-1]\frac{dp}{p} \tag{2.12}$$

where Z is the compressibility factor is found through the real gas EoS and is dependent on temperature and pressure

$$Z(T,p) = \frac{pV(T,p)}{RT} \qquad (2.13)$$

Similarly, the residual enthalpy in a constant temperature process is found to be

$$\frac{H^R}{RT} = -T \int_0^p \left(\frac{\partial Z}{\partial T}\right)_p \frac{dp}{p} \qquad (2.14)$$

It is therefore shown how an EoS in the form of the compressibility factor Z is used within the residual property formulation to determine the thermodynamic properties of a pure species. An analytical EoS makes the job of an engineer more efficient by allowing for the quantitative assessment of thermodynamic properties of real fluids through the use of Equations (2.12) through (2.14).

The excess properties of a liquid quantify the deviation of real liquid mixture behavior from ideal solution behavior. For instance, the molar Gibbs energy of an ideal solution is the weighted sum of energies from the pure liquids added to an entropic contribution

$$G^{id} = \sum_i x_i G_i + RT \sum_i x_i \ln x_i \qquad (2.15)$$

where x_i is the mole fraction of species i in the liquid phase. The excess molar Gibbs energy of a mixture is defined as follows

$$G^E(T,p,\mathbf{x}) = G(T,p,\mathbf{x}) - G^{id}(T,p,\mathbf{x}) \qquad (2.16)$$

where each system is at the same temperature, pressure, and composition $\mathbf{x}$. To develop mathematical expressions usable by engineers, it is necessary to consider the FPR of the total Gibbs energy, similar to Equation (2.5). For an ideal solution,

$$d\left[\frac{(G^t)^{id}}{RT}\right] = \frac{(V^t)^{id}}{RT}dp - \frac{(H^t)^{id}}{RT^2}dT + \sum_i \frac{\mu_i^{id}}{RT}dn_i \qquad (2.17)$$

where the chemical potential of species i in an ideal solution is denoted as μ_i^{id}.

Subtracting this expression from the FPR for a real mixture system gives the total excess

Gibbs energy

$$d\left[\frac{\left(G^t\right)^E}{RT}\right] = \frac{V^t - \left(V^t\right)^{id}}{RT}dp - \frac{H^t - \left(H^t\right)^{id}}{RT^2}dT + \sum_i \frac{\mu_i - \mu_i^{id}}{RT}dn_i \qquad (2.18)$$

Including excess property notation gives

$$d\left[\frac{\left(G^t\right)^E}{RT}\right] = \frac{\left(V^t\right)^E}{RT}dp - \frac{\left(H^t\right)^E}{RT^2}dT + \sum_i \frac{\mu_i - \mu_i^{id}}{RT}dn_i \qquad (2.19)$$

Within this expression, define the natural log of the activity coefficient of species i, $\ln\gamma_i$,

as proportional to the difference in chemical potentials of the real and ideal mixture

systems

$$RT\ln\gamma_i = \mu_i - \mu_i^{id} \qquad (2.20)$$

This relation quantitatively describes the different way a species behaves in a real

mixture as opposed to an ideal solution and is used to determine what phase is preferable

to a species. By inserting Equation (2.20) into Equation (2.19) and taking the derivative

with respect to amount of species, n_i, the activity coefficient is related to the total excess

Gibbs energy

$$\ln\gamma_i = \left\{\frac{\partial\left[\left(G^t\right)^E / RT\right]}{\partial n_i}\right\}_{p,T,n_j} \qquad (2.21)$$

Inserting the original expressions for the total excess Gibbs energy and the total Gibbs

energy of an ideal solution, the activity coefficient is found

$$\ln \gamma_i = \left[\frac{\partial \left(G^t / RT \right)}{\partial n_i} \right]_{p,T,n_j} - \left[\frac{\partial \left(G_i^t / RT \right)}{\partial n_i} \right]_{p,T} - \ln x_i \qquad (2.22)$$

where G_i^t is the total Gibbs energy of pure species i.

With development of models of the Gibbs energy for pure liquids and liquid mixtures, expressions are attainable for the activity coefficients of species within the system and, therefore, the thermodynamic behavior of those species in solution.

2.3 Classical Treatment of VLE Behavior

The activity coefficient plays a very important role in determining the vapor/liquid equilibrium (VLE) curves of mixture systems of interest to engineers. At low to moderate pressures, the engineering expressions relating the liquid and vapor compositions to the system temperature and pressure are called the gamma-phi formulation

$$y_i \Phi_i p = x_i \gamma_i p_i^{sat} \qquad (2.23)$$

where y_i is the mole fraction of species i in the vapor phase, p_i^{sat} is the saturated vapor pressure of species i at the system temperature and Φ_i contains factors that describe the nonidealities of species i within the system. These include vapor phase nonidealities accounted for by the virial EoS and liquid volume corrections accounted for by the Poynting factor. This quantity is given by

$$\Phi_i = \exp \frac{\left(B_{ii} - V_i^l \right) \left(p - p_i^{sat} \right) + \frac{p}{2} \sum_j \sum_k y_j y_k \left(2 \delta_{ji} - \delta_{jk} \right)}{RT} \qquad (2.24)$$

where B_{ii} is the pure species second virial coefficient, V_i^l is the molar liquid volume, and

δ_{mn} is a difference of pure and mixture second virial coefficients

$$\delta_{mn} = 2B_{mn} - B_{mm} - B_{nn} \qquad (2.25)$$

For a binary mixture, Equation (2.24) reduces to

$$\Phi_1 = \exp\frac{\left(B_{11} - V_1^l\right)\left(p - p_1^{sat}\right) + py_2^2\delta_{12}}{RT} \qquad (2.26)$$

for species 1, and the expression for species 2 is found by exchanging subscripts. In practice, the thermodynamic variables for VLE systems are temperature, pressure and composition. The dependence of the terms in the gamma-phi formulation is as follows

$$\Phi_i = f_\Phi\left(T, p, y_1, y_2, ..., y_{N-1}\right) \qquad (2.27)$$

$$\gamma_i = f_\gamma\left(T, p, x_1, x_2, ..., x_{N-1}\right) \qquad (2.28)$$

$$p_i^{sat} = f\left(T\right) \qquad (2.29)$$

Iterative procedures have been developed to determine full VLE tables and graphs with limited information by using the functionalities above. The following procedures are documented (Smith, et al. 1996), therefore the algorithms are not presented here. There is a slight discrepancy, since here γ_i is considered dependent on pressure. The documented procedure approximates γ_i as only dependent on temperature and liquid composition. The general dependence can be worked into the documented procedures by evaluating the activity coefficients with each iterated pressure.

The bubble point (BUBL) procedures calculate the composition of the vapor phase within the first bubble that forms within the liquid phase, while the dew point (DEW) procedures calculate the liquid composition of the first liquid drop that forms

from the vapor phase. The most common information is the liquid composition and either temperature (for an isothermal VLE graph) or pressure (for an isobaric VLE graph). The BUBL p procedure allows for pressure and vapor composition to be found with knowledge of the liquid composition and temperature. The BUBL T procedure allows for temperature and vapor composition to be found with knowledge of the liquid composition and pressure. The implementation of the DEW procedures is less common, since experimental data rarely includes only vapor phase compositions. Difficulty is encountered when trying to keep the entire headspace at uniform conditions to avoid fluctuations in density or concentrations. The DEW p procedure calculates the liquid composition and pressure with knowledge of the vapor composition and temperature. The DEW T procedure calculates the liquid phase composition and the temperature with knowledge of the vapor compositions and pressure.

2.4 Notes on the Critical Point of a Fluid

Most pure fluids and fluid mixtures have a particular set of conditions where classical mathematical formulations of thermodynamic properties break down. The critical point for a pure species is a set of conditions, usually given as the critical temperature and pressure, T_C and p_C, that denote the highest temperature and pressure where the liquid phase and the gas phase of a fluid have a distinct interface. Below such conditions, VLE may exist. Experimentally properties for a pure species EoS at the critical point have been found, namely

$$\left(\frac{\partial p}{\partial V} \right)_T = 0 \text{ and } \left(\frac{\partial^2 p}{\partial V^2} \right)_T = 0 \text{ at } p = p_C \text{ and } T = T_C \qquad (2.30)$$

To get a sense of the difficulties around the critical point, consider that the isothermal compressibility of a fluid, the change in volume when applying a pressure at constant temperature, is mathematically expressed as

$$\kappa = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T \qquad (2.31)$$

This expression is thereby infinite at the critical point, given by the conditions stated in Equation (2.30), meaning that one can theoretically apply pressure to a fluid without resistance. This is counterintuitive, since real world experience shows that a fluid resists occupying a smaller space by applying a force outward on the container. At the critical point, this is not the case.

Molecular-level characteristics of a species do not determine the thermodynamics around the critical point, since all fluids and fluid mixtures have critical points. Nor do the molecular properties affect the way the fluid approaches the critical point, either isothermally or isobarically. Such species-specific and composition-specific characteristics do determine the actual values of the critical temperature and pressure. Mathematical treatments for an EoS and a Gibbs energy expression do exist, but these are arduous to use and are high in theory. Systems far from the critical point are subject to the classical treatments described above and throughout this work. Introductory information of critical phenomena and its relation to phase transitions is explained (Stanley, 1971), and nonclassical approaches resolving some issues with the critical region are offered (Anisimov and Sengers, 2000).

2.5 Summary

Engineers benefit when the thermodynamic behavior of fluids and fluid mixtures are reduced to mathematical formulations. Equations of state can be used to determine real gas properties. Gibbs energy expressions can be used to determine the behavior of liquid mixtures. Although these expressions are more difficult to determine around the critical region, methods do exist and offer a quantitative description of the thermodynamics there. If one is able to determine the Gibbs energy for a generalized fluid system at any temperature, pressure and composition, it is found that both the EoS and activity coefficients are available for use in calculations. Theoretical frameworks that attempt to relate fundamental molecular behavior to macroscopic system properties exist. Statistical thermodynamics, an interpretation of molecular behavior that results in macroscopic thermodynamic properties, is the bridge between the very small entities and the large properties with which engineers are familiar.

CHAPTER 3

STATISTICAL THERMODYNAMICS

Statistical thermodynamics allows for the information at the molecular level to be translated into the macroscopic thermodynamics that engineers are interested in. The focus of macroscopic property prediction is to find the most probable energy levels of the system using a statistical analysis of the myriad combinations of molecular energy states, since the sheer number of molecules in a typical engineering system includes Avogadro's number of molecules, N_A, and a system degeneracy (the combinations of molecular states that result in the same system energy) of order 10^{N_A}. Knowledge of these most probable states leads to the macroscopic system properties engineers use in practice.

This chapter reviews concepts within physical chemistry and molecular physics that are not usually taught to undergraduate chemical engineers. The derivation of Boltzmann's distribution of quantized energy states available to a system is discussed, and expressions relating these distributions to macroscopic thermodynamics are established. The maximum term approximation for the summation within the partition function is expressed, as it will serve an important role in the calculation of thermodynamic properties. The origin of the virial equation of state, a theoretically rigorous equation of state, is noted, and the role of intermolecular potentials within gaseous systems is presented. Classical physics expressions for the interaction of molecules and statistical simplifications of these expressions are reviewed. The derivations and notation of the statistical thermodynamics in this chapter is a restatement

of work within a standard textbook (McQuarrie, 2000), as too are the expressions of the intermolecular potential functions (Hirschfelder, Curtiss, & Bird, 1964).

3.1 Boltzmann's Distribution

To start, consider the classical example of determining the thermodynamic properties of a system given the temperature, total volume and total number of molecules, N. With these conditions, the theory of quantum mechanics ensures that the states available to the system are defined. If the most probable states are found, then the properties of the system at these states are calculated, as they are assumed to reflect the equilibrium properties of the real system.

In time, the macroscopic state cycles through its available quantum states. Let the time-dependent motion of states be represented as an ensemble, a large collection of systems under the same thermodynamic conditions, where each system represents a state visited by the real system. Analysis thereby begins with gathering all the information about the ensemble. Each system in this ensemble is in a state j with a total system energy E_j^t. The number of systems that occupy the same energy state is given by a_j. The total energy of the ensemble is therefore

$$\mathfrak{E} = \sum_j a_j E_j^t \tag{3.1}$$

The number of systems in the ensemble must be very large, given the number of degeneracies discussed above. This quantity is

$$\mathfrak{A} = \sum_j a_j \tag{3.2}$$

Given system conditions and the above constraints, the most probable energy state is able to be found.

The most probable state is directly related to the distribution of states in the ensemble. The number of ways the systems in the ensemble can be distributed is a reflection of how often a distribution of states is encountered upon analysis. This number is given by the combinatorial formula

$$\Omega(\mathbf{a}) = \frac{\mathfrak{A}!}{\prod_j a_j!} \qquad (3.3)$$

where $\mathbf{a} = (a_1, a_2, ...)$. If the number of systems in the ensemble is sufficiently large, which is prescribed above, the distribution of systems corresponding to the largest number of ways effectively represents the distribution of systems. One can assume the existence of this distribution, labeled $\mathbf{a}^*$. Of course, the distribution $\mathbf{a}^*$ must be constrained by Equations (3.1) and (3.2).

The task now is to quantify $\mathbf{a}^*$. This is accomplished by maximizing the ways function (or here, the natural log of the ways function, which is entirely equivalent due to the monotonic nature of the natural log operator) using the constraints above and the method of Lagrange undetermined multipliers. Maximizing $\ln \Omega$ involves taking the derivatives of all a_j and constraining the distribution to Equations (3.1) and (3.2). The result is given as

$$\frac{\partial}{\partial a_j}\left[\ln \Omega(\mathbf{a}) - \alpha \sum_k a_k - \beta \sum_k a_k E_k^t \right] = 0 \qquad (3.4)$$

for each j. In this treatment, the undetermined multipliers are α and β, and the task will be to find what these quantities are. Solving for the elements of the most probable

distribution, a_j^*, gives a relationship between the multipliers and the energy state E_j^t. The result follows, for each j

$$a_j^* = e^{-(\alpha+1)}e^{-\beta E_j^t} \tag{3.5}$$

We can further evaluate by summing Equation (3.5) over all j and utilizing Equation (3.2). This gives an expression isolating α

$$e^{(\alpha+1)} = \frac{\sum_j e^{-\beta E_j^t}}{\mathfrak{A}} \tag{3.6}$$

The probability of finding a system in state j, given by the occupation number of state j in the most probable distribution, is found by combining Equations (3.5) and (3.6)

$$P_j = \frac{a_j^*}{\mathfrak{A}} = \frac{e^{-\beta E_j^t}}{\sum_k e^{-\beta E_k^t}} \tag{3.7}$$

This is the relationship that leads to the probability distribution of quantum states.

The function in the denominator of Equation (3.7) is a common relation within statistical mechanics and is called the canonical ensemble partition function, Q. It is defined as

$$Q(N,V^t,T) = \sum_j e^{-\beta E_j^t(N,V^t)} \tag{3.8}$$

The partition function may also be expressed by grouping energy states with the same energy and summing through the available energy levels

$$Q(N,V^t,T) = \sum_{k(\text{levels})} \Omega_k e^{-\beta E_k^t(N,V^t)} \tag{3.9}$$

where Ω_k is the degeneracy of energy level k

$$\Omega_k = \Omega\left(N, V^t, E_k^t\left(N, V^t\right)\right) \tag{3.10}$$

Therefore, the probability now of finding a system at energy level k is

$$P_k = \frac{\Omega_k e^{-\beta E_k^t}}{\sum_i \Omega_i e^{-\beta E_i^t}} \tag{3.11}$$

Probabilities in the form of Equation (3.11) and the canonical partition function, Equation (3.10), are central quantities when relating the molecular quantum states to the macroscopic properties.

3.2 Relationships between Partition Functions and Macroscopic Properties

With the expressions for the most probable energy level, the macroscopic system properties desired by engineers are nearly calculable. The last hurdle is to relate the probabilities to macroscopic properties and, in the process, to determine the second multiplier β. One must begin by considering the average energy of a system, the weighted average of all the energy states in the ensemble. This is expressed as

$$\overline{E}^t = \sum_j P_j E_j^t = \sum_j \frac{E_j^t e^{-\beta E_j^t}}{Q} \tag{3.12}$$

Recall that the total internal energy of a system, U^t, is the energy of the molecules within a system: the kinetic energy of translation; internal rotations and vibrations; and the potential energy they experience through interactions. These effects are exactly described in the quantum state, therefore Equation (3.12) is exactly the internal energy of the macroscopic system.

$$U^t = \overline{E}^t \tag{3.13}$$

Another relation between statistical thermodynamics and macroscopic thermodynamics reveals the definition of β. First, recall that the differential change in energy needed to reversibly expand a volume at a given pressure

$$dU^t = pdV^t \tag{3.14}$$

For a system within the ensemble, this relation is given by

$$dE_j^t = p_j dV^t \tag{3.15}$$

The subscript exists on the pressure here because the pressure for a given differential volume change is dependent on the system quantum state. Rearrangement isolates this pressure

$$p_j = \left(\frac{\partial E_j^t}{\partial V^t} \right)_N \tag{3.16}$$

The pressure familiar to engineers is the thermodynamic pressure, and this quantity is given as the ensemble average of the system pressures, similar to Equation (3.12)

$$p = \bar{p} = \sum_j P_j p_j = \sum_j \frac{1}{Q} \left(\frac{\partial E_j^t}{\partial V^t} \right)_N e^{-\beta E_j^t} \tag{3.17}$$

Consider now the derivative of the average energy, Equation (3.12), with respect to the volume, holding N and β constant

$$\left(\frac{\partial \bar{E}^t}{\partial V^t} \right)_{N,\beta} = -\bar{p} + \beta \frac{1}{Q} \sum_j \left(\frac{\partial E_j^t}{\partial V^t} \right)_N E_j^t e^{-\beta E_j^t} - \beta \bar{E}^t \bar{p} \tag{3.18}$$

This expression is simplified when one considers the derivative of the average pressure, Equation (3.17), with respect to β while holding N and V^t constant

$$\left(\frac{\partial \bar{p}}{\partial \beta} \right)_{N,V^t} = \bar{E}^t \bar{p} - \frac{1}{Q} \sum_j \left(\frac{\partial E_j^t}{\partial V^t} \right)_N E_j^t e^{-\beta E_j^t} \tag{3.19}$$

Equation (3.19) inserted into Equation (3.18) results in

$$\left(\frac{\partial \bar{E}^t}{\partial V^t}\right)_{N,\beta} = -\bar{p} - \beta\left(\frac{\partial \bar{p}}{\partial \beta}\right)_{N,V^t} \tag{3.20}$$

Relating this to a similar expression derived from the FPR of the internal energy

$$\left(\frac{\partial U^t}{\partial V^t}\right)_{T,N} = -p + T\left(\frac{\partial p}{\partial T}\right)_{N,V^t} \tag{3.21}$$

suggests that $\beta = \text{constant}/T$. This constant is universal for all systems considered, and it is called Boltzmann's constant, named after a founding contributor to statistical thermodynamics. The undetermined multiplier is therefore

$$\beta = \frac{1}{k_B T} \tag{3.22}$$

where the numerical value of Boltzmann's constant has been determined to be

$$k_B = 1.3807 \times 10^{-23} \text{ J/K} \tag{3.23}$$

The natural variables of the canonical partition function suggest a relationship between Q and the Helmholtz energy A. Development of the statistics and relations of the FPR in classical thermodynamics result in a straightforward relationship

$$A^t = -k_B T \ln Q \tag{3.24}$$

The pressure of a system describable by a canonical ensemble is found using the FPR of the Helmholtz energy, Equation (2.6)

$$p = -\left(\frac{\partial A^t}{\partial V^t}\right)_{T,N} = k_B T\left(\frac{\partial \ln Q}{\partial V^t}\right)_{T,N} \tag{3.25}$$

The chemical potential for a given species i, utilized within engineering models, is given by

$$\mu_i = -k_B T \left(\frac{\partial \ln Q}{\partial N_i} \right)_{T, V', N_{j \ne i}}$$ (3.26)

For the other macroscopic thermodynamic functions of interest, one must sum over the Boltzmann-like factor of interest to achieve the appropriate variables. For instance, if one wished to have the partition function depend on p instead of V', one finds

$$G'(N, T, p) = -k_B T \ln \Delta(N, T, p) = -k_B T \ln \sum_{V'} \sum_{E'} \Omega e^{-\frac{E'}{k_B T}} e^{-\frac{pV'}{k_B T}}$$ (3.27)

where Δ is called the isothermal/isobaric, or Gibbs, ensemble partition function. The summations here are over all possible system volumes and energy levels. Also of interest is the partition function the grand canonical ensemble, Ξ, where N is exchanged with the chemical potential, μ

$$pV' = k_B T \ln \Xi(\mu, T, V') = k_B T \ln \sum_{N} \sum_{E'} \Omega e^{-\frac{E'}{k_B T}} e^{\frac{\mu N}{k_B T}}$$ (3.28)

One is able to go backwards from the canonical ensemble and remove the summation over energies, thus having the thermodynamic variables N, V', and E'. This function is the degeneracy of the system of energy E' and is related to the entropy of the system through Boltzmann's formula

$$S' = k_b T \ln \Omega(N, V', E')$$ (3.29)

Since the focus of this work is around the Gibbs function and its associated partition function Δ, thermodynamic functions found using Equation (3.27) and the FPR in Equation (2.4) are of interest. The total volume of the system, important in research of an EoS, is given by

$$V^{t} = \left(\frac{\partial G^{t}}{\partial p} \right)_{T,N} = -k_{B}T \left(\frac{\partial \ln \Delta}{\partial p} \right)_{T,N} \tag{3.30}$$

The chemical potential is important within the VLE framework, and this is given by

$$\mu_{i} = \left(\frac{\partial G^{t}}{\partial N_{i}} \right)_{T,p} = -k_{B}T \left(\frac{\partial \ln \Delta}{\partial N_{i}} \right)_{T,p} \tag{3.31}$$

where N_{i} is the number of molecules of species i in the system.

To facilitate calculation of properties using the partition function approach, the maximum term within the summation is used to represent the entire partition function. For instance, consider writing the canonical partition function as

$$Q\left(N,V^{t},T \right) = \sum_{E^{t}} \tau_{Q}\left(N,V^{t},T;E^{t} \right) \tag{3.32}$$

where

$$\tau_{Q}\left(N,V^{t},T;E^{t} \right) = \Omega\left(N,V^{t},E^{t} \right) e^{-E^{t}/k_{B}T} \tag{3.33}$$

It is assumed that the terms in the summation are dominated by a single term, similar to the assumption made with the occupancy of states and Equation (3.3). The function representing the summand, τ_{Q}, is maximized with respect to energy to find this maximum term. The partition function, Equation (3.32), is therefore assumed to be

$$Q\left(N,V^{t},T \right) = \tau_{Q}\left(N,V^{t},T;E^{t*} \right) \tag{3.34}$$

where E^{t*} maximizes τ_{Q}. Similarly with the Gibbs ensemble, the summation within the partition function is expressible as the maximum term. Considering the partition function

$$\begin{aligned} \Delta\left(N,T,p \right) &= \sum_{V^{t}} \sum_{E^{t}} \Omega\left(N,V^{t},E^{t} \right) e^{-E^{t}/k_{B}T} e^{-pV^{t}/k_{B}T} \\ &= \sum_{V^{t}} \sum_{E^{t}} \tau_{\Delta}\left(N,T,p;V^{t},E^{t} \right) \end{aligned} \tag{3.35}$$

The maximum of τ_G is found by taking the derivatives with respect to E^t and V^t and setting those to zero

$$\left(\frac{\partial \tau_\Delta}{\partial V^t}\right)_{N,T,p,E^t} = 0 \text{ and } \left(\frac{\partial \tau_\Delta}{\partial E^t}\right)_{N,T,p,V^t} = 0 \tag{3.36}$$

The partition function is therefore expressible as

$$\Delta(N,T,p) = \tau_\Delta\left(N,T,p;V^{t*},E^{t*}\right) \tag{3.37}$$

thereby giving the most probable energy state and volume when considering an ensemble of systems at a given temperature, pressure and composition.

3.3 Virial Equation of State

The virial equation of state is a Taylor's series expansion of system pressure in terms of system density around the low density limit. This expansion is the only theoretically correct representation of the gas state that exists in thermodynamics. The macroscopic thermodynamic representation is given by

$$p = \frac{RT}{V}\left(1 + \frac{B}{V} + \frac{C}{V^2} + \frac{D}{V^3} + ...\right) \tag{3.38}$$

where B, C, and D are virial coefficients dependent on temperature. At subcritical temperatures, the series converges only for gas-like volumes. At supercritical temperatures, the series converges for all volumes. The number of terms necessary for an accurate representation of a system depends on the pressure and on the substance. For systems with pressures up to 20 bar, an accurate volume can usually be calculated with the inclusion of the B/V term. For system with pressures up to 50 bar, terms up to C/V^2 must be included.

To derive the virial coefficients, one must consider the grand canonical ensemble, Equation (3.28). A series expansion of the ensemble and relation to Equation (3.38) yields rigorous definitions in terms of intermolecular potentials. These definitions relate the coefficient of order $(-n)$ to the molecular interactions of systems involving n and fewer molecules. The result for B involves expressions for molecular interactions in systems with two molecules, given by

$$B(T) = -2\pi \int_0^\infty \left(e^{-u_{12}(r_{12})/k_B T} - 1\right) r_{12}^2 dr_{12} \qquad (3.39)$$

where $u_{12}(r_{12})$ is the intermolecular potential depending on the distance between molecule 1 and 2 and $r_{12} = |\mathbf{r_1} - \mathbf{r_2}|$ is the distance between molecules 1 and 2. The third virial coefficient, C, involves interactions within systems up to three molecules. This expression is somewhat more involved, since there are three binary interaction terms and one three-body interaction term. If one neglects the three-body interaction term, the expression for the third virial coefficient is simply

$$C(T) = \frac{-1}{3V} \iiint \left(e^{-u_{12}(r_{12})/k_B T} - 1\right)\left(e^{-u_{23}(r_{23})/k_B T} - 1\right)\left(e^{-u_{13}(r_{13})/k_B T} - 1\right) d\mathbf{r_1} d\mathbf{r_2} d\mathbf{r_3} \qquad (3.40)$$

Here, the integrals are taken over all possible positions of the three molecules in the system. If the three-body interaction is included in the formulation, C is then given by

$$C(T) = \frac{-1}{3V} \iiint \left[\begin{array}{l} \left(e^{-u_{12}(r_{12})/k_B T} - 1\right)\left(e^{-u_{23}(r_{23})/k_B T} - 1\right)\left(e^{-u_{13}(r_{13})/k_B T} - 1\right) \\ + e^{\frac{-u_{12}(r_{12})-u_{23}(r_{23})-u_{13}(r_{13})}{k_B T}} \left(e^{-\delta(r_{12},r_{13},r_{23})/k_B T} - 1\right) \end{array} \right] d\mathbf{r_1} d\mathbf{r_2} d\mathbf{r_3} \qquad (3.41)$$

where $\delta(r_{12}, r_{13}, r_{23})$ describes the energetics evolved from the three molecules given their binary distances. These expressions reflect that, up to pressure of 20 bar, two-body

interactions play a significant role in the gas phase thermodynamics. Up to 50 bar, three-body interactions are significant influences on the properties of the gas system.

3.4 Intermolecular Interactions

Statistical thermodynamics and lattice-fluid theory reduce the problem of finding the quantum states of a system into the problem of molecular interactions. Interactions change the system state by stabilizing or destabilizing the configurational energy. Intermolecular potentials are necessary for the calculation of virial coefficients of real fluids and in approximating the system quantum states within partition function expressions. Such forces are modeled by separating the effects into the Coulombic effects, the effects at short interaction distances and the effects at long interaction distances.

Coulombic effects dominate the interactions between molecules that have a non-zero electronic charge, e.g. systems with ionic species. The energy between two molecules with charges is given by the Coulombic potential

$$u_{qq} = \frac{q_1 q_2}{r_{12}}$$ (3.42)

where q_i represents the charge of molecule i. If the charges are of opposite sign, the molecules attract one another due to a greater reduction of system energy at shorter interaction distance. If the charges are of the same sign, the interaction energy is positive, and the molecules will repel one another due to a more stable interaction energy at larger distances.

Short-range interactions between neutral molecules are repulsive. This is a manifestation of the Pauli exclusion principle, which states that quantum particles (here,

electrons) cannot occupy the same quantum state. For closed-shell molecules, the electron clouds around the nuclei overlap, thereby occupying the same space and attempting to occupy the same quantum state. This is energetically unfavorable, therefore the molecules repel. However when the electron clouds overlap, electrons have more space to occupy, which energetically is favorable. This results in an attractive force, albeit small. This is called the exchange contribution and cannot be described classically, since such an effect has no analogue in the particle description of electrons. The combined attractive and repulsive effects at short-range is sometimes called the exchange-repulsion effect.

Long-range interactions between neutral molecules are attractive. These effects are further separable into the electrostatic, induction and dispersion contributions. The electrostatic contribution involves the interactions of multipole moments (charge, dipole, quadrupole and higher moments), which arise from the asymmetrical distribution of the negatively charged electron cloud and the positively charged nuclei. Induction is when the multipole moments of one molecule induce a shift in the electron cloud of the second molecule, thus causing a short-lived attraction similar to the interaction between dipoles. Dispersion arises when one electron cloud induces another to shift, and in response, the shifted electron cloud causes the originating cloud to shift, thus creating an alternating motion that leads to short-lived dipole moments of opposite orientation. The latter two contributions are influenced by the polarizability of the molecule, which is the measure of how easily swayed the electron cloud is when under the influence of an external electric field.

The electrostatic interactions are calculated using Coulomb's law, Equation (3.42). Partial charges are placed within the molecule to reproduce the multipole moments. This results in orientation-dependent expressions for intermolecular interaction energies. For instance, the charge-dipole interaction is given by

$$u_{q\mu} = -\frac{q_1 \mu_2}{r_{12}^2} \cos \varphi_2 \qquad (3.43)$$

where μ_i is the magnitude of the dipole moment of molecule i, and φ measures the angle of the dipole off the line connecting the partial charge and the center of the dipole vector. The dipole-dipole interaction is given by

$$u_{\mu\mu} = -\frac{\mu_1 \mu_2}{r_{12}^3} \left[2 \cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2 \cos(\theta_1 - \theta_2) \right] \qquad (3.44)$$

where θ is the angle of the dipole moment off the plane perpendicular to the line connecting the centers of the dipole moments.

For long-range interactions, molecules can rotate without being influenced by the potential wells of certain stable orientations between molecules. These interactions may be considered as orientation-averaged interactions. This is done by integrating Boltzmann factors over all orientations to determine a partition function for the system, and then dividing the partition function weighted by the energy, similar to Equation (3.12). In general, the orientation-averaged interaction energy for an intermolecular potential u_{xy} is given by

$$\bar{u}_{xy}(r) = \frac{\iint u_{xy} e^{-u_{xy}(r)/k_B T} d\omega_1 d\omega_2}{\iint e^{-u_{xy}(r)/k_B T} d\omega_1 d\omega_2} \qquad (3.45)$$

where $\omega_i = \sin\theta_i d\theta_i d\varphi_i$. For large enough distances or high enough temperatures, the Boltzmann factors can be expanded in terms of $1/k_BT$

$$e^{-u_{xy}/k_BT} = 1 - \frac{u_{xy}}{k_BT} + \frac{1}{2}\left(\frac{u_{xy}}{k_BT}\right)^2 - \cdots \tag{3.46}$$

The result when applying the charge-dipole interaction of Equation (3.43) into Equation (3.45) yields

$$\bar{u}_{q\mu}(r_{12}) = -\frac{1}{3k_BT}\frac{q_1^2\mu_2^2}{r_{12}^4} \tag{3.47}$$

For the dipole-dipole interaction, Equations (3.44) and (3.45) give

$$\bar{u}_{\mu\mu}(r_{12}) = -\frac{2}{3k_BT}\frac{\mu_1^2\mu_2^2}{r_{12}^6} \tag{3.48}$$

The induction contribution is dependent on the magnitude of the electric multipoles and the polarizability of the electron cloud of the affected molecule. All orders of multipole moments can induce a dipole in an electron cloud. Induction caused by a charge results in the intermolecular potential describable by

$$u_{q\alpha} = -\frac{q_1^2\alpha_2}{2r_{12}^4} \tag{3.49}$$

where α_i is the polarizability of molecule i with the induced moment. Induction caused by a dipole moment is given by

$$u_{\mu\alpha} = -\frac{\mu_1^2\alpha_2\left(3\cos^2\varphi_1 + 1\right)}{2r_{12}^6} \tag{3.50}$$

The orientation-averaged expression is found by applying Equation (3.45). The result of this is

$$\overline{u}_{\mu\alpha} = -\frac{\mu_1^2 \alpha_2}{r_{12}^6}$$

(3.51)

Dispersion interactions are significant because they are the main force of attraction between non-polar molecules. Spherically symmetric molecules do condense into a liquid phase; therefore they do have an interaction that results in attraction. These interactions have been deemed London-type dispersive interactions and are named after the scientist who applied quantum mechanical treatments to explain this phenomenon (Gray and Gubbins, 1984). The interaction is described by

$$u_{\alpha\alpha} = -\frac{3}{2} \frac{\alpha_1 \alpha_2}{r_{12}^6} \left(\frac{h\nu_{0,1} h\nu_{0,2}}{h\nu_{0,1} + h\nu_{0,2}} \right)$$

(3.52)

where h is Planck's constant and $\nu_{0,i}$ is a characteristic ground-state electronic frequency of molecule i. If one assumes that $h\nu_{0,i}$ is roughly described by the ionization potential, $I_{0,i}$, then the relation becomes

$$u_{\alpha\alpha} = -\frac{3}{2} \frac{\alpha_1 \alpha_2}{r_{12}^6} \left(\frac{I_{0,1} I_{0,2}}{I_{0,1} + I_{0,2}} \right)$$

(3.53)

To calculate the interaction energy between two molecules, the effects described above are summed together

$$u\left(\mathbf{r}_1, \mathbf{r}_2\right) = u_{ex/rep} + u_{electrostatic} + u_{induction} + u_{dispersion}$$

(3.54)

where the Coulombic interaction is included in the electrostatic contributions and, in general, the orientation-dependent expressions are used.

3.5 Summary

Statistical thermodynamics offers engineers a bridge between the molecular description of matter and the macroscopic description of systems. The partition function and the probability distribution of quantum states allows for a fuller understanding of how molecular-level properties, such as state distributions and interaction energies, affect volumetric properties and chemical potentials.

The volumetric behavior of a gaseous system is shown to be directly influenced by the characteristics of interaction potentials involving only a few molecules. Molecular properties used in these interaction formulas, such as the multipoles and polarizabilities, have in the past been found through experiment. More recently, these properties can be calculated numerically by utilizing quantum chemistry from its first principles to give electron density profiles of molecules and how they may be influenced by electric fields. These properties are becoming accurately represented through quantum chemistry, in turn making such information more readily accessible to engineers for use in these interaction formulas and within partition functions.

CHAPTER 4

LATTICE-FLUID THEORY

The amount of information necessary to use the general partition function approach in the description of fluid systems is immense. Even for the system in the gas phase, the ease of the virial equation series solution is lost as one considers multiple body interactions. Ideally, the partition function approach would describe all states available to a fluid, whether in the gas phase or the liquid phase. To accomplish this, assumptions must be made to consider a smaller number of dominant molecular effects most important to the macroscopic description of the system.

One method for describing fluids within a system space is to imagine them occupying a lattice. Lattice statistics make available closed-form, analytical equations describing the dominant effects within the partition function formulation. The details of the lattice need not be explicit, as long as the sites are uniformly distributed throughout the space, whether on fixed coordinates or not. Methods applied within lattice-fluid theory make the description of fluids more tractable and result in qualitative descriptions of vapor-liquid equilibrium for pure species and mixture systems. These theoretical methods are also flexible enough to incorporate properties of real molecules, thus making the models useful in the description of real systems.

This chapter focuses on lattice-fluid concepts and current approaches to the fluid problem. Fluids on lattices have been modeled both by fully occupying the lattice and by occupying a lattice also occupied by vacancies. Basic assumptions to simplify the partition functions for use within the theory are presented for a system with a fully

occupied lattice. A combinatorial approach to the entropy neglecting contributions from the interaction energies, thus describing the random state of the fluid, is formulated. The interaction energies are then included and adjustments are made to describe the system in the more realistic, nonrandom state. The same process is followed for systems where vacancies are included in the lattice, and the volumetric properties of lattice fluids are considered.

Lattice-fluid theory is the foundation for the application of the molecular-level structural and energetic properties later in this work. It is here that both the pure species equation of state and the activity coefficient of species in solution are able to be calculated using the same model. The goal is to make available a single set on modeling equations that engineers can employ in the study of pure and mixture fluid systems.

4.1 Partition Functions of Lattice-Fluid Theory

The initial systems considered involve fully occupied lattices, since it is thought that such systems serve as reasonable approximations to the liquid state. Since the number of lattice sites is known from the beginning, the volume of the system (a liquid-like volume) is thereby known also. These models attempt to evaluate the terms within Equation (3.9), the canonical partition function.

The lattice-fluid partition function is assumed to be simpler than the general canonical partition function. Consider a lattice-fluid system where only two different species exist, molecules of species A and molecules of species B. Instead of being interested in the entire system energy, only the total configurational energy W' is considered within the partition sum (Hill, 1960)

$$Q\left(N_A, N_B, V^t, T\right) = Q_{internal} \sum_{W^t} \Omega e^{-W^t/k_B T} \tag{4.1}$$

where $Q_{internal}$ represents the contribution of the energy modes internal to the molecule and assumed separable from configurational effects. The total configurational energy W^t is given by the interaction of nearest neighbors. For a binary system,

$$W^t = \sum_i \sum_j N_{ij} \varepsilon_{ij} = N_{AA} \varepsilon_{AA} + N_{AB} \varepsilon_{AB} + N_{BA} \varepsilon_{BA} + N_{BB} \varepsilon_{BB} \tag{4.2}$$

where N_{ij} is the number of nearest-neighbor interactions between molecules i and j, and where ε_{ij} is the corresponding interaction energy. Although separated explicitly, it is assumed that $N_{ij} = N_{ji}$ and $\varepsilon_{ij} = \varepsilon_{ji}$. When the maximum term method described in Equation (3.34) is applied, the notation within the partition function reduces

$$Q = Q_{internal} \Omega e^{-W^t/k_B T} \tag{4.3}$$

where the $*$ associated with the maximum term approximation is henceforth omitted.

The systems later considered are those lattice fluids that include vacancies. The number of vacancies is considered a variable that is to be determined by system temperature and pressure, as well as implicitly by the nature of molecules mixed. Thereby, the Gibbs ensemble in Equation (3.27) is used, and the maximum term approximation is taken as above

$$\Delta\left(N_A, N_B, p, T\right) = \Delta_{internal} \Omega e^{-W^t/k_B T} e^{-pV^t/k_B T} \tag{4.4}$$

The total volume of the system is given by the volumes of the molecules and unoccupied lattice sites

$$V^t = \sum_i N_i b_i = N_0 b_0 + N_A b_A + N_B b_B \tag{4.5}$$

where b_i is the volume of molecule i, N_0 is the number of vacancies on the lattice, and b_0 is the volume of a vacant lattice site.

These interpretations of the partition functions allow the problem to be reduced to determining the most favorable set of interaction numbers N_{ij}, the number of vacancies N_0 and the degeneracy of the system given that interaction profile. The latter problem is now addressed.

4.2 The Ideal Solution and the Athermal Mixture

The first attempt to determine the degeneracy is to consider systems where the interaction energies have no bearing on the thermodynamics. These systems are called athermal, where either $W' = 0$ or the configurational energies of the pure systems being mixed and the resulting mixture systems are the same (Guggenheim, 1944a), as in the case of mixing two alcohol species of similar sizes. The only contributions to the partition function in Equation (4.3) here are the internal term and the degeneracy term

$$Q = Q_{internal}\Omega^{ath} \tag{4.6}$$

Since the internal contributions will cancel upon mixing, they will be left out of subsequent expressions.

An ideal solution describable by Equation (2.15) considers a system where the geometries of the molecules are so similar that size and shape effects have a negligible contribution to the system entropy. Therefore, simple mixing is considered in the degeneracy, and the partition function is given by (Hill, 1960)

$$Q = \Omega^{ath} = \frac{N!}{N_A!N_B!} \tag{4.7}$$

This formula is the same as a model problem in the field of combinatorial analysis. This is the value of the number of ways one can draw $N = N_A + N_B$ balls from an urn with two different color balls, N_A balls of one color and N_B balls of the other. The system is said to possibly exist in Ω ways, since this is the number of ways the molecules can exist in the system while yielding an equivalent system energy.

For polymer systems where the size and shape differences between the macromolecule and the solvent are large, the degeneracy described by Equation (4.7) does not adequately depict the thermodynamics of the athermal mixture. This problem must be approached by considering the probabilities of placing a polymer onto a lattice systematically. These probabilities are then computed as a function of lattice-site occupation, assuming the other species on the lattice is a monomer (or nonexistent, therefore considered to be vacancies). An original assumption in this treatment is that the probability of placing the next segment of a polymer on a lattice is assumed to be the fraction of unoccupied sites within the entire system, without consideration as to whether an adjacent site is available for a monomer to be placed (Flory, 1942). The number of ways N_A monomers and N_B polymers occupying r_B segments is placed on the lattice is given by

$$\Omega_{Flory}^{ath} = \left(\frac{z_0 - 1}{N_t}\right)^{(r_B-1)N_B} \left(\frac{r_B^{r_B N_B}}{2^{N_B} N_B!}\right)\left[\frac{(N_t/r_B)!}{(N_t/r_B - N_B)!}\right]^{r_B} \tag{4.8}$$

where z_0 is the coordination number of the lattice and $N_t = N_A + r_B N_B$ is the total number of lattice sites in the system.

A more general treatment of the mixture of a macromolecular species and monomers has been conducted (Guggenheim, 1944a). The approach analyzes the

mixture of molecules composed of segments of different sizes as they occupy a lattice

The ways function for any number of different species on a lattice is given by

$$\Omega_{Guggenheim}^{ath} = \prod_i \varpi_i^{N_i} \frac{N_t!}{\prod_i N_i!} \left(\frac{Q_t!}{N_t!}\right)^{z_0/2} \tag{4.9}$$

where ϖ_i is a symmetry factor for molecule i, Q_t is related to the total number of

nearest-neighbor interactions in the system

$$Q_t = q_A N_A + q_B N_B \tag{4.10}$$

and q_i is such that $z_0 q_i = z_i$ is the number of external contacts for molecule i. Within

the original derivation, q_i is defined through the relation

$$\frac{z_0}{2}(r_i - q_i) = r_i - 1 \tag{4.11}$$

Equation (4.9) has been considered as a generalization of Equation (4.8), where the Flory

expression is recovered from the Guggenheim expression by allowing for the lattice

coordination number to approach infinity, $z_0 \rightarrow \infty$ (Sanchez and Lacombe, 1976). This

is the case because of the placement probabilities assumption described above (Sayegh

and Vera, 1980); to guarantee that there is always an adjacent lattice site available when

placing a polymer onto a lattice, that site must be a nearest neighbor to all the sites on the

lattice. Also, Equation (4.8) has been shown to be recovered from Equation (4.9) at the

low volume fraction occupancy of a given species in the presence of vacant lattice sites

(Martinez, 1995).

The relationship between the number of external contacts and the number of

occupied lattice sites, here given by Equation (4.11), assumes that each monomer of the

polymer occupies a single lattice site, that exactly two contact (one from each monomer)

is used to connect the monomers, and that the molecules are not cyclic. A factor is included into Equation (4.11) that allows for the description of molecules that violate the assumptions above (Staverman, 1950). This more general expression is

$$l_i = \frac{z_0}{2}(r_i - q_i) - (r_i - 1) \tag{4.12}$$

where l_i quantifies the bulkiness of the molecule. The revised degeneracy of the system described with Equation (4.12) is

$$\ln \Omega_{Staverman}^{ath} = \ln \Omega_{Guggenheim}^{ath} + \sum_i l_i N_i \ln N_t \tag{4.13}$$

A review of expressions that describe an athermal mixture (Sayegh and Vera, 1980) suggest the use of Equation (4.13) over the expressions given by Equations (4.8), (4.9), and others (Tompa, 1956; Donohue and Prausnitz, 1975). It is shown that Equation (4.13) offers an accurate description of systems describable by more complicated methods, such as those based within computational thermodynamics, while still retaining a simple form. Furthermore, it has been shown that Equation (4.13) improves the calculation of cavity formation free energies in liquid systems (Lin and Sandler, 1999a, Lin and Sandler, 1999b), and has been chosen to represent the athermal contributions within COSMO-based models (Lin and Sandler, 2002; Klamt, et al. 2002).

A generalized derivation of Equation (4.9) without assuming Equations (4.11) or (4.12) is offered in Appendix A. The motivation is to make available an expression without defining an arbitrary value such as l_i, which usually becomes a correlating parameter in engineering models of real fluids.

4.3 The Nonathermal Mixture

The nonathermal mixture considers the interaction energies between the molecules in the above statistics. However, to determine the maximum term as given by Equation (4.3), one must determine the interaction numbers that gives the maximum of the expression $\Omega e^{-W'/k_B T}$. This is not straightforward because of the difficulties of describing the entropy of a real system, even though the athermal configurational energy is explicit. Although the only closed-form representation of the entropy term up to this point is that of the athermal case, this is enough to determine the entropy of the nonathermal mixture.

The original statements of the entropy consider the degeneracy strictly a function of the number of molecules. For the binary system

$$\Omega = \Omega^{ath}\left(N_A, N_B\right) = \Omega^{ath}\left(N_i\right) \tag{4.14}$$

It is assumed that the degeneracy is a function of the interaction numbers also (Guggenheim, 1944b)

$$\Omega^{ath}\left(N_i\right) \rightarrow \Omega\left(N_i, N_{ij}\right) \tag{4.15}$$

This assumption makes the athermal degeneracy, as it stands, not adequate to incorporate interaction effects.

Here, a guess at the form of $\Omega\left(N_i, N_{ij}\right)$ is made. Consider a binary mixture system describable by the canonical ensemble. Consider also a fully occupied lattice. There exists a number of interactions that will occur on this lattice, dependent on the number of molecules A and B, the size and bulkiness of the molecules, and the number of interactions that A and B participate in. Let I represent the total number of interactions, given by

$$I = \sum_i \sum_j N_{ij} = N_{AA} + N_{AB} + N_{BA} + N_{BB} \tag{4.16}$$

These I interactions may be depicted as a lattice that is superimposed on the lattice that molecules occupy. These interaction lattice points thereby are occupied by interactions denoted AA, AB, BA, and BB. A first assumption is to state that these interaction points are distributable as molecules are on a lattice. The number of ways by this assumption is given by an expression similar to Equation (4.7), which is

$$\Omega'(N_i, N_{ij}) = \frac{I!}{\prod_i \prod_j N_{ij}!} = \frac{I!}{N_{AA}! N_{AB}! N_{BA}! N_{BB}!} \tag{4.17}$$

A descriptive argument as to why this method overcounts the number of ways is documented (Hill, 1960).

Another assumption on the form of $\Omega(N_i, N_{ij})$ must be made. It is assumed that the degeneracy is separable into two contributions: one strictly in terms of the numbers of molecules, the other in terms of the first assumption described above (Hill, 1960)

$$\Omega(N_i, N_{ij}) = K(N_i)\Omega'(N_i, N_{ij}) \tag{4.18}$$

The proportionality factor K is found by normalizing the sum over all possible sets of N_{ij}, which should equate to the athermal contribution above

$$\sum_{\text{all sets of } N_{ij}} \Omega(N_i, N_{ij}) = \Omega^{ath}(N_i) = K(N_i) \sum_{\text{all sets of } N_{ij}} \Omega'(N_i, N_{ij}) \tag{4.19}$$

This process is generally described for molecules (Guggenheim, 1944b) and interacting surface segments (Kehiaian, et al. 1978). The proportionality factor K, for any number of molecules in the system, is found explicitly using Lagrange undetermined multipliers (Knox, et al. 1984). The expression is given by

$$K(N_i) = \Omega^{ath} \left[\frac{\prod\limits_i (z_i N_i/2)!}{I!} \right]^2 \qquad (4.20)$$

where I is expressible through the numbers of molecules and their numbers of interactions z_i

$$I = \sum_i \frac{z_i N_i}{2} \qquad (4.21)$$

and where the number of interactions attributable to species i is given by

$$\frac{z_i N_i}{2} = \sum_j N_{ij} \qquad (4.22)$$

Therefore, the partition function that describes the lattice fluid where any number of species occupies all sites is given by (Knox, et al. 1984)

$$Q(N_i, V^t, T) = Q_{internal} \sum_{\text{all sets of } N_{ij}} \Omega^{ath}(N_i) \frac{\left[\prod\limits_i (z_i N_i/2)! \right]^2}{I! \prod\limits_i \prod\limits_j N_{ij}!} e^{-W^t(N_{ij})/k_B T} \qquad (4.23)$$

With this partition function, the problem of maximizing the term within the summation is considered. The objective is to find the set of N_{ij} (for a binary system, the N_{AA}, N_{AB}, N_{BA} and N_{BB}) that maximizes the term. These numbers are under the constraints given by Equations (4.2), (4.16), (4.22) and by the equality

$$N_{ij} = N_{ji} \qquad (4.24)$$

This problem has been solved indirectly (Guggenheim, 1944b; Kehiaian, 1978) and through the using Lagrange undetermined multipliers (Knox, et al. 1984; Knox, 1987). It is found that the numbers of interactions in the nonathermal system must satisfy what is called the quasi-chemical relationship

$$\frac{N_{ij}N_{ji}}{N_{ii}N_{jj}} = e^{-\frac{2\varepsilon_{ij}-\varepsilon_{ii}-\varepsilon_{jj}}{k_B T}} \qquad (4.25)$$

where the quantity within the exponential is referred to as the interchange energy. Once this has been established, all unknowns within the system of equations can be solved for. The derivation for the quasi-chemical equations in its most general form is offered in Appendix B. The solution to these equations then gives thermodynamic properties of the lattice-fluid system through the techniques described in Section 3.2.

4.4 Vacancies on a Lattice and Equations of State

The premise under most of the early work within lattice-fluid theory centers on the description of polymer fluids. The polymer is modeled as a large number of segments, each occupying one lattice site, while another, monomer-sized species, occupied the remaining sites. Since these models consider the athermal solution, the energies of interaction between monomer-polymer segments are treated as equivalent to polymer-polymer interactions. Therefore, these monomers have been considered vacancies or small solvent molecules by those developing the theories. Assuming these monomers are vacancies, it has been shown that statistics involving the surface area contacts (Guggenheim, 1944a; Staverman, 1950) are a general theoretical treatment of this problem, while earlier works (Flory, 1942; Huggins, 1942) are approximations at low polymer density (Martinez, 1995).

Use of the Gibbs partition function, Equation (4.4), allows for the effect of vacancies on a lattice to be explicit. The difference to the canonical partition function approach is that, to eliminate the summation over all volumes as given in Equation (3.35),

the number of vacancies N_0 must also maximize the term. It turns out that this process reveals the relation between the pressure, volume, and temperature of the system, thus gives an equation of state. Recall that using the Gibbs partition function allows for a description of the volumetric properties of the fluid, given by Equation (3.30).

Allowing vacancies on a lattice has become an interesting technique in modeling volumetric properties for pure species and mixture systems. One way is to consider the vacancies randomly distributed throughout the lattice (Sanchez and Lacombe, 1976). This approach begins with the athermal statistics given by Equation (4.9) and considers a binary system of molecules and vacancies

$$\Omega^{ath} = \varpi^N \frac{(N_0 + rN)!}{N_0! N!} \left[\frac{(N_0 + qN)!}{(N_0 + rN)!} \right]^{z_0/2} \tag{4.26}$$

where ϖ, N, r, and q are quantities of the pure species in the system. Using Stirling's approximation and taking the coordination number $z_0 \to \infty$, a simplified form of the degeneracy emerges

$$\lim_{z_0 \to \infty} \Omega^{ath} = \left(\frac{1}{\phi} \right)^{N_0} \left(\frac{r\varpi}{e^{r-1}} \frac{1}{\phi} \right)^N \tag{4.27}$$

where ϕ is the volume fraction of the species on the lattice

$$\phi = \frac{rN}{N_0 + rN} \tag{4.28}$$

Equation (4.27) corresponds to the ways given by the Flory degeneracy in Equation (4.8). Inserting Equations (4.2), (4.5), and (4.27) into Equation (4.4) and maximizing with respect to N_0 gives the EoS (Sanchez and Lacombe, 1976)

$$\tilde{\rho}^2 + \tilde{p} + \tilde{T}\left[\ln\left(1-\tilde{\rho}\right)+\left(1-\frac{1}{r}\right)\tilde{\rho}\right]=0 \qquad (4.29)$$

where the dimensionless variables are given by

$$\tilde{\rho} = \frac{Nrb_0}{V} \qquad (4.30)$$

$$\tilde{p} = \frac{2}{z_0}\frac{pb_0}{\varepsilon} \qquad (4.31)$$

$$\tilde{T} = \frac{2}{z_0}\frac{k_BT}{\varepsilon} \qquad (4.32)$$

where ε is the interaction energy between two molecules.

Further work of including vacancies in the lattice includes them in a nonrandom manner (Smirnova and Victorov, 1987; Taimoori and Panayiotou, 2001; Panayiotou, 2003b), where the notation in the subsequent expressions reflects the latter references.

The vacancies recently have been considered a separate species in the lattice-fluid mixture, occupying space and interacting with nearest neighbors. These interactions contribute no energy to the configuration of the lattice. The resulting EoS when the entropy of Equation (4.27) is applied is given by

$$\tilde{p} + \tilde{T}\left[\ln\left(1-\tilde{\rho}\right)+\left(1-\frac{1}{r}\right)\tilde{\rho}+\frac{s}{2}\ln\Gamma_{00}\right]=0 \qquad (4.33)$$

where the reduced variables are the same as above, s is the average number of contacts for a lattice site, and Γ_{00} is a nonrandomness factor found by solving the following equations (Taimoori and Panayiotou, 2001)

$$\left(1-\tilde{\rho}\right)\Gamma_{00} + \tilde{\rho}\Gamma_{0r} = 0 \qquad (4.34)$$

$$\Gamma_{0r} = \frac{2}{1+\left[1-4\tilde{\rho}(1-\tilde{\rho})(1-\tau^2)\right]^{1/2}} \qquad (4.35)$$

$$\tau = e^{\varepsilon/2k_BT} \qquad (4.36)$$

When $\Gamma_{00} = 1$, the vacancies are said to be distributed randomly, as in the case described above. For the case where Equation (4.26) describes the entropy, the EoS is given by (Smirnova and Victorov, 1987; Panayiotou, 2003b)

$$\tilde{p} + \tilde{T}\left[\ln(1-\tilde{\rho}) + \frac{z}{2}\ln\left(1+\frac{q}{r}\tilde{\rho}-\tilde{\rho}\right) + \frac{z}{2}\ln\Gamma_{00}\right] = 0 \qquad (4.37)$$

4.5 Summary

Lattice-fluid theory allows for the simplification of the fluid problem. The assumption that the internal molecular modes are independent of the other molecules in the system allows for focus to be applied on the numbers and types of interactions in the partition function. The reduced problem of the athermal system gives the foundation for expressions that describe the nonathermal system. Vacancies are included in the lattice in the attempt to describe both liquid and gas phases with the same model.

Lattice-fluid theory includes molecular properties, such as volumes and surface area (reflected in r and q) that can be evaluated approximately through correlation or rigorously calculated from quantum chemistry. Also, the energetics of the system depends primarily on nearest-neighbor binary interactions, which allows again for descriptions ranging from empirical to theoretical. Structural, electrostatic and energetics quantities will be the focus of a large majority of this work, after an assessment of the ways these properties have been formulated in engineering models.

CHAPTER 5

ENGINEERING SOLUTIONS TO THE
VAPOR/LIQUID EQUILIBRIUM PROBLEM

To make the classical and statistical theory useful for applications, engineers elucidate concepts to bridge the molecular level and the macroscopic system level. Engineers need not know the absolute values of internal energy and entropy, but how these properties behave when a system changes temperature and pressure. Approximations must be made at the molecular level to separate the dominating effects from the incidental effects that most affect changes at the macroscopic level.

To accomplish this, engineers have developed phenomenological models to account for the most important molecular-level effects. The VLE problem has served as a measure of how well a model encapsulates the dominant effects, mainly because errors may cancel out in the application of Equation (2.22).

Assumptions are made to find a closed-form expression for the excess Gibbs energy. Firstly, although encountered in practice and explicit in the mathematical framework described above, engineers have assumed that the excess volume is negligible in a mixing process. This allows for the for the excess Gibbs energy to be approximated by the excess Helmholtz energy

$$G^E \simeq A^E \tag{5.1}$$

Therefore, the total Gibbs energy is approximately the total Helmholtz energy and its corresponding partition function

$$G^t \simeq A^t = -k_B T \ln Q \tag{5.2}$$

Secondly, the quantum energy levels within the isolated molecules of the system are assumed to not change through the mixing process, as per the assumption made within Section 4.1. This allows for the separation of the partition function as mentioned by Equation (4.3).

The engineering models presented in the chapter take these assumptions and describe the entropic and energetic effects of mixing, which in turn yields VLE behavior. Lattice-fluid modeling is employed due to the ease with which system level properties are described using statistics. The first part of this chapter involves the combinatorial entropy contribution to the excess Gibbs energy formulas. The local composition concepts are introduced, and early attempts to include contributions of interaction energies are presented. Theoretical problems within these approaches have arisen as these methods have been studied, and these are reviewed here. The more rigorous lattice-fluid theory in Chapter 4, the quasi-chemical method, has been developed in response to these criticisms, and both molecular and functional group models that utilize the theory are presented. Finally, the COSMO-based methods are introduced due to their recent acceptance by the engineering community. The identity between the quasi-chemical statistics and the COSMO-based statistics is established.

5.1 Athermal Effects and Engineering G^E Equations

For the simplest case, consider a mixture resulting in an ideal solution. The sizes and geometries of the species are the same, and W' for the mixture system and for the pure systems are the same. A mixture system that is not affected by changes in W' upon

mixing is called an athermal mixture. The difference between the Gibbs energy of the mixture system and the weighted averages of the pure systems is found to be

$$G - \sum_i x_i G_i = \sum_i x_i \ln x_i \qquad (5.3)$$

This is, of course, the identical expression given for an ideal solution, given by Equation (2.15).

For mixtures of molecules that are of different sizes, the composition is not considered the greatest factor in determining the entropy. The occupancy of the lattice is considered more important, and the entropy of mixing is determined to be the number of ways the lattice sites are occupied (Flory, 1942; Huggins, 1942). Assuming an athermal mixture, the difference in mixture and pure Gibbs energies is found to be

$$G - \sum_i x_i G_i = RT \sum_i x_i \ln \phi_i \qquad (5.4)$$

where ϕ_i is the volume fraction of species i in the system. Equation (5.4) reduces to the ideal solution expression in Equation (5.3) if all the sizes of the molecules are the same; mathematically, this means that $r_i = r_j$ for all the molecules in the system.

A more rigorous approach in determining the entropy from a mixing process for an athermal mixture is accomplished by analyzing the mixture of molecules composed of segments of different sizes as they occupy a lattice (Guggenheim, 1944a). The resulting difference in mixture and pure Gibbs energies of the athermal mixture system is

$$G - \sum_i x_i G_i = RT \sum_i x_i \ln \phi_i + RT \sum_i x_i \frac{z_0 q_i}{2} \ln \frac{\theta_i}{\phi_i} \qquad (5.5)$$

where θ_i is the surface area fraction of species i in the system, given by

$$\theta_i = \frac{(z_0 q_i) N_i}{\sum_j (z_0 q_j) N_j} = \frac{q_i N_i}{\sum_j q_j N_j} \tag{5.6}$$

Note the first term of Equation (5.5) is identical to that of the Flory expression, while the last term contains effects due to surface area differences.

5.2 Residual Effects and Engineering G^E Equations

The athermal system expressions above give idealized descriptions on which the effects within a real system, namely the interaction energies, are built. The ideal basis considers fluids existing in a random state, while the influence of interaction energies brings the fluid into a nonrandom state. It is assumed in the following models that the interaction energies from nearest-neighbor interactions are the only significant contributions to W', and these may be described by Equation (4.2). This affects the local composition of molecules around a particular molecule, which is different than the bulk composition and depends on magnitudes and signs of interaction energies.

The first widely used equation that exploits the local composition concept is the Wilson equation (Wilson, 1964). Here, the ideal solution relation of Equation (5.3) is used as a basis, where the latter composition is replaced by the local composition

$$G - \sum_i x_i G_i = \sum_i x_i \ln x_{ii} \tag{5.7}$$

where x_{ij} represents the local composition of j molecules surrounding i molecules. The local compositions are assumed to be related to the actual composition by the following relation

$$\frac{x_{ij}}{x_{ii}} = \frac{x_j V_j^l}{x_i V_i^l} e^{-(\varepsilon_{ij} - \varepsilon_{ii})/k_B T} \tag{5.8}$$

This equation has been used by engineers to correlate VLE data with the $n(n-1)$ parameters, where n is the number of species in a system. Each parameter contains the exponential in Equation (5.8). Therefore, for a binary system, the two parameters are

$$P_{12} = e^{-(\varepsilon_{12}-\varepsilon_{11})/k_B T}, P_{21} = e^{-(\varepsilon_{21}-\varepsilon_{22})/k_B T} \tag{5.9}$$

For a ternary mixture, there are six parameters: 1-2; 1-3; 2-1; 2-3; 3-1; 3-2.

The Universal Quasi-Chemical (UNIQUAC) equation (Abrams and Prausnitz, 1975) is an oft-used model that combines the athermal contribution from Guggenheim, Equation (5.5), and a residual contribution, modeled by a modified Wilson equation

$$G - \sum_i x_i G_i = RT \sum_i x_i \ln \phi_i + RT \sum_i x_i \frac{z_0 q_i}{2} \ln \frac{\theta_i}{\phi_i} + RT \sum_i x_i q_i \ln \frac{\theta_{ii}}{\theta_i} \tag{5.10}$$

where θ_{ij} is the local surface area fraction of j molecules around i molecules. This quantity is given in an expression similar to the local composition, Equation (5.8), except for the factor of $z_0/2$ within the exponential function

$$\frac{\theta_{ij}}{\theta_{ii}} = \frac{\theta_j}{\theta_i} e^{-\frac{z_0}{2} \frac{\varepsilon_{ij}-\varepsilon_{ii}}{k_B T}} \tag{5.11}$$

Similar to Wilson's equation, there are $n(n-1)$ parameters that are correlated to experimental data.

In conjunction with the development of UNIQUAC, a group-contribution method called the UNIQUAC Functional-group Activity Coefficients (UNIFAC) method (Fredenslund, et al. 1975) takes most of the same features as UNIQUAC except that the residual contribution focuses on the interactions between functional groups. These functional group interaction parameters, similar to those in Wilson and UNIQUAC, are

correlated to known experimental systems and used in systems that lack experimental data. The expression for the difference in Gibbs energies is given by

$$G - \sum_i x_i G_i = RT \sum_i x_i \ln \phi_i + RT \sum_i x_i \frac{z_0 q_i}{2} \ln \frac{\theta_i}{\phi_i}$$
$$+ RT \sum_k Q_k y_k \ln \frac{y_{kk}}{y_k} - RT \sum_i x_i \sum_k Q_k y_k^{(i)} \ln \frac{y_{kk}^{(i)}}{y_k^{(i)}}$$

(5.12)

where k cycles through the different groups in the system, Q_k is the surface area parameter of a group analogous to q_i, y_k is the composition of group k in the system, y_{kl} is the local composition of group l around group k, and the superscript $^{(i)}$ denotes the quantity as it exists in pure species i. The local composition in this functional group framework is found by Equation (5.11), which is similar to the UNIQUAC model

$$\frac{y_{kl}}{y_{kk}} = \frac{y_l Q_l}{y_k Q_k} e^{-\frac{\varepsilon_{kl} - \varepsilon_{kk}}{k_B T}}$$

(5.13)

The parameters within UNIFAC the form of those in Equation (5.9). These are found through correlation of known experimental results and are assumed to be transferable, which then allows the model to be applied to systems with no experimental VLE data.

The UNIFAC model has been the most widely used predictive tool for VLE and other engineering systems involving phase equilibrium. The number of groups within the method grows as experimental data for new systems are made available. The Dortmund Data Bank has been employed to fit parameters for 45 major group definitions with temperature dependence within a modified UNIFAC method (Gmehling, et al. 1993). The modification increases the number of energetic parameters for each pairing of functional groups to six, where it is two in the original UNIFAC method. A recent

assessment reveals that the total numbers of parameters for the UNIFAC and modified

UNIFAC methods are 168 and 612, respectively (Lin and Sandler, 2002).

5.3 Criticisms of the UNIQUAC/UNIFAC Method

The local composition concept employed within the UNIQUAC/UNIFAC method and its

various implementations have been scrutinized. Of foremost concern within these

models has been the violation of a basic mass-balance constraint (McDermott and Ashton,

1977), which resides within the definition of the local composition. When it is related to

N_{ij}, the number of interactions between molecule i and molecule j , x_{ij} is found to be

$$x_{ij} = \frac{N_{ij}}{z_0 q_i N_i / 2} \tag{5.14}$$

The mass-balance constraint requires that the number of ij interactions equal the number

of ji interactions as in Equation (4.24). This links the local compositions by the relation

$$\frac{z_0 q_i N_i}{2} x_{ij} = N_{ij} = N_{ji} = \frac{z_0 q_j N_j}{2} x_{ji}$$

Instead of relating the local compositions in this manner, engineers have related them

through the energetics in Equations (5.8), (5.11) and (5.13). The errors arising from the

lack of a mass-balance constraint are not explicit in the correlations, as they are contained

within the regressed parameters.

The UNIQUAC relation originally is derived using the concept of one-fluid

theory (Abrams and Prausnitz, 1975), where the statistics of the mixture system were

derived as the molecules existed in the same fluid. A second derivation of UNIQUAC

(Maurer and Prausnitz, 1978) offers a consistent derivation under the above stated

criticism by employing the two-fluid theory. The two-fluid theory supposes that the changes in the external energy, $\Delta W'$, evolved from a binary mixture process is given by

$$\Delta W' = \frac{z_0 N_1}{2}\left(\theta_{11}\varepsilon_{11}^{(1)} + \theta_{21}\varepsilon_{21}^{(1)} - \varepsilon_{11}^{(0)}\right) + \frac{z_0 N_2}{2}\left(\theta_{22}\varepsilon_{22}^{(2)} + \theta_{12}\varepsilon_{12}^{(2)} - \varepsilon_{22}^{(0)}\right) \qquad (5.15)$$

where $\varepsilon_{ij}^{(0)}$, $\varepsilon_{ij}^{(1)}$, $\varepsilon_{ij}^{(2)}$ denote an interaction energy between i and j in the mixture state, in fluid 1 and fluid 2, respectively. These latter fluids do not necessarily represent pure fluid states. It is further assumed that the homogeneous interactions are equivalent

$$\varepsilon_{11}^{(1)} = \varepsilon_{11}^{(0)}, \varepsilon_{22}^{(2)} = \varepsilon_{22}^{(0)}$$

while the inhomogeneous interaction is not

$$\varepsilon_{21}^{(1)} \neq \varepsilon_{12}^{(2)}$$

This assumption leads to the two different energetic parameters for a single ij interaction. This adds to the capability of the model, yet it renders the description theoretically incorrect and empirical (Kehiaian, et al. 1978; Klamt, et al. 2002).

The mass-balance constraint and a reinterpretation of the factor multiplying the residual contribution have been applied to the UNIQUAC expression (Knox, 1982). The resulting expression for the local surface area fraction from (5.11) is now given by

$$\frac{\theta_{ij}}{\theta_{ii}} = \frac{\theta_j}{\theta_i}e^{-\frac{\varepsilon_{ij}-\varepsilon_{ii}}{k_B T}} \qquad (5.16)$$

and the full expression for the Gibbs energy from (5.10) is now

$$G - \sum_i x_i G_i = RT\sum_i x_i \ln\phi_i + RT\sum_i x_i \frac{z_0 q_i}{2}\ln\frac{\theta_i}{\phi_i} + RT\sum_i x_i \frac{z_0 q_i}{2}\ln\frac{\theta_{ii}}{\theta_i} \qquad (5.17)$$

which simplifies to

$$G - \sum_i x_i G_i = RT\sum_i x_i \ln\phi_i + RT\sum_i x_i \frac{z_0 q_i}{2}\ln\frac{\theta_{ii}}{\phi_i} \qquad (5.18)$$

It has been shown that this expression fits VLE data about as well as the original UNIQUAC model (Knox, 1982).

5.4 The Quasi-Chemical Approach

Local composition models exist where the mass-balance constraint has been built into the derivation, thus alleviating the criticisms noted above. This approach corresponds to the quasi-chemical approach described in Section 4.3. Initial theoretical exercises on fluid systems are given for polysegmented molecules (Guggenheim, 1944b; Guggenheim, 1952). This approach has been applied for real systems where the interaction energies are correlated to reproduce data (Barker and Smith, 1954). A derivation of the approach has been offered to include a relation between random and nonrandom effects within the entropy term for any number of components in the system (Kehiaian et at, 1978). The nonrandom entropy is assumed to be a proportion of the random entropy, and an iterative solution to the nonlinear set of equations has been found (Knox, et al. 1984). The derivation of the same nonlinear set of equations is accomplished using surface segments within a Conductor-Like Screening Model (COSMO) (Lin and Sandler, 2002; Klamt, et al. 2002). Throughout this work, the partition function derivation and accompanying notation of Knox and coworkers is employed. A generalized derivation of this method is offered in Appendix B.

The expression for the excess molar Gibbs energy of a system describable through the quasi-chemical approach is similar to that given in Equation (5.18) combined with Equations (2.15) and (2.16)

$$\frac{G^E}{RT} = \sum_{i=1}^{n} x_i \ln\frac{\phi_i}{x_i} + \sum_{i=1}^{n} \frac{z_i x_i}{2} \ln\frac{x_{ii}}{\phi_i} \tag{5.19}$$

where $z_i = z_0 q_i$ is the number of contacts on molecule i. In this expression, the local compositions x_{ii} cannot be found explicitly. These values are a solution to the set of nonlinear equations given by

$$\theta_i x_{ij} = \theta_j x_{ji}, (i, j = 1, 2, ..., n) \tag{5.20}$$

$$\sum_{j=1}^{n} x_{ij} = 1, (i = 1, 2, ..., n) \tag{5.21}$$

$$\frac{x_{ij} x_{ji}}{x_{ii} x_{jj}} = \tau_{ij}^2, (i, j = 1, 2, ..., n) \tag{5.22}$$

where the area fraction of the molecule is found in a similar manner to Equation (5.6)

$$\theta_i = \frac{z_i x_i}{\sum_{j=1}^{n} z_j x_j} \tag{5.23}$$

and where the interactions are implicit within the τ_{ij} expression

$$\tau_{ij} = e^{-\frac{2\varepsilon_{ij} - \varepsilon_{ii} - \varepsilon_{jj}}{2 k_B T}} \tag{5.24}$$

The athermal contribution in the above model is given by Equation (5.5). This quasi-chemical model has $n(n-1)/2$ parameters for a mixture of n species, a reduction by a factor of 2 due to the recognition of the mass-balance constraint. The interaction parameters within this engineering method are the τ_{ij}. This model has been shown to perform comparably to the UNIQUAC expression while only needing half the number of interaction parameters, although difficulty is encountered when attempting to correlate highly non-ideal mixtures, such as alcohol/alkane mixtures (Knox, et al. 1984).

The excess molar Gibbs energy model of Equation (5.19) also exists as a group-contribution method (Knox, 1987). The excess Gibbs energy for that model is given by

$$\frac{G^E}{RT} = \sum_{i=1}^{n} x_i \ln\frac{\phi_i}{x_i} + \sum_{i=1}^{n} x_i \frac{z_i}{2}\ln\frac{\theta_i}{\phi_i} + \sum_{i=1}^{n}\sum_{k=1}^{m}\frac{Z_k v_{i,k} x_i}{2}\ln\left(\frac{y_{kk}/\Theta_k}{y_{kk}^{(i)}/\Theta_k^{(i)}}\right) \qquad (5.25)$$

where m is the number of different group types in the system, Z_k is the number of contacts on group k, $v_{i,k}$ denotes the number of group k within molecule i, Θ_k is the surface area fraction of group k. The molecular surface area fraction is given through group quantities by

$$\theta_i = \frac{\displaystyle\sum_{k=1}^{m} x_i v_{i,k} Z_k}{\displaystyle\sum_{j=1}^{n}\sum_{k=1}^{m} x_j v_{j,k} Z_k} \qquad (5.26)$$

The molecular volume fractions ϕ_i are found using volume properties of the groups by

$$\phi_i = \frac{\displaystyle\sum_{k=1}^{m} x_i v_{i,k} R_k}{\displaystyle\sum_{j=1}^{n}\sum_{k=1}^{m} x_j v_{j,k} R_k} \qquad (5.27)$$

where R_k is the number of lattice sites occupied by group k. The surface area fraction of group k is given by

$$\Theta_k = \frac{\displaystyle\sum_{i=1}^{n} x_i v_{i,k} Z_k}{\displaystyle\sum_{i=1}^{n}\sum_{l=1}^{m} x_i v_{i,l} Z_l} \qquad (5.28)$$

The local compositions y_{kk} in the excess Gibbs expression in Equation (5.25) also cannot be included explicitly. The system of equations that yield the local compositions are given by

$$\Theta_k y_{kl} = \Theta_l y_{lk}, \left(k,l = 1, 2, ..., m\right) \qquad (5.29)$$

$$\sum_{l=1}^{m} y_{kl} = 1, \left(k = 0, 1, ..., m\right)$$ (5.30)

$$\frac{y_{kl}y_{lk}}{y_{kk}y_{ll}} = \tau_{kl}^{2}, \left(k, l = 1, 2, ..., m\right)$$ (5.31)

The resulting expression for the activity coefficient of species i in the mixture is given by

$$\ln \gamma_{i} = \ln \gamma_{i}^{ath} + \sum_{k=1}^{m} \frac{Z_{k}\nu_{i,k}}{2} \ln \left(\frac{y_{kk}/\Theta_{k}}{y_{kk}^{(i)}/\Theta_{k}^{(i)}}\right)$$ (5.32)

where γ_{i}^{ath} is the activity coefficient of molecule i in an athermal solution and traditionally holds the contribution of the first two terms within Equation (5.25). Again, this system of equations yields $m(m-1)/2$ interaction parameters for the m functional groups within the system. This method is able to predict VLE data for systems without experimental data, similar to the use of UNIFAC, although the number of correlated interaction parameters is not large. Correlated interaction parameters between functional groups show physical significance by trending monotonically with the number of hydrogen atoms within a CH_{n} group (Knox, 1987).

5.5 Defining Functional Groups using Computational Chemistry

In order to minimize the number of experiments in the evaluation of interaction parameters within group-contribution methods, the complete transferability of groups has to date almost always been assumed. By definition, a group that is transferable has the same intermolecular and intramolecular properties in all applications. Thus the individual group properties are presumed to be independent of the other groups that exist

in the system, as well as independent of the other groups constituting the molecule. Within activity coefficient models, where interaction effects are important, group-group interaction energies have been considered transferable, just as with all other group properties.

The development of a group-contribution method for evaluation of system properties must first begin with a search for the proper group definitions. A group can be as small as a single atom, or it can be a combination of several atoms, or it can consist of an entire molecule. A successful additivity scheme that predicts thermophysical property data (Benson, 1976) defines a group as a heavy (non-hydrogen atom) atom of a linear or cyclic molecule and all the atoms whose only bond or bonds are with the heavy atom (mainly hydrogen atoms). Other heavy atoms bonded to the group heavy atom are included in the definition for nominal purposes.

The defining of groups has been arbitrary, usually resulting in definitions that suit particular systems and allow for the best predictive results. Where groups in certain molecules are clearly not transferable, they are replaced by larger, more specific groups that do not occur in as many molecules. These definitions then become somewhat particular to the molecules investigated, and the advantage of adopting a group-contribution technique is lost. However, the problem for smaller group definitions, that the group properties vary between molecules, remains (Wu and Sandler, 1991a). Thus it is impractical to rely upon a group-contribution method that relies upon universal properties for small functional groups.

The use of computational chemistry has been proposed as a tool for ending the arbitrary methods for defining groups within engineering models (Wu and Sandler,

1991a). It is suggested that, to expect the transferability of a group, the geometry of the constituent atoms should be equivalent amongst all species when occurring. Also, Mulliken population analysis (Mulliken, 1955) is used to examine charge distributions within molecules by assigning effective partial charges to atoms. Groups are defined as collections of atoms with a near neutral Mulliken charge. More recently, the neutrality criterion has been revised to one that accepts a consistent distribution of charge over the atoms as indicative of transferability (Lin and Sandler, 2000). Some groups satisfying the criteria are the same as those for UNIFAC, while new group definitions, usually a combination of smaller groups, are employed to satisfy the Mulliken charge criterion. A result of this study states that methylene (CH_2) groups should be included with highly electronegative groups (OH, NH_2, NO_2, CHO) to create an approximately transferable group.

5.6 An Overview of COSMO-based Methods

A successful attempt to evaluate VLE with computational methods and correlated model parameters has been proposed (Klamt, 1995; Klamt et al 1998). The COSMO for Real Solvent (COSMO-RS) approach avoids the problems of past group-contribution methods by avoiding the need to define functional groups altogether. A molecule within a condensed state is assumed to be solvated within a dielectric continuum. A surface-charge distribution, referred to as the σ-profile, is calculated for each molecule using computational chemistry. The surface-charge distribution is partitioned into segments, and molecular interactions are accounted for by the juxtaposition of these segments. The types of binary surface segment interactions are assigned in a similar manner of

minimizing the partition function with respect to N_{ij}, and properties such as vapor pressure, partition coefficients and energy of hydration become available.

The parameterization of the original COSMO-RS scheme involves two structural parameters for each of three elements (H , C , and O) and eight global model parameters. The structural parameters for the elements N and Cl were determined after fixing the parameters of the other elements. Therefore, for molecules containing the above elements, the model requires 18 global parameters, much less than the hundreds of parameters necessary in the UNIFAC methods.

The procedure of parameterizing and optimizing the model has led to some compromising of the basic model premises. Although claimed that the model parameters are close to their theoretical estimate, a pair of parameters, r_{av} and r_{eff}, which are equivalent in theory, differ after optimization by a factor of three (Klamt, et al. 1998). The sampling of properties around a molecule, such as for the screening charge density in the COSMO-RS model, has been criticized as being dependent on the distance from the nuclei and on the distribution of sampling points (Sigfridsson and Ryde, 1998). Theoretical problems also result from the choice of radii for the elements. It had been found that charge density outside the approximating sphere for an atom distorts the screening charge density. Also, the interaction distance described by the radii of the two interacting components may not correspond to the actual interaction distance in the real fluids, especially in instances of hydrogen bonding.

It is noted that the COSMO-RS method does not satisfy the Gibbs-Duhem relation, a basic requirement for expressions that calculate activity coefficients γ_i (Lin and Sandler, 2002). This relation states, for a binary mixture,

$$x_1 \frac{\partial \ln \gamma_1}{\partial x_1} + x_2 \frac{\partial \ln \gamma_2}{\partial x_1} = 0 \qquad (5.33)$$

To satisfy this relation, the activity coefficient for molecule i in a mixture is proposed to be

$$\ln \gamma_i = \ln \gamma_i^{ath} + \sum_{k=1}^{m} P_k^{(i)} \left(\sum_{j=1}^{m} v_{i,j} \right) \ln \left(\frac{\Gamma_k}{\Gamma_k^{(i)}} \right) \qquad (5.34)$$

where $v_{i,k}$ here denotes the number of segments of type k within molecule i, thereby giving the total number of segments contributed by molecule i by the inner summation, $P_k^{(i)}$ is the probability of finding a segment of type k within pure species i, and Γ_k is deemed the activity coefficient of surface segment k. The activity coefficient of segment k is found by solving the nonlinear system of equations, given by

$$\frac{1}{\Gamma_k} = \sum_{l=1}^{m} P_l \Gamma_l \tau_{kl}, \left(k = 1, 2, ..., m \right) \qquad (5.35)$$

This methodology, combining the concept of interacting surface segments with thermodynamically consistency of Equation (5.33) is called the COSMO Segment Activity Coefficient (COSMO-SAC) model (Lin and Sandler, 2002). Soon after, the COSMO Surface-Pair Activity Coefficient Equation (COSMOSPACE) method (Klamt, et al. 2002), which utilizes the same statistics, has been proposed.

5.7 The Statistics of Interacting Surface Segments

The interacting surface segment model is the contribution to the residual partition function within the COSMO-based models. However, the link between functional group interaction schemes and surface segment interactions (Kehiaian, et al. 1978) seems to be

made far earlier than the development of COSMO-SAC and COSMOSPACE. The more recent derivation of the surface segment statistics is presented here (Klamt, et al. 2002).

Consider a set of surface segments, where there are m different types of segments. Let M_i be the number of segments of type i. The goal here is to determine the activity coefficient of a general segment type i under the influence of the interactions of other segments while the macroscopic system is in the liquid phase. The molecular activity coefficient will then be available by summing the functional group contributions, as in other excess Gibbs energy models.

The chemical potential for a surface segment of type i is found through the canonical partition function of Equation (3.26). Since the number of segments of the system is so large (on the order of the number of molecules within the system), the derivative in the chemical potential relation is approximated by a finite difference

$$\mu_i = -k_B T \frac{d \ln Q}{dM_i} \simeq -k_B T \frac{\ln Q - \ln Q_{(-i)}}{M_i - (M_i - 1)} = -k_B T \ln \frac{Q}{Q_{(-i)}} \qquad (5.36)$$

where the subscript $(-i)$ denotes the system with one less segment of type i. Also removing a segment of type j follows the procedure in Equation (5.36)

$$\mu_i + \mu_j = -k_B T \ln \frac{Q}{Q_{(-i,-j)}} \qquad (5.37)$$

Now the goal is to clarify the ratio of partition functions given here.

Consider a similar system as above, except where the surface segments are distinguishable. The partition function for this new system of segments is very similar to that above, except for a familiar factor that includes indistinguishability in the above case. The partition function of this new case, Q', in terms of the earlier partition function is

$$Q' = Q\prod_{k=1}^{m} M_k! = QM_{ind} \tag{5.38}$$

Therefore, the chemical potential of segment i in terms of Q' is given by inserting Equation (5.38) into Equation (5.36)

$$\mu_i = -k_B T \ln \frac{Q'M_{ind(-i)}}{Q'_{(-i)}M_{ind}} = -k_B T \ln \frac{Q'}{Q'_{(-i)}} + k_B T \ln M_i = \mu'_\alpha + k_B T \ln M_i \tag{5.39}$$

where $M_{ind}/M_{ind(-i)} = M_i$ and the segment removed from the system of distinguishable objects is called α. The removal of α from the system of distinguishable objects is implicit within the chemical potential μ'_α, and this must be determined.

The partition function of the distinguishable system includes terms involving segment α and all the interacting segments. It is presumed that these effects are separable from the partition function in such a way as to relate the whole partition function with that where α is removed (Klamt, et al. 2002). This is given by

$$Q' = \sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^{M} 2 \frac{M}{2} e^{\frac{-\varepsilon'_{\alpha\beta}}{k_B T}} Q'_{(-\alpha,-\beta)} \tag{5.40}$$

where the factors of 2 and $M/2$ account for the indistinguishability of the $\alpha\beta$ ordering and the $M/2$ interaction sites where the interaction may take place, respectively, and where $\varepsilon'_{\alpha\beta}$ is the interaction energy between the segments. Inserting Equation (5.37) with primed notation into Equation (5.40) and simplifying yields

$$Q' = \sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^{M} M e^{\frac{-\varepsilon'_{\alpha\beta} + \mu'_\alpha + \mu'_\beta}{k_B T}} Q' \tag{5.41}$$

Isolating μ'_α gives

$$\mu'_\alpha = -k_B T \ln \left(\sum_{\substack{\beta=1 \\ \beta \neq \alpha}}^{M} e^{\frac{-\varepsilon'_{\alpha\beta} + \mu'_\beta}{k_B T}} \right) - k_B T \ln M \tag{5.42}$$

Here, the model notation must revert back to the original ensemble, where the surface segments of a given type are indistinguishable. This is accomplished by inserting the relationship between μ'_α and μ_i found in Equation (5.39) and allowing the summation to run over all types with the appropriate M_j multiplying the summand, since there are M_j terms in the summation within Equation (5.42) that contribute effects from type j. These manipulations result in the following

$$\mu_i = -k_B T \ln \left(\sum_{j=1}^{m} M_j e^{\frac{-\varepsilon_{ij} + \mu_j - \ln M_j}{k_B T}} \right) + k_B T \ln \frac{M_i}{M} \tag{5.43}$$

where $\varepsilon'_{\alpha\beta} = \varepsilon_{ij}$ when the segment named β is of type j. Simplifying this gives

$$\mu_i = -k_B T \ln \left(\sum_{j=1}^{m} e^{\frac{-\varepsilon_{ij} + \mu_j}{k_B T}} \right) + k_B T \ln x_i \tag{5.44}$$

The definition of the activity of surface segment i, given by Γ_i, closely resembles that for molecules, given by Equation (2.20), except that the reference system here is not an ideal solution but a solution fully occupied by segments i

$$k_B T \ln \Gamma_i = \mu_i - \mu_{ii} - k_B T \ln x_i \tag{5.45}$$

where μ_{ii} is the chemical potential of segment i in a system of pure segments i, and where the final term is included due to the composition difference in the real and reference states (a problem not encountered in Equation (2.20)). By Equation (5.44), it follows that

$$\mu_{ii} = \frac{\varepsilon_{ii}}{2}$$ (5.46)

To find the relation for the activity coefficient of the segments, Equations (5.44) through (5.46) are combined to yield the final result

$$\frac{1}{\Gamma_i} = \sum_{j=1}^{m} x_j \tau_{ij} \Gamma_j, (i = 1, 2, \ldots m)$$ (5.47)

The segment activity coefficients within Equations (5.47) are used to calculate the activity of the molecule in a way similar to the engineering excess Gibbs energy formulations in Section 5.4. The activity for the molecule is given by

$$\ln \gamma_i = \ln \gamma_i^{ath} + \sum_{j=1}^{n} \nu_{i,j} \ln \frac{\Gamma_j}{\Gamma_j^{(i)}}$$ (5.48)

where $\nu_{i,j}$ is the number of segments of type j in molecule i, and the superscript (i) denotes the quantity as it exists in pure species i. This relation is quite similar to that given by Equation (5.32) where the quasi-chemical equations and the local composition model are considered.

5.8 Surface Segment Statistics and the Quasi-Chemical Equations

The relationship between the result of the surface segment statistics, given by Equation (5.48) and that of the quasi-chemical equations, given by Equation (5.32) suggests a strong resemblance in concepts. If one were to consider that each contact on the external surface area of a functional group is equivalent to a surface segment, and that all the contacts on a particular functional group are the same surface segment type, then it follows that

$$\nu_{i,j}^{(\text{group})} Z_j = \nu_{i,j}^{(\text{segment})}$$ (5.49)

It now seems plausible that if one were to relate the local composition fractions and areas within the functional group argument to the activity coefficient of a surface segment by the relation

$$\sqrt{\frac{y_{jj}}{\Theta_j}} = \Gamma_j \qquad (5.50)$$

then Equations (5.48) and (5.32) are exactly equivalent. The activity coefficient for a species through the quasi-chemical approach is given by

$$\ln \gamma_i = \ln \gamma_i^{ath} + \sum_j \nu_{i,j} Z_j \ln \frac{\Gamma_j}{\Gamma_j^{(i)}} \qquad (5.51)$$

This conclusion is also achieved when the nonlinear equations within the quasi-chemical model, Equations (5.29) through (5.31), are simplified. What one finds is that these m^2 equations are reducible to the m equations necessary to find the surface segment activity coefficients in Equation (5.47), as long as the relationship within Equation (5.50) is used. This simplification results in the following nonlinear system

$$\frac{1}{\Gamma_i} = \sum_j \theta_j \tau_{ij} \Gamma_j, \left(i = 0,1,...,n\right) \qquad (5.52)$$

This is accomplished for a system with a limited number of functional groups in Appendix B. A different derivation that relates the quasi-chemical approach to the statistics within COSMO-based models has also been offered (Panayiotou, 2003a).

5.9 Summary

Engineers have taken large steps in the past 30 years in formulating a workable description of the liquid state. These efforts have yielded strong fundamental work in the nature of entropy in a fluid system and how the interactions between molecules bring a level of order to the system arrangements.

The methods based on local compositions, namely the UNIQUAC and UNIFAC methods, have given engineers a very powerful tool in the prediction of mixture system behavior. These tools are still referenced today, even though the criticisms of the method have rendered the theoretical framework incorrect. Enough flexibility has been built into the model to allow engineers to predict the thermodynamics of a large number of fluid systems.

The faults of these methods, however, render them unusable if a rational approach is taken to build more predictive capabilities into the methods. The physical reasoning behind the number of contacts and the number of lattice sites of a group, as well as the energetics within the model, is compromised by the lack of mathematical consistency in the derivation of the methods. There is no guarantee that these quantities, if calculated at a more fundamental level, would improve the predictive capabilities of the method since their physical significance is lost.

The quasi-chemical approach offers the consistency that the early local-composition models are lacking. Since no compromises had been made during the derivation of the expressions, it is at least possible to rationalize the improvement of these models by the use of physically-significant structural and electrostatic properties of the functional groups. The success of the COSMO-based methods shows the power of

the quasi-chemical method in the application of information from computational chemistry.

Quantum and computational chemistry is explored further to find first-principle quantities similar to the surface segment information in COMSO-based methods. However, the information of interest in this work is related to the information commonly used within past group-contribution methods: structural and electrostatic properties of molecules that lead to the interaction energies of functional groups.

CHAPTER 6

QUANTUM CHEMISTRY

Application of quantum chemistry allows for the determination of atomic and molecular properties from first principles. Insightful information about molecules is available, even when starting from the basic concepts of subatomic particles and a very simplistic interaction law, Coulomb's Law. Such information is more accessible now through the explosive growth of computational power and the refinement of computational algorithms. The theory of quantum chemistry also leads to insights in statistical mechanics, such as the quantized states of system and the manifestations of interacting electron clouds.

In this chapter, the most basic problems of quantum chemistry, the hydrogen atom and the molecular Hamiltonian, are outlined. Of specific interest are the functions used to represent the electron density within an atomic system, since these will be employed in an engineering model that approximates electron densities for functional groups. A classical solution method for the molecular Hamiltonian, called the Hartree-Fock Self Consistent Field method, is presented in its pedagogical form. Also, mathematical methods on how to extract information from the solution of the Hamiltonian are reviewed. The more complete presentation of these derivations and solution methods are documented in standard textbooks (Levine, 2000; Szabo and Ostlund, 1982). A more contemporary interpretation of molecular properties through Atoms in Molecules theory (Bader, 1990) is also reviewed.

6.1 Hamiltonian for the Hydrogen Atom and Its Solution

The hydrogen atom is one of the simplest problems within quantum chemistry, and one of the few real systems that is completely soluble. The problem is to determine the energy states of a single electron existing around a single proton. The only force between the two is the force of attraction arising from the Coulombic expression, Equation (3.42).

In quantum chemistry, the solution of problems begins with the formulation of the Hamiltonian of the system. The Hamiltonian is the sum of the kinetic and potential energies, K and V, thereby accounting for all the energy within a system

$$\hat{H} = \hat{K} + \hat{V} \tag{6.1}$$

The notation includes the $^\wedge$ symbol to signify the use of operators that act upon the wavefunction of the system Ψ. The Hamiltonian is important because the energy of the system is revealed by this eigenvalue equation

$$\hat{H}\Psi = E\Psi \tag{6.2}$$

When expanded, this relation is either expressible as a partial differential equation where $\hat{H}$ contains derivatives acting on the function Ψ, or a linear algebra equation where $\hat{H}$ is a matrix acting on the vector Ψ. In this work, Equation (6.2) is considered a partial differential equation. The goals for all problems expressed in the form of Equation (6.2) are to determine the Hamiltonian of the system using classical mechanical arguments and to solve for the wavefunction analytically or numerically.

The wavefunction Ψ is the most important description of the system and its behavior in space and time. It holds all the information about all the measurable properties of that system over all time. However, in this work the wavefunctions of equilibrium systems are independent of time and describe stationary states. The

wavefunction within the framework of quantum chemistry require it to be a well-behaved, mathematical function. Firstly, the function must be continuous and differentiable over all the variables within the system. Secondly, the function must be quadratically integrable, where $\int \Psi^* \Psi d\tau$ integrated over the entire variable ranges must be finite. The integrand here is the complex conjugate of the wavefunction Ψ^* multiplying the wavefunction itself Ψ. The product is integrated over the space within the bounds of the variables, a differential element of which is given by $d\tau$. When integrated over all space, the wavefunction of one particle is not only finite but also normalized, satisfying the relationship

$$\int \Psi^* \Psi d\tau = 1 \tag{6.3}$$

A normalization constant is usually included in the wavefunction to ensure this.

The Hamiltonian and wavefunction for the hydrogen atom is a classical problem offering insight to the solution of the molecular Hamiltonian. The general Hamiltonian keeps track of the kinetic energies of both the proton and the electron, and the potential energy of the system due to their Coulombic interaction. In a reduced problem, the only coordinates of interest are those that express the position of the electron relative to the proton. This Hamiltonian thereby only considers the kinetic energy of the electron around a motionless proton and the potential energy from the Coulombic interaction

$$-\frac{\hbar^2}{2\mu}\left(\nabla^2 \psi\right) - \frac{e^2}{r}\psi = E\psi \tag{6.4}$$

where ψ is the wavefunction for the hydrogen atom system, $\hbar = h/2\pi$ is the definition of h-bar, $\mu = m_p m_e / \left(m_p + m_e\right)$ is the reduced mass of the proton-electron system, e is the

charge of one electron, $r = \sqrt{x^2 + y^2 + z^2}$ is the distance between the electron and the proton, and ∇^2 is the Laplacian operator, given in Cartesian coordinates by

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{6.5}$$

The proton is assumed to exist at the origin, and the Cartesian coordinates describe the position of the electron about the origin.

Two more simplifications are made before solving Equation (6.4). Firstly, atomic units are defined to simplify the notation. Multiplying through by $\mu/\hbar^2$ gives

$$-\frac{1}{2}\left(\nabla^2 \psi\right) - \frac{e^2 \mu}{\hbar^2}\frac{1}{r}\psi = \frac{\mu}{\hbar^2}E\psi \tag{6.6}$$

Define the reduced energy $\tilde{E}$ as

$$\tilde{E} = \frac{\mu}{\hbar^2}E \tag{6.7}$$

Also define the reduced interaction distance $\tilde{r}$ as

$$\frac{1}{\tilde{r}} = \frac{e^2 \mu}{\hbar^2}\frac{1}{r} \tag{6.8}$$

to give the partial differential equation in simplified form

$$-\frac{1}{2}\left(\nabla^2 \psi\right) - \frac{1}{\tilde{r}}\psi = \tilde{E}\psi \tag{6.9}$$

The reduced energy and distance are now in atomic units. These units have been deemed hartrees and bohrs, respectively, named after significant contributors to the field. The conversion factors for hartrees are

$$1 \text{ hartree} = 627.15 \frac{\text{kcal}}{\text{mol}} = 2625.5 \frac{\text{kJ}}{\text{mol}} = 27.2114 \text{ eV} \tag{6.10}$$

The main conversion factor for the bohr is

$$1 \text{ bohr} = 0.529177 \text{ Å} \qquad (6.11)$$

Secondly, the symmetry of the problem suggests an approach on the PDE in spherical coordinates. The Laplacian in Equation (6.5) in spherical coordinates is given by

$$\nabla^2 = \frac{\partial^2}{\partial \tilde{r}^2} + \frac{2}{\tilde{r}} \frac{\partial}{\partial \tilde{r}} + \frac{1}{\tilde{r}^2} \frac{\partial^2}{\partial \theta^2} + \frac{1}{\tilde{r}^2} (\cot \theta) \frac{\partial}{\partial \theta} + \frac{1}{\tilde{r}^2 \sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \qquad (6.12)$$

This simplification allows for the wavefunction to be solved in a more natural coordinate system for this problem.

The problem when posed in spherical coordinates allows for the technique of separation of variables to be used. Using Equation (6.12) with Equation (6.9) and separating variables, the solution to the hydrogen atom is given as a product of radial and angular portions

$$\psi_{nlm} = R_{nl} (\tilde{r}) Y_l^m (\theta, \varphi) \qquad (6.13)$$

where Y_l^m is the spherical harmonic function, and the integers n, l, and m are quantum numbers that denote the quantum state of the electron. Ranges for these numbers are determined in the solution, and these are given by

$$\begin{aligned} n &= 1, 2, 3, \ldots \\ l &= 0, 1, \ldots, n-1 \\ m &= -l, -l+1, \ldots, l-1, l \end{aligned} \qquad (6.14)$$

The radial portion of the wavefunction, R_{nl}, is found to be

$$R_{nl} (\tilde{r}) = e^{-\tilde{r}/n} \tilde{r}^l \sum_{j=0}^{n-l-1} b_j \tilde{r}^j \qquad (6.15)$$

where the series coefficients within the summation are found using a recursion relation

$$b_{j+1} = \frac{2}{n} \frac{1+l+j+n}{(j+1)(j+2l+2)} b_j \qquad (6.16)$$

Each combination of quantum numbers within Equation (6.14) yields an energy state. The reduced energy for a set of quantum numbers is given by

$$\tilde{E} = \frac{-1}{2n^2} \tag{6.17}$$

The degeneracy of the energy level is reflected in the number of combinations of l and m possible for a given n.

Each combination of quantum numbers in Equation (6.14) also yields a function called an atomic orbital, a function in space that reflects the energy state of the electron. The simplest of these orbitals is the ground state of the hydrogen atom, the $1s$ orbital

$$\psi_{100} = \psi_{1s} = \frac{1}{\sqrt{\pi}} e^{-\tilde{r}} \tag{6.18}$$

This function does not have angular dependencies, since the spherical harmonic for $l = m = 0$ is constant with respect to the angles. The only series coefficient within the radial contribution is b_0, and this is used solely to normalize the wavefunction as necessary by Equation (6.3).

Here are several examples of atomic orbitals with $m = 0$. The following wavefunctions include the depiction through the quantum numbers and through the commonly used alphanumeric names of the orbitals. The $2s$ orbital is

$$\psi_{200} = \psi_{2s} = \frac{1}{4\sqrt{2\pi}} \left(2 - \tilde{r}\right) e^{-\tilde{r}/2} \tag{6.19}$$

The $3s$ orbital is similar, except for the polynomial radial portion

$$\psi_{300} = \psi_{3s} = \frac{1}{81\sqrt{3\pi}} \left(27 - 18\tilde{r} + 2\tilde{r}^2\right) e^{-\tilde{r}/3} \tag{6.20}$$

The $2p_z$ orbital is an example of a state that is dependent on an angle

$$\psi_{210} = \psi_{2p_z} = \frac{1}{4\sqrt{2\pi}} \tilde{r} e^{-\tilde{r}/2} \cos\theta \tag{6.21}$$

For the cases when $m \neq 0$, the spherical harmonics contain imaginary numbers. Real-valued wavefunctions are attainable from these expressions by determining the correct linear combination that eliminates the imaginary parts. Two real orbitals, $2p_x$ and $2p_y$, are found in this manner

$$\frac{1}{\sqrt{2}} \left(\psi_{21(-1)} + \psi_{211} \right) = \psi_{2p_x} = \frac{1}{4\sqrt{2\pi}} \tilde{r} e^{-\tilde{r}/2} \sin\theta \cos\phi$$

$$\frac{1}{i\sqrt{2}} \left(\psi_{21(-1)} - \psi_{211} \right) = \psi_{2p_y} = \frac{1}{4\sqrt{2\pi}} \tilde{r} e^{-\tilde{r}/2} \sin\theta \sin\phi \tag{6.22}$$

Higher orbitals, such those on the $n = 3$ energy level, are also found this way

$$\psi_{3p_x} = \sqrt{\frac{2}{\pi}} \frac{1}{81} (6 - \tilde{r}) \tilde{r} e^{-\tilde{r}/3} \sin\theta \cos\phi$$

$$\psi_{3d_{yz}} = \sqrt{\frac{2}{\pi}} \frac{1}{81} \tilde{r}^2 e^{-\tilde{r}/3} \sin\theta \cos\theta \sin\phi \tag{6.23}$$

Orbitals are also expressible in Cartesian coordinates, which are helpful within computational schemes. The s-type orbitals are expressed in Cartesian form by replacing $\tilde{r}$ with the definition of distance

$$\tilde{r} = \sqrt{\tilde{x}^2 + \tilde{y}^2 + \tilde{z}^2} \tag{6.24}$$

The directionality of the orbitals for those where $l \neq 0$ becomes more obvious when portions of the spherical expressions are expressed in Cartesian coordinates

$$\tilde{x} = \tilde{r} \sin\theta \cos\varphi$$
$$\tilde{y} = \tilde{r} \sin\theta \sin\varphi \tag{6.25}$$
$$\tilde{z} = \tilde{r} \cos\theta$$

The motivation behind the common names of the above orbitals becomes apparent when Cartesian factors are included:

$$\psi_{2P_z} = \frac{1}{4\sqrt{2\pi}} \tilde{z} e^{-\tilde{r}/2} \tag{6.26}$$

$$\psi_{2P_x} = \frac{1}{4\sqrt{2\pi}} \tilde{x} e^{-\tilde{r}/2} \tag{6.27}$$

$$\psi_{2P_y} = \frac{1}{4\sqrt{2\pi}} \tilde{y} e^{-\tilde{r}/2} \tag{6.28}$$

$$\psi_{3P_x} = \sqrt{\frac{2}{\pi}} \frac{1}{81} (6 - \tilde{r}) \tilde{x} e^{-\tilde{r}/3} \tag{6.29}$$

$$\psi_{3d_{yz}} = \sqrt{\frac{2}{\pi}} \frac{1}{81} \tilde{y}\tilde{z} e^{-\tilde{r}/3} \tag{6.30}$$

6.2 Molecular Hamiltonian and Its Solution

The molecular system is of interest in this work, since nearly all of the compounds encountered in practice are not constructed from a single atom, especially not a single hydrogen atom. The molecular Hamiltonian is the generalization of the atomic Hamiltonian expressed in Equation (6.9)

$$\hat{H} = -\frac{m_e}{2}\sum_\alpha \frac{1}{m_\alpha}\nabla_\alpha^2 - \frac{1}{2}\sum_i \nabla_i^2 + \sum_\alpha \sum_{\beta>\alpha} \frac{Z_\alpha Z_\beta}{\tilde{r}_{\alpha\beta}} - \sum_\alpha \sum_i \frac{Z_\alpha}{\tilde{r}_{\alpha i}} + \sum_i \sum_{j>i} \frac{1}{\tilde{r}_{ij}} \tag{6.31}$$

where α and β range over the nuclei, ∇_α^2 is the Laplacian operator with respect to coordinates of nucleus α, Z_α represents the atomic number of nucleus α, $\tilde{r}_{kl}$ is the reduced interaction distance between entities k and l, and i and j range over the electrons. From left to right, this equation holds the mathematical expressions for the kinetic energy of the nuclei, the kinetic energy of the electrons, and the potential energies from nucleus-nucleus interactions, from nucleus-electron interactions and from electron-electron interactions. The nuclei and electrons are here treated as point masses.

The solution of the problem expressed in the molecular Hamiltonian is extremely complicated, since the wavefunction of the system involves the location of a large number of electrons that are under the influence of a complex field dependent on the location of the nuclei. A classical assumption, called the Born-Oppenheimer approximation, is made to simplify Equation (6.31) by decoupling the problem of the nuclei positions from the problem of the electron positions. The Born-Oppenheimer approximation comes from the concept that the mass of an electron is several orders of magnitude less than the mass of any nucleus, even a single proton. Therefore, since $m_e \ll m_\alpha$ for all α, the magnitude of first term of Equation (6.31), which contributes the kinetic energy of the movement of the nuclei, vanishes. This leaves

$$\hat{H} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_\alpha \sum_{\beta > \alpha} \frac{Z_\alpha Z_\beta}{\tilde{r}_{\alpha\beta}} - \sum_\alpha \sum_i \frac{Z_\alpha}{\tilde{r}_{\alpha i}} + \sum_i \sum_{j > i} \frac{1}{\tilde{r}_{ij}} \tag{6.32}$$

This Hamiltonian can now be separated into two major contributions

$$\hat{H} = \hat{H}_{el} + \sum_\alpha \sum_{\beta > \alpha} \frac{Z_\alpha Z_\beta}{\tilde{r}_{\alpha\beta}} \tag{6.33}$$

where the nucleus-nucleus interactions are found in the latter term, and the positions of the electrons relative to one another and the nuclei contribute to the electronic Hamiltonian

$$\hat{H}_{el} = -\frac{1}{2}\sum_i \nabla_i^2 - \sum_\alpha \sum_i \frac{Z_\alpha}{\tilde{r}_{\alpha i}} + \sum_i \sum_{j > i} \frac{1}{\tilde{r}_{ij}} \tag{6.34}$$

Equation (6.9) is recovered from Equation (6.34) if one nucleus with atomic number $Z_1 = 1$ and one electron exists in the system.

The separation in Equation (6.33) allows for the solution of the wavefunction solely describing electron position, while the locations of atomic nuclei are considered

parameters in the problem. This approach is intuitively correct because of the differences in velocities of electrons and nuclei in regular molecular motion. The speed with which electrons move is so much faster than nuclei that for normal nuclear movements due to vibration or rotation, electrons nearly instantaneously react and alter their state at all points in time to accommodate the nuclei locations.

Of specific importance is the location of the ground-state positions of the nuclei. This equilibrium geometry is the configuration that allows for the lowest energy configuration of all the possible geometric configurations of the nuclei and the electrons. This geometry reveals the energy minima around which molecular vibrational modes act and from which rotational modes depart. This configuration serves as the most probable configuration for the isolated molecule within the gas phase.

Solving the wavefunction for the Hamiltonian given in Equation (6.34) gives the electronic energy of the molecular system and the electron density profile around the nuclei. This information yields molecular properties of use within the statistical frameworks described above.

6.3 Calculation of Properties within Quantum Chemistry

Determining the wavefunction is the most powerful objective when one desires the properties of a system. The wavefunction holds all the measurable information of the system. Quantum chemistry includes methods to extract this information.

Every operator within quantum chemistry is directly related to a property within classical physics. For instance, the Hamiltonian operator $\hat{H}$ is related to the total system energy. When the operator acts upon the wavefunction describing the stationary state of

a system, as in Equation (6.2), the energy of the system is found. Another way of extracting this information is by taking the average value of that property within the system. For a general operator $\hat{F}$, the mean value of the corresponding classical property is given by

$$\bar{F} = \int \Psi^* \hat{F} \Psi d\tau \qquad (6.35)$$

The integration is taken over all the values of the coordinates. With regards to the system energy, the mean value described by the stationary state Ψ is given by

$$\bar{E} = \int \Psi^* \hat{H} \Psi d\tau = \int \Psi^* E \Psi d\tau = E \int \Psi^* \Psi d\tau = E \qquad (6.36)$$

Replacing $\hat{H}\Psi$ with $E\Psi$ is a consequence of Equation (6.2). Another example is the position operator. Assume one wants to know the average x-coordinate of the particle of the system. This is given by

$$\bar{x} = \int \Psi^* \hat{x} \Psi d\tau = \int \Psi^* x \Psi d\tau \qquad (6.37)$$

No further simplification is possible in this example, since the wavefunction is not necessarily an eigenfunction of the x-position operator, $\hat{x}$, and since the functionality of Ψ is not explicit here. Any operator can be included in Equation (6.35) to determine an average property value for the system.

A different notation is commonly used to express the integration of an operator in the form of Equation (6.35). The bracket notation is given by

$$\bar{F} = \langle \Psi | \hat{F} | \Psi \rangle \qquad (6.38)$$

If Ψ is an eigenfunction of the operator $\hat{F}$, the eigenvalue is factored out of the integral, and the notation is reduced

$$\bar{F} = \langle \Psi | \hat{F} | \Psi \rangle = \langle \Psi | F | \Psi \rangle = F \langle \Psi | \Psi \rangle = F \qquad (6.39)$$

The last simplification is because the normalization property of Equation (6.3), which yields $\langle \Psi | \Psi \rangle = 1$. In bracket notation, the calculation described in Equation (6.36) is depicted by

$$\bar{E} = \langle \Psi | \hat{H} | \Psi \rangle = \langle \Psi | E | \Psi \rangle = E \langle \Psi | \Psi \rangle = E \qquad (6.40)$$

For parameterized functions that approximate real wavefunctions, the variation theorem is used to determine the how good the approximation is. For example, if the task is to find the ground-state energy eigenvalue for a given Hamiltonian, the following inequality holds

$$\langle \Psi' | \hat{H} | \Psi' \rangle \geq E_0 \qquad (6.41)$$

where Ψ' represents an approximation to the real wavefunction, and E_0 is the lowest possible energy eigenvalue. The theorem also states that the wavefunction form and parameters that come closest to the ground-state energy, of course from above, yields the best approximation to the actual wavefunction. The parameters within the wavefunction are thereby found by minimizing the variation integral, the left side of Equation (6.41), with respect to the parameters.

6.4 The Hartree-Fock Self-Consistent Field Method

The problem posed in the molecular Hamiltonian is considerably simpler to solve if the Coulombic interactions between electrons are omitted. Equation (6.34) is reduced to

$$\hat{H} = -\frac{1}{2} \sum_i \nabla_i^2 - \sum_\alpha \sum_i \frac{Z_\alpha}{\tilde{r}_{\alpha i}} \qquad (6.42)$$

and the solution is simply the product of hydrogen-like orbitals given by Equation (6.13)

$$\Psi = \prod_i \psi_{n,l,m_i} \left(\tilde{r}_i, \theta_i, \phi_i \right) = \prod_i R_{n,l_i} \left(\tilde{r}_i \right) Y_{l_i}^{m_i} \left(\theta_i, \phi_i \right) \qquad (6.43)$$

Of course the Coulombic interaction between the electrons is a very strong influence on the molecular system and, thereby, the molecular Hamiltonian. The solution given by Equation (6.43) still serves as a workable first-order approximation to the real wavefunction.

To gain a more precise solution, changes to this approximation must be applied. The form of an improved solution to the molecular Hamiltonian retains the product of a radial portion multiplying the spherical harmonic

$$\Psi = \prod_i s_i \left(\tilde{r}_i, \theta_i, \phi_i \right) \qquad (6.44)$$

where the functional form of s_i is not necessarily the same as the explicit function given by Equations (6.13) and (6.15). The function does retain the exponential function in distance $e^{-r/n}$ and parameters that are able to change to accommodate physical effects, such as the Coulombic effects between electrons and nuclear charge screening. To find the best values for the parameters within Equation (6.44), the variation theorem described above is used.

The Hartree-Fock Self-Consistent Field (HF) method is an iterative algorithm that guesses the wavefunctions to Equation (6.34) by attempting to adjust the wavefunction description of one electron while the other electron wavefunctions are not changing. Consider Equation (6.44) the guess for the system wavefunction. The wavefunction of electron 1, given as $s_1 \left(\tilde{r}_1, \theta_1, \phi_1 \right)$ is subjected to a potential field V_1 created by remaining $n_e - 1$ electrons and the nuclei

$$V_1\left(\tilde{r}_1,\theta_1,\phi_1\right) = \sum_{i=2}^{n_e} u_{qq}\left(\tilde{\mathbf{r}}_1,\tilde{\mathbf{r}}_i\right) + \sum_{\alpha} \frac{Z_\alpha}{\tilde{r}_{\alpha 1}} = \sum_{i=2}^{n_e} \frac{q_1 q_i}{\tilde{r}_{1i}} + \sum_{\alpha} \frac{Z_\alpha}{\tilde{r}_{\alpha 1}} \tag{6.45}$$

Here, let the charges of all electrons except for that of electron 1 be represented with their wavefunctions. This is accomplished by representing the electrons as clouds of electron density ρ, the square of their wavefunctions. An integral over all space of the electron cloud will represent the total charge q_i of the electron in (6.45), thus leading to the relation

$$\sum_{i=2}^{n_e} \frac{q_1 q_i}{\tilde{r}_{1i}} = q_1 \sum_{i=2}^{n_e} \int \frac{\rho_i\left(\tilde{r}_i,\theta_i,\phi_i\right)}{\tilde{r}_{1i}} dv_i = \sum_{i=2}^{n_e} \int \frac{\left|s_i\left(\tilde{r}_i,\theta_i,\phi_i\right)\right|^2}{\tilde{r}_{1i}} dv_i \tag{6.46}$$

where dv_i is the differential volume element of electron i, and the charge on electron 1 is unity. The potential field becomes

$$V_1\left(\tilde{r}_1,\theta_1,\phi_1\right) = \sum_{i=2}^{n_e} \int \frac{\left|s_i\left(\tilde{r}_i,\theta_i,\phi_i\right)\right|^2}{\tilde{r}_{1i}} dv_i + \sum_{\alpha} \frac{Z_\alpha}{\tilde{r}_{\alpha 1}} \tag{6.47}$$

The result of this transformation is that the Hamiltonian representing the state of electron 1 is represented as a PDE similar to Equation (6.9), except for the dependence on the potential field

$$-\frac{1}{2}\nabla_1^2 s_1^{\{i\}} - V_1\left(\tilde{r}_1,\theta_1,\phi_1\right) s_1^{\{i\}} = \varepsilon_1 s_1^{\{i\}} \tag{6.48}$$

where the wavefunction representing electron 1 has a superscript $^{\{i\}}$ to denote the ith iterated solution to the wavefunction, and where ε_1 denotes the energy eigenvalue of electron 1 in the PDE. It is advantageous to have the potential in Equation (6.48) be a spherically symmetric potential field. This allows for the wavefunction $s_1^{\{i\}}\left(\tilde{r}_1,\theta_1,\phi_1\right)$ to

be rigorously separable into a radial portion and a spherical harmonic, as in Equation (6.13). To achieve this, V_1 is averaged over all angles

$$V_1(\tilde{r}_1) = \frac{\displaystyle\int_0^{2\pi}\int_0^{\pi} V_1(\tilde{r}_1,\theta_1,\phi_1)\sin\theta_1 d\theta_1 d\phi_1}{\displaystyle\int_0^{2\pi}\int_0^{\pi}\sin\theta_1 d\theta_1 d\phi_1} \tag{6.49}$$

This assumption is called the central field approximation and reduces the PDE to

$$-\frac{1}{2}\nabla_1^2 s_1^{\{1\}} - V_1(\tilde{r}_1)s_1^{\{1\}} = \varepsilon_1 s_1^{\{1\}} \tag{6.50}$$

Now, the iterated solution can be separated into a radial portion and a spherical harmonic.

To continue toward the solution, the wavefunction of the second electron is iterated under influence of the field of the remaining $n_e - 1$ electrons, including the new wavefunction for electron 1. The field is described by

$$V_2(\tilde{r}_2,\theta_2,\phi_2) = \int \frac{\left|s_1^{\{1\}}(\tilde{r}_1,\theta_1,\phi_1)\right|^2}{\tilde{r}_{12}}dv_1 + \sum_{i=3}^{n_e}\int \frac{\left|s_i(\tilde{r}_i,\theta_i,\phi_i)\right|^2}{\tilde{r}_{2i}}dv_i + \sum_{\alpha}\frac{Z_\alpha}{\tilde{r}_{\alpha 2}} \tag{6.51}$$

The central field approximation is used here to get a PDE of the form of Equation (6.50). Once the iterated solution of electron 2, $s_2^{\{1\}}(\tilde{r}_2,\theta_2,\phi_2)$, is found, electron 3 is considered, and so on and so forth.

This process continues until a first iterated solution for each electron is found. Once this occurs, the process begins again at electron 1 to achieve a second iterated solution, $s_1^{\{2\}}(\tilde{r}_1,\theta_1,\phi_1)$. All wavefunctions of the electrons are considered again. Additional iterations are performed until convergence of each wavefunction to a solution.

The solution method to find the wavefunction described in Equation (6.44) offers a pedagogical example for the HF method. Although a reasonable choice, the product of

purely radial functions with spherical harmonics, deemed spatial orbitals, does not account for spin effects of electrons. The proper wavefunction guess is actually an asymmetric product of spin orbitals, spatial orbitals multiplied by the spin wavefunction (either up-spin or down-spin) to achieve the proper description of the orbitals within which electrons occupy.

The energy of the system, the original eigenvalue of Equation (6.42), in the HF method is found to be

$$E_{HF} = 2\sum_{i=1}^{n_e/2}\left\langle s_i \left| \hat{h}_i \right| s_i \right\rangle + \sum_{i=1}^{n_e/2}\sum_{j=1}^{n_e/2}\left(2J_{ij} - K_{ij}\right) + \sum_{\alpha}\sum_{\beta>\alpha}\frac{Z_\alpha Z_\beta}{\tilde{r}_{\alpha\beta}} \tag{6.52}$$

where $\hat{h}_i$ is called the core Hamiltonian for electron i and is given by

$$\hat{h}_i = -\frac{1}{2}\nabla_i^2 - \sum_{\alpha}\frac{Z_\alpha}{r_{\alpha i}} \tag{6.53}$$

J_{ij} is called the Coulomb integral between electrons i and j, given by

$$J_{ij} = \left\langle s_i(1)s_j(2)\left| \tilde{r}_{12}^{-1} \right| s_i(1)s_j(2) \right\rangle = \iint \frac{\left|s_i(1)\right|^2 \left|s_j(2)\right|^2}{\tilde{r}_{12}}dv_1 dv_2 \tag{6.54}$$

and K_{ij} is called the exchange integral, given by

$$K_{ij} = \left\langle s_i(1)s_j(2)\left| \tilde{r}_{12}^{-1} \right| s_j(1)s_i(2) \right\rangle = \iint \frac{s_i(1)^* s_j(2)^* s_j(1)s_i(2)}{\tilde{r}_{12}}dv_1 dv_2 \tag{6.55}$$

The Coulomb integral is called such because it is the quantum mechanical equivalent to the Coulombic interaction between two electrons, except in quantum mechanics the electrons are treated as probability densities and not particles. The exchange integral is called such because it reflects the existence of the electrons partly in their own orbitals and party in the other electron's orbital, thus seemingly exchanging positions. The

summations containing the Coulomb and exchange integrals are taken over the different spatial orbitals the electrons occupy. The contribution of the Coulomb integral is doubled to reflect the occupancy of two electrons in each spatial orbital, one with up-spin and one with down-spin.

The integrals given in Equations (6.54) and (6.55) are encountered frequently within computational chemistry. These are called two-electron integrals, and multitudes of these are calculated within the computer routines that employ the HF and other solution methods. This work mainly deals with spatial orbitals. Therefore, the above integrals acting on spatial orbitals are typically denoted by

$$\left(ij|kl\right) = \left\langle s_i(1)s_j(2)\left|\tilde{r}_{12}^{-1}\right|s_k(1)s_l(2)\right\rangle = \int\int \frac{s_i(1)^* s_j(2)^* s_k(1)s_l(2)}{\tilde{r}_{12}} dv_1 dv_2 \quad (6.56)$$

This notation is sometimes called the 1212 (one-two-one-two) notation, since the orbital labels alternate between electrons, beginning with electron 1. In 1212 notation, the Coulomb integral and the exchange integral are expressed as

$$J_{ij} = \left(ij|ij\right) \quad (6.57)$$

$$K_{ij} = \left(ij|ji\right) \quad (6.58)$$

Another two-electron integral encountered in computations and this work is called quantum-hybrid integrals. An examples of one such integral is

$$L_{ij} = \left(ii|ij\right) \quad (6.59)$$

Since the orbitals are real functions describing spatial orbitals, the complex conjugate of the orbital function is equal to the orbital function. Different permutations within the integrals, and therefore within the parenthesis correspond to equivalent integrals. The

equivalence relations for the Coulombic, exchange, and hybrid integrals, respectively, are given by

$$J_{ij} = \left(ij\middle|ij\right) = \left(ji\middle|ji\right) = J_{ji} \tag{6.60}$$

$$K_{ij} = \left(ij\middle|ji\right) = \left(ii\middle|jj\right) = \left(jj\middle|ii\right) = \left(ji\middle|ij\right) = K_{ji} \tag{6.61}$$

$$L_{ij} = \left(ii\middle|ij\right) = \left(ii\middle|ji\right) = \left(ij\middle|ii\right) = \left(ji\middle|ii\right) \tag{6.62}$$

Note here that within the quantum-hybrid integral notation, switching subscripts does not imply equality in the expressions, as in the Coulomb and exchange integrals

$$L_{ij} \neq L_{ji} \tag{6.63}$$

An important quality of the HF method is that the Hamiltonian of the system and the wavefunction that is found to best approximate the solution satisfy the assumptions of the variation theorem, Equation (6.41). The HF energy therefore is always greater than the actual energy of the system, unless when the energies are equal and the approximating wavefunction of the system is exactly the real wavefunction. A wavefunction approximation that gives a lower energy than previous approximations implies an improved description.

6.5 Partitioning of the Electron Density using Atoms in Molecules Theory

It has been theoretically proven that all the properties of a molecule can be found through the electron density ρ of a molecule (Hohenberg and Kohn, 1964). This is helpful, since concepts like an atom within a molecule or bonds between these atoms are not obvious from the wavefunction, which is dependent on six variables (three position, three

momentum) for each particle in the system and exists in complex space. Atoms and bonds do have meaning in real space and are reflected in the electron density.

For a system with multiple electrons, ρ is expressed as the integration of the square of the wavefunction of the system over all electron spins and coordinates except one

$$\rho(\mathbf{r}) = \int \cdots \int |\Psi|^2 \, d\mathbf{r}_2 d\mathbf{r}_3 \ldots d\mathbf{r}_{n_e} \qquad (6.64)$$

The electron density is a real-valued function over all space and describes the probability of finding an electron at the differential space around the position $\mathbf{r}$.

For example, the positions of the nuclei are able to be found through the electron density by the existence of local maxima. The type of atomic nucleus at a particular position is related to the electron density of the position by (Bader, 1990)

$$\rho(\mathbf{r}_a) \approx \frac{Z_a^3}{2} \qquad (6.65)$$

The existence of bonds within a molecule is also seen within the electron density. A bond exists on a line between two atoms if there exists a saddle point in the electron density. This point is a local minimum on the line connecting the two nuclei, while it is a local maximum along the two directions perpendicular to the bond.

The electron density also holds information on how to partition a molecule into representative constituents. In theory, any partitioning is allowable. Within a partition, the average value for a classical property is calculable using Equation (6.35) where the space of integration is the partition. However, there is no guarantee that the properties found within an arbitrary partition yields physically significant properties. Atoms in

Molecules (AIM) theory (Bader, 1990) establishes that surfaces that define physically-significant partitions of the molecule satisfy the zero-flux condition

$$\nabla \rho(\mathbf{r}) \cdot \mathbf{n}(\mathbf{r}) = 0 \qquad (6.66)$$

where $\nabla \rho$ is the gradient of the electron density and $\mathbf{n}(\mathbf{r})$ is the vector normal to the bounding surface at $\mathbf{r}$. A system that is bound by a surface or surfaces that satisfies Equation (6.66) is called a proper open system. Boundary conditions satisfying Equation (6.66) yield the same mathematical consequences as the boundary conditions of isolated quantum mechanical systems. Therefore, one can expect the same physical significance in an average property calculation within a proper open system as for any general quantum system. Also, the properties within a proper open system are said to be transferable from one quantum system to another if the electron density functionality is preserved.

Partitioning the electron density of a molecule with zero-flux surfaces results in a set of subsystems each containing one nucleus. Under AIM theory, these quantum atoms are thereby defined to be the electron density within the partitioned space and the nucleus that resides within the boundaries. Quantum atoms are the objects that, when assembled, create molecules. These atoms have properties that, when combined appropriately, yield the property of the molecule.

The properties for a quantum atom defined in AIM theory are calculated in the same manner as the properties of any isolated quantum system, by Equation (6.35), except that the integration is taken over the partition of the atom in real space

$$\overline{F}_A = \int \Psi^* \hat{F} \Psi d\tau_A = \int \rho \hat{F} d\tau_A \qquad (6.67)$$

where $d\tau_A$ is a differential element within atom A. These properties are additive; the sum of a given property over all the atoms in a molecule gives the molecular property. For atoms not wholly bounded by a surface satisfying Equation (6.66), as most atoms are, the integration bounds would be infinite on some rays. For computational purposes, the boundary is taken to be a distance at which the electron density is negligible. However, for structural properties, a bounding isodensity surface of somewhat more significant density, e.g. the 0.001 au surface, is included to create a finite space for the atom.

Atomic properties of interest in engineering applications are separated into structural and electrostatic properties. The volume of the atom is the space bounded by zero-flux boundaries and the isodensity surface. The total surface area is the size of the zero-flux boundaries and the isodensity surface. The exposed surface area is solely the area of the bounding isodensity surface. The partial charge of the atom is the amount of electron density with the bounding surfaces. The dipole and quadrupole moments detail the displacement of the center of negative charge off the nucleus.

The energy of an atomic basin is not as straightforward as above. It is not possible to separate the Hamiltonian of the system into parts that reveal the contributions of atomic basins without violating the indistinguishability of the electrons in the molecule. An intermediate function, the energy density, is defined, and this quantity is integrated within Equation (6.67) to give the energy of an atom within AIM theory. This energy, as with the other properties of atoms, is additive and gives the energy of the molecule when added together.

6.6 Summary

First-principles knowledge of molecules and electron densities allows for the detailed study of molecules at very basic levels. Although the rigorous orbital information given by the solution of the hydrogen atom is not available for a molecular system, such functional forms are applicable to the solutions of molecular wavefunctions and, later, to the approximations of electron densities involved in intermolecular interactions.

The HF method is an example of how complex problems of a finite number of discrete entities can be iteratively solved. This method offers a first look of what orbitals within a molecule may look like, and it serves as a first approximation to more contemporary and complete solutions to the molecular Hamiltonian: the Møller-Plesset methods described later. Frequently used integrals, such as the two-electron integrals, have been defined within the context of the HF derivations, as these will be used extensively in the study of molecular interactions.

AIM theory offers a concise way of describing portions of molecules through the electron density, rather than the wavefunction or orbital description. Rigorous definitions of the structural and electrostatic properties, and the subsequent calculations of such properties, are usable by engineers through the group-contribution methodology, where portions of molecules are assembled and their contributions are added to give an approximation to the macroscopic property. AIM theory states that the summation for the functional group properties, when found rigorously, exactly gives the molecular property. This is developed further in this work to give thermodynamic properties for macroscopic systems.

CHAPTER 7

COMPUTATIONAL CHEMISTRY

User-friendly software packages employing computational chemistry algorithms are allowing scientists and engineers to use powerful numerical routines without being fully knowledge of quantum theory. The software computes approximations of the solution to the molecular Hamiltonian, thereby giving a numerical solution to the wavefunction. The approximated wavefunction is assumed to be the product of molecular orbitals, and each of these orbitals is approximated as a series of functional forms similar to orbitals within a hydrogen atom. All these methods are made available by Gaussian 98W (Gaussian Inc., 1998), a program that holds most of the contemporary techniques for the calculation of a molecular wavefunction. The full wavefunction is also made available from the program, making molecular and functional group properties calculable using more specific software.

This chapter reviews theoretical methods used within contemporary computational chemistry. Density functional theories are considered and compared to classical *ab inito* methods and perturbation theories. The functions that are used to approximate wavefunctions are reviewed, as these functional forms are important to the analysis of the resulting wavefunction. How Gaussian 98W is used within this work, including a list of important keywords and sample input files, is presented.

7.1 *ab initio* versus Density Functional Theory

As of now, methods that offer a solution or solution algorithm to the problem of determining the absolute energy of a molecule and the form of the electron density are separable into three categories. The fastest and least rigorous methods are called semi-empirical methods, since the Hamiltonian is not treated in its original terms and two-electron integrals are not calculated at every step. Modern schemes in this category include the AM1 (Dewar, et al. 1985) and PM3 schemes (Stewart, 1989). The other two methods are more strongly rooted in theory. Methods based on density-functional theory (DFT) look to determine the electron density from first principles, while *ab initio* methods attempt to calculate the system wavefunction from first principles.

The motivation behind the study of DFT methods is that the electron density holds all the information necessary to determine the properties of a molecule (Hohenberg and Kohn, 1964). This eliminates the need to determine the wavefunction and its numerous degrees of freedom. In contrast, the electron density is merely a function in the three spatial coordinates. A theorem within DFT states that, if one knows the ground-state electron density, ρ_0, of a quantum system, the number of electrons within the system, n_e, is calculable

$$\int \rho_0 d\tau = n_e \tag{7.1}$$

Also, one may determine the position of the nuclei in a molecular system uniquely using ρ_0, as shown in Section 6.5.

An algorithm to find the ground-state electron density of a molecular system has been proposed (Kohn and Sham, 1965). The method separates the properties of a real

system into the properties of non-interacting electrons and non-ideal effects. The ground-state energy of the molecule is thereby separated into four contributions

$$E_0[\rho_0] = \int \rho_0(\mathbf{r}) v(\mathbf{r}) d\mathbf{r} + \bar{K}_{ni}[\rho_0] + \frac{1}{2}\int \frac{\rho_0(\mathbf{r}_1)\rho_0(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2 + E_{xc}[\rho_0] \qquad (7.2)$$

where $E_0[\rho_0]$ is the ground state energy and is a functional of ρ_0, the first term is the electron density interacting with the field v created by the nuclei, the second term is the mean kinetic energy of the non-interacting system, $\bar{K}_{ni}$, the third term is the classical expression of electron-electron repulsion in terms of the smeared out electron clouds, and the last term is the exchange-correlation functional, E_{xc}, which holds the nonidealities of the kinetic energy and the electron-electron repulsion within the real, interacting system. All the terms in the energy expression are found rigorously, except for E_{xc}.

The problem within DFT is reduced to finding the form of E_{xc}. The first approximation offered is the local-density approximation (Hohenberg and Kohn, 1964) given by

$$E_{xc}^{LDA}[\rho_0] = \int \rho_0(\mathbf{r}) E_{xc}^{UEG}(\rho_0(\mathbf{r})) d\mathbf{r} \qquad (7.3)$$

where E_{xc}^{UEG} is the exchange-correlation within a uniform electron gas and depends on the magnitude of the electron density. This expression is shown to be a good approximation for slowly varying electron densities within real systems. A more general form of Equation (7.3) attempts to incorporate more complicated electron densities by including the gradient of the electron density explicitly in the integrand

$$E_{xc}[\rho_0] = \int f(\rho_0(\mathbf{r}), \nabla \rho_0(\mathbf{r})) d\mathbf{r} \qquad (7.4)$$

A more modern approach is to separate the exchange contribution from the correlation contribution and assume that the exchange within DFT contributes the same energy within the Hartree-Fock energy of Equation (6.52). Functionals using this concept are called hybrid functionals, and the most commonly used DFT methods today, such as B3LYP (Becke, 1992) and B3PW91 (Perdew, et al. 1996) employ hybrid functionals. The above approximations to E_{xc} usually contain parameters founds empirically, whether correlated to experimental data or high level *ab initio* calculations.

A review of modern computational methods summarizes the strengths and weaknesses of DFT approaches (Head-Gordon, 1996). Since DFT strives for the electron density directly rather than the wavefunction representation, fewer functions (smaller basis sets, discussed in subsequent sections) are needed to reproduce solutions (Johnson, et al. 1993). Also, the calculation method employed is similar to that of the HF methods, and this is faster than methods that employ perturbation or correlation corrections (such as MP2, which is addressed in the following section). For a given amount of computational power and time, this allows for calculations on molecules similar in size to those available for the HF method and larger than those available to more accurate *ab initio* methods. The results for ground-state geometries and energies are similar to that of the MP2 method, which are better than those found with the HF method.

The downsides of the DFT theory lie in the exchange-correlation terms approximated by Equations (7.3) and (7.4). Since arbitrary forms of these equations are introduced into the theory, questions arise as to whether this is a method still based on first principles (Levine, 2000). The variation theorem, a version of which exists for DFT, cannot be employed in DFT due to the use of arbitrary forms. Also, improvements

within the expression for E_{xc} cannot be made in a systematic manner as in perturbation methods. New functionals are introduced and must be tested with calculation, since no theoretical framework exists to assess the benefits of one functional over others.

For these reasons, perturbation methods are used in this work instead of DFT methods. With increased computational power and refined algorithms, higher perturbations are able to be made on current, already quite accurate, results of molecular properties.

7.2 Perturbation Theory and the Møller-Plesset Expansion

Some basic assumptions made to facilitate calculation within the HF method in Section 6.4 render the results somewhat undesirable. In the above presentation, there had been no mention of the HF energy, Equation (6.42), being exactly the eigenenergy in the molecular Hamiltonian. Errors arise from the central-field approximation in Equation (6.49) and in the treatment of each electron separately in the iterative process. Electron-electron repulsions are included in the central-field approximation, but only in an average way.

Methods have been developed to alleviate some of the error from the averaging procedure in the HF method. Most methods use the HF energy as a first-order approximation to the real energy, and small correcting factors are then added to gradually reduce the errors. Perturbation treatments of the ground-state wavefunction and the ground-state energy for the electronic Hamiltonian in Equation (6.34) aim to express Ψ_0 and E_0 as a series expansion in λ. First, the Hamiltonian of the system is separated into an unperturbed portion $\hat{H}^0$ and a perturbation $\hat{H}'$ whose effect on the system is smaller

in magnitude. The latter is thereby multiplied by the perturbing parameter to denote this. The electronic Hamiltonian for the molecule, expressed henceforth as $\hat{H}$, becomes

$$\hat{H} = \hat{H}^0 + \lambda\hat{H}' \tag{7.5}$$

and the Hamiltonian equation for the system in its ground state is

$$\left(\hat{H}^0 + \lambda\hat{H}'\right)\Psi_0 = E_0\Psi_0 \tag{7.6}$$

Next, the assumption is made that the wavefunction and energy are functions of λ

$$\Psi_0\left(\mathbf{r}_1,\mathbf{r}_2,...\right) \rightarrow \Psi_0\left(\lambda,\mathbf{r}_1,\mathbf{r}_2,...\right) \text{ and } E_0 \rightarrow E_0\left(\lambda\right) \tag{7.7}$$

These functions are expanded as a Taylor's series around $\lambda = 0$

$$\Psi_0 = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!}\left(\frac{\partial^i\Psi_0}{\partial\lambda^i}\right)\bigg|_{\lambda=0} = \sum_{i=0}^{\infty} \lambda^i\Psi_0^{[i]}$$

$$E_0 = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!}\left(\frac{d^iE_0}{d\lambda^i}\right)\bigg|_{\lambda=0} = \sum_{i=0}^{\infty} \lambda^iE_0^{[i]} \tag{7.8}$$

Inserting these expansions into Equation (7.6) and expressing the functions out to the second order gives the following relationship

$$\left(\hat{H}^0 + \lambda\hat{H}'\right)\left(\Psi_0^{[0]} + \lambda\Psi_0^{[1]} + \lambda^2\Psi_0^{[2]} + ...\right)$$
$$= \left(E_0^{[0]} + \lambda E_0^{[1]} + \lambda^2 E_0^{[2]} + ...\right)\left(\Psi_0^{[0]} + \lambda\Psi_0^{[1]} + \lambda^2\Psi_0^{[2]} + ...\right) \tag{7.9}$$

This method allows for the systematic solution to the terms in the Taylor's expansions by equating like powers of λ. Equating the terms multiplied by λ^0 on the left with those on the right gives

$$\hat{H}^0\Psi_0^{[0]} = E_0^{[0]}\Psi_0^{[0]} \tag{7.10}$$

This reduced problem states that the first-order approximation to the ground-state wavefunction and ground-state energy evolve from the unperturbed Hamiltonian. Notice

this is the expression when the perturbation parameter λ is set to zero. This means that all perturbations are excluded, thus revealing the unperturbed problem.

Equating terms multiplied by λ^1 gives the relation

$$\hat{H}'\Psi_0^{[0]} + \hat{H}^0\Psi_0^{[1]} = E_0^{[1]}\Psi_0^{[0]} + E_0^{[0]}\Psi_0^{[1]} \qquad (7.11)$$

This is soluble once the unperturbed problem is solved. Expressions for the zeroth-order wavefunction and energy are inserted into Equation (7.11) to yield a soluble problem for $\Psi_0^{[1]}$ and $E_0^{[1]}$. The derivations for these general expressions are not shown here, but the results are

$$E_0^{[1]} = \left\langle \Psi_0^{[0]} \middle| \hat{H}' \middle| \Psi_0^{[0]} \right\rangle$$

$$\Psi_0^{[1]} = \sum_{i \neq 0} \frac{\left\langle \Psi_0^{[0]} \middle| \hat{H}' \middle| \Psi_0^{[0]} \right\rangle}{E_0^{[0]} - E_i^{[0]}} \Psi_i^{[0]} \qquad (7.12)$$

where the change in subscripts denotes system states with energies higher than that of the ground state.

It is seen within Equation (7.9) how the problems containing the higher-order corrections are posed. As suggested by the solution to the first-order problem in Equation (7.12), knowledge of both the lower-order solutions of the ground state and of all possible excited states are necessary. This complexity continues to increase for all higher-order solutions, making the determination of further corrections difficult.

The Møller-Plesset (MP) Perturbation Theory (Møller and Plesset, 1932) expresses the molecular Hamiltonian as an unperturbed portion and a perturbation in order to apply the method described above. The unperturbed portion is chosen to be

$$\hat{H}^0 = \sum_{i=1}^{n_e} -\frac{1}{2}\nabla_i^2 + \sum_{i=1}^{n_e}\sum_{j=1}^{n_e}\left(\hat{j}_j(i) - \hat{k}_j(i)\right) \qquad (7.13)$$

where the Coulomb operator $\hat{j}_j(i)$ acting on a general orbital f is given by

$$\hat{j}_j(1)f(1) = f(1)\int\frac{\left|\psi_j(2)\right|^2}{\tilde{r}_{12}}dv_2 \tag{7.14}$$

and where the exchange operator $\hat{k}_j(i)$ acting on a general orbital f is given by

$$\hat{k}_j(1)f(1) = \psi_j(1)\int\frac{\psi_j^*(2)f(2)}{\tilde{r}_{12}}dv_2 \tag{7.15}$$

The perturbation in the problem is thereby the difference between the actual electronic

Hamiltonian of the system, Equations (6.34) and (7.13)

$$\hat{H}' = \hat{H} - \hat{H}^0 = \sum_{i=1}^{n_e-1}\sum_{j=i+1}^{n_e}\frac{1}{\tilde{r}_{ij}} - \sum_{i=1}^{n_e}\sum_{j=1}^{n_e}\left(\hat{j}_j(i) - \hat{k}_j(i)\right) \tag{7.16}$$

Applying the perturbation theory, the zeroth-order problem has the same form as

Equation (7.10). Once the zeroth-order problem is solved, the first-order correction is

found using Equation (7.12). The application of MP perturbation theory states that the

energy of the system taken to the first-order correction is exactly the Hartree-Fock energy

$$E_{HF} = E_0^{[0]} + E_0^{[1]} \tag{7.17}$$

Any improvement over the result from the HF method is made beginning with the

calculation of the second-order correction to the energy and the system wavefunction.

Such a treatment is called a calculation at the MP2 (Møller-Plesset, 2nd order) level of

theory. Calculations at higher levels are possible, and the theoretical levels are called the

MPn levels of theory.

Calculations of energies and wavefunctions for molecular systems using the MPn

methods give a better representation of the environment electrons encounter. Not only

are the electrons subject to the normal Coulombic interactions, the excited states of the

molecule are also included in the solution and interact with the ground-state configuration. Necessary in the calculation of higher approximations of the wavefunction and energy are the wavefunctions of the excited states within lower approximations, as shown in Equation (7.12). This added information changes the states of the electrons by bringing in quantum effects that excited levels bear on the system in the ground state.

For whatever level of theory applied, a HF solution must be calculated to assemble the first-order wavefunction for MPn. The most common calculations are the MP2 levels, and they require more computational effort than HF and DFT methods. Further improvements can be made with MP3 and MP4 calculations. MP5 and further calculations can be done, but these are extremely expensive with present day computational resources.

7.3 Basis Sets

Up to this point, the theory behind calculating the energy and wavefunction of a molecular system has been the focus. Approximate solution methods exist that give a good representation of the molecular structure and the electron density profile upon solving the molecular Hamiltonian. General statements about wavefunctions have been used in the derivations of these methods. Actual guesses of the form of these wavefunctions are needed to facilitate calculations and to express the electron density for further use in molecular modeling. Two main types of functional forms are used in the expression of the wavefunction: Slater-type orbitals (STOs) and Gaussian-type functions (GTFs). Also, these expressions are used in two types of representations of the orbitals within which electrons reside: atomic orbitals (AOs) and molecular orbitals (MOs).

As in the first wavefunction guess in the HF method, Equation (6.44), a product of hydrogen-like orbitals gives a possible functional form for orbitals within the molecular system. The hydrogen-like wavefunctions take the form of Equation (6.13) with the radial contribution being a product of a polynomial and exponential in terms of $\tilde{r}$. A STO closely resembles a hydrogen-like orbital by taking a simplified form

$$s_{nlm} = \frac{(2\zeta)^{n+1/2}}{\sqrt{(2n!)}} \tilde{r}^{n-1} e^{-\zeta \tilde{r}/n} Y_l^m (\theta, \phi) \tag{7.18}$$

where ζ is the STO exponent that determines the spatial extent of the wavefunction, and the prefactor normalizes the wavefunction. Like the hydrogen-like atom, n, l, and m are quantum numbers that determine the state of the orbital. The radial portion of Equation (7.18) takes the form of a gamma distribution regularly used in statistics and probability theory. The orbital exponents are used as a parameter either to minimize the variation integral (and thus minimize the system energy) or to represent a single orbital by a linear combination of STOs with judiciously chosen exponents.

An orbital of the form of Equation (7.18) give a very good function representation of real orbitals. However, calculations of the above numerical methods using STOs are computationally intense. Specifically, the calculation of two-electron integrals of Equation (6.56) is quite complicated for the non-spherical STOs. To alleviate the computational time issue with STOs, functions containing a normal distribution rather than a gamma distribution are used (Boys, 1950). These GTFs take the form

$$g_{ijk} = \left(\frac{2\xi}{\pi}\right)^{3/4} \left[\frac{(8\xi)^{i+j+k} i! j! k!}{(2i!)(2j!)(2k!)}\right]^{1/2} \tilde{x}^i \tilde{y}^j \tilde{z}^k e^{-\xi \tilde{r}^2} \tag{7.19}$$

where ξ is the GTF exponent that determines the spatial extent of the wavefunction, where all the coordinates are distances from the nucleus on which the orbital resides, and where the prefactor normalizes the wavefunction. The integers i, j, and k are chosen so the GTF has the same functionality as the orbital it represents. For the case $i + j + k = 0$, the orbital takes a spherically symmetric form and thus represents an s-type orbital. For the case $i + j + k = 1$, the GTF has one of the three Cartesian coordinates multiplying the radial factor. This function has the form of the orbitals in Equations (6.26) through (6.29), and therefore this GTF represents a p-type orbital. Similarly for the case $i + j + k = 2$, where there will be either two Cartesian coordinates or a power of two on one of the coordinates. This represents a d-type orbital not unlike the function in (6.30). Integers that sum to three represent an f-type orbital and so on.

A comparison between the functionalities of a STO and a GTF are shown in Figure 7.1. The STO reflects the functionality of a hydrogen-like orbital, with the discontinuous first derivative at the nucleus $(r = 0)$, while the GTF reflects the functionality of a normal distribution. Of more importance than the behavior at the origin is the behavior at the tail (larger r) of the wavefunction and the electron densities. At these distances, the differences between the functions become small, and with the addition of more GTFs as a description of a STO, this difference vanishes.

The utility of GTFs to represent orbitals lies in the ease with which a two-electron integral like Equation (6.56) is computed. This is due to the mathematical identity which allows a product of two normal distributions to be rewritten as a single normal distribution
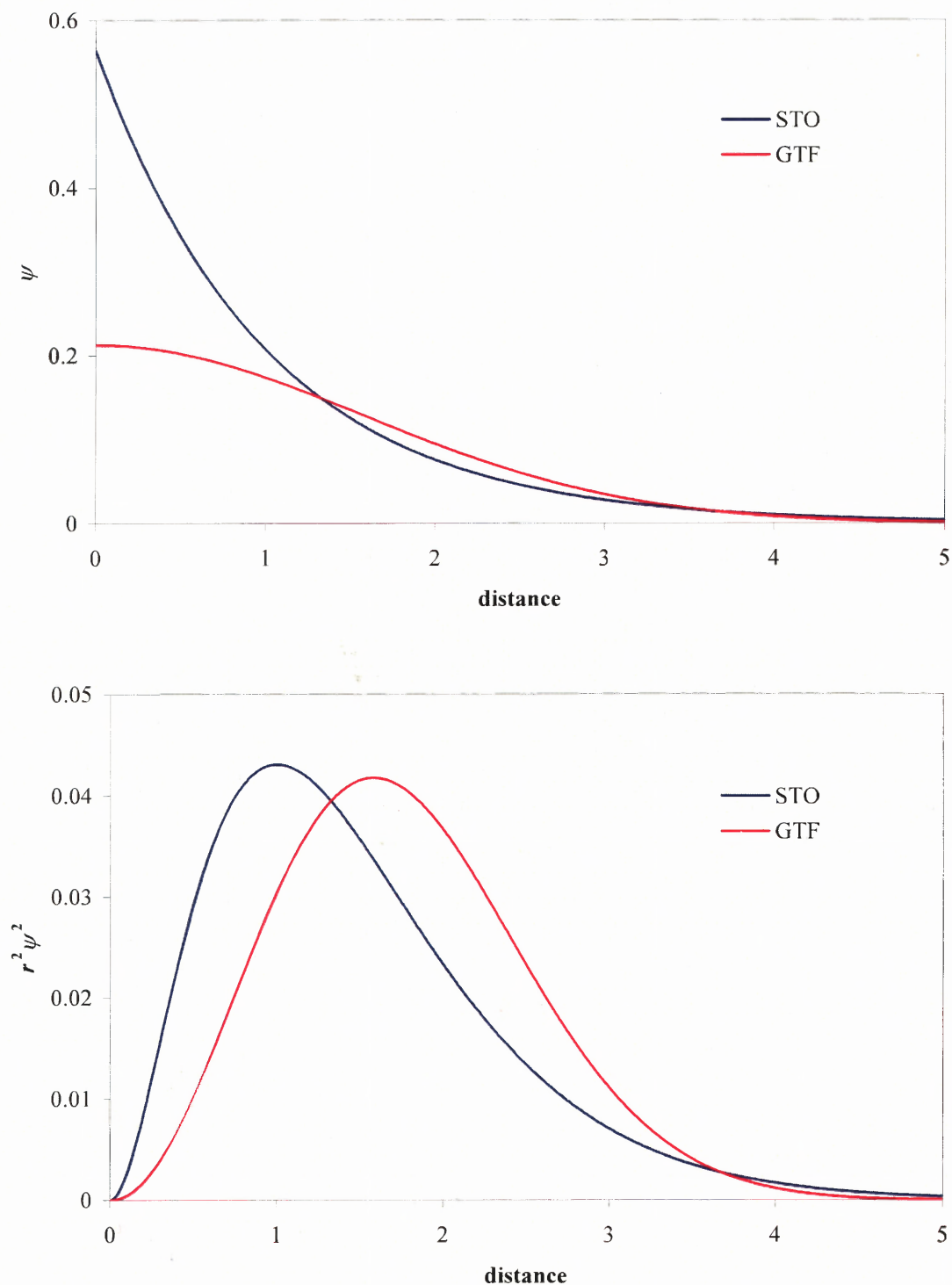
Figure 7.1 Comparison of a STO $(\zeta = 1.0)$ and a GTF $(\xi = 0.2)$. The upper graph compares the wavefunctions, while the lower graph compares the radial distribution functions.

$$C_A g_{1s} \left(\mathbf{r} - \mathbf{R}_A \right) C_B g_{1s} \left(\mathbf{r} - \mathbf{R}_B \right) = C_A C_B C_{AB} g_{1s} \left(\mathbf{r} - \mathbf{R}_P \right) \tag{7.20}$$

where C_A and C_B contain normalization constants and axial factors for the uncombined GTFs, $\mathbf{R}_A$ and $\mathbf{R}_B$ are the location of the orbital centers, C_{AB} is a prefactor that emerges from the combination and is given by

$$C_{AB} = e^{\frac{-\xi_A \xi_B}{\xi_A + \xi_B} |\mathbf{R}_A - \mathbf{R}_B|^2} \tag{7.21}$$

and where the location of the intermediate center $\mathbf{R}_P$ is given by

$$\mathbf{R}_P = \frac{\xi_A \mathbf{R}_A + \xi_B \mathbf{R}_B}{\xi_A + \xi_B} \tag{7.22}$$

An important integral encountered throughout this work is the overlap integral, S, between two $1s$-type orbitals. This expression is given by Equation (7.20) integrated through all space. For non-normalized GTFs

$$S = \left\langle g_{1s} \left(\mathbf{r} - \mathbf{R}_A \right) \middle| g_{1s} \left(\mathbf{r} - \mathbf{R}_B \right) \right\rangle = \left\langle g_{1s} \left(A \right) \middle| g_{1s} \left(B \right) \right\rangle = \left(\frac{\pi}{\xi_A + \xi_B} \right)^{\frac{3}{2}} e^{\frac{-\xi_A \xi_B}{\xi_A + \xi_B} r_{AB}^2} \tag{7.23}$$

where r_{AB} is the distance between the centers of the orbitals. Another important integral in this work is the electron-nucleus integral. For non-normalized $1s$-type GTFs, this expression is given by

$$\left\langle g_{1s} \left(A \right) \middle| -Z_C / r_{1C} \middle| g_{1s} \left(B \right) \right\rangle = \frac{-2\pi}{\xi_A + \xi_B} Z_C e^{\frac{-\xi_A \xi_B}{\xi_A + \xi_B} r_{AB}^2} F_0 \left(\left(\xi_A + \xi_B \right) r_{PC}^2 \right) \tag{7.24}$$

where r_{PC} is the distance between the nucleus center and the point described by Equation (7.22), and where F_0 contains the error function and is a function defined as

$$F_0 \left(t \right) = \frac{1}{2} \sqrt{\frac{\pi}{t}} erf \left(\sqrt{t} \right) \tag{7.25}$$

A product of four different GTFs in the two-electron integral given by Equation (6.56) is reducible to one GTF with an exponent containing contributions of the original four exponents. The integrals reduce to concise mathematical expressions readily available for calculation. The general expression for a two-electron integral for four different non-normalized $1s$-type GTFs is given by

$$\left\langle g_{1s}\left(A\right)g_{1s}\left(B\right)\middle|r_{12}^{-1}\middle|g_{1s}\left(C\right)g_{1s}\left(D\right)\right\rangle = \frac{2\pi^{5/2}}{\left(\xi_A+\xi_B\right)\left(\xi_C+\xi_D\right)\sqrt{\xi_A+\xi_B+\xi_C+\xi_D}}$$
$$\times e^{\frac{-\xi_A\xi_B}{\xi_A+\xi_B}r_{AB}^2}\,e^{\frac{-\xi_C\xi_D}{\xi_C+\xi_D}r_{CD}^2}$$
$$\times F_0\left(\frac{\left(\xi_A+\xi_B\right)\left(\xi_C+\xi_D\right)}{\xi_A+\xi_B+\xi_C+\xi_D}r_{PQ}^2\right) \tag{7.26}$$

where the center $\mathbf{R}_Q$ is defined in (7.22) where A and B are replaced by C and D.

The above integrals are the basis for computations of integrals with any type of orbital considered. For integrals with higher orbitals, a derivative is taken with respect to the position of the orbital center (Boys, 1960). For example, the overlap integral (7.23) of an s-type orbital with a p_z-type orbital is given by

$$\left\langle g_s\left(A\right)\middle|g_{P_z}\left(B\right)\right\rangle = \frac{1}{2\xi_B}\left(\frac{\partial}{\partial B_z}\left\langle g_s\left(A\right)\middle|g_s\left(B\right)\right\rangle\right) \tag{7.27}$$

where B_z is the z-coordinate for the center of the p_z-type orbital, and the function upon which the differential is taken is given by Equation (7.23). The direction of the p-type orbital is altered by changing the axis along which the derivative is taken. To determine the overlap of two p-type orbitals, another derivative is taken with respect to the center location of the orbital labeled A in the given direction. In general, integrals with higher orbitals are found by applying the proper derivatives to Equations (7.23), (7.24) and

(7.26). A generalized set of integral equations with any type of GTF-representation of orbitals has been compiled for use within numerical methods (Browne and Poshusta, 1962).

Wavefunctions containing STOs do not have such a simplified treatment when the above integrals are calculated. This complication has eliminated the use of STOs in most modern calculations of HF, DFT and MPn energies and wavefunctions, except for the small molecules. However, the functionality of a STO is very important and is attempted to be preserved. Within calculations of wavefunctions the exponents of a linear combination of GTFs are selected to approximate the functional form of a STO. An example is the STO-3G basis set, where a $1s$ STO is approximated with the form

$$s_{1s} \approx \sum_{i=1}^{3} c_i \chi_{1s,i} \tag{7.28}$$

where χ_i is a GTF called a primitive and has a different exponent ξ_i. The coefficients c_i and the primitive exponents ξ_i are found by maximizing the integral $\left\langle s \left| \sum_i c_i \chi_i \right. \right\rangle$. This allows the approximate representation of theoretically appealing STOs with more efficient GTFs.

The sum of functions in Equation (7.28) is considered a series representation of the orbital. In this view, a better representation can be made if more functions are included. For example, if an orbital on a hydrogen atom within a molecule is desired, the reasonable starting point for representation of that orbital is a sum of three primitives, as in Equation (7.28). A better representation of orbitals about the hydrogen includes six primitives. And although not clearly suggested by the problem, a more exhaustive representation is to include primitives in the form of higher orbitals, such as a set of

p-type primitives, a set of *d*-type primitives and so on. These add-ons to the original set of orbital functions give a more complete description of the space and energy levels an electron may occupy around the hydrogen atom. It is readily accepted that the more functions included in the orbital description, the better the approximation to the real system.

The number and types of primitive Gaussian functions used to describe orbitals around atoms within molecules is called the basis set. A large number of examples have been developed for use in present computational chemistry methods (Hehre, et al. 1986). Common basis sets used in calculations are abbreviated to define how large these functions are. For example, a common basis set called 3-21G uses a linear combination of three primitives for non-valence, or inner-shell, orbitals (one $1s$ orbital for the row Li-Ne; one $1s$, one $2s$, and three $2p$ orbitals for the row Na-Ar; etc.), and a combination of a single GTF with a linear combination of two primitives for the valence orbitals (one $1s$ orbital for H and He; one $2s$ and three $2p$ orbitals for the row Li-Ne; etc.). A similarly defined basis set is the 6-31G, where a larger number of primitives are included in the inner-shell and valence orbital descriptions.

Additions to these basis sets can better describe specific physical effects that electrons encounter within a molecular environment. Polarization functions are added to expand the space in which an electron resides and accommodate for the field imposed by other atoms in the molecule. The inclusion of polarization functions in the above basis sets is denoted by a * after the G. The basis set 6-31G* adds six *d*-type GTFs on the rows Li-Ne and Na-Ar. A second star, as in the 6-31G** basis set, adds also three *p*-type functions to H and He. This basis function is also called 6-31G(d,p) to explicitly

show the types of polarization functions. Also made available in calculations are further polarization orbitals of the f-type on the rows Li-Ne and Na-Ar and d-type for H and He. Such a basis set is denoted by 6-31G(f,d).

Advantages are found when very diffuse orbitals (GTFs with small exponents) are added to the description. This addition to the above basis sets is denoted by a + in the name. Therefore, 6-31+G includes one s-type and three p-type GTFs with small exponents on all non-hydrogen atoms. The 6-31++G basis set also adds a diffuse s-type orbital on the hydrogen atoms.

A more systematic approach to a series-like description of an orbital exists in the correlation-consistent, polarized valence, n zeta (cc-pVnZ) basis sets (Dunning, 1989), where n=D,T,Q,5,6 describes the multiplicity of the valence shell (double, triple, quadruple, etc.). The statement correlation-consistent means that the basis set is useful in theoretical treatments that describe the interactions of excited states, such as the MPn methods. The polarized functions described above, previously denoted by ** or (d,p), are included in this basis set without denotation. The cc-pVnZ set has available diffuse functions similar to the + and ++ notation above. To include these, an AUG- is added as a prefix in the basis set name. Thus, a double zeta basis set with polarization functions is denoted as the AUG-cc-pVDZ basis set.

The multiplicity of the valence shell, or the zeta, is where the systematic addition of functions can be examined. As calculations with higher zeta multiplicities are performed, inferences can be made on the greater multiplicities, such as an infinite multiplicity that would exhaust all possible wavefunction descriptions and give the actual

wavefunction behavior. The number of functions in the cc-pVnZ sets dwarfs that of the 6-31G set, so a large computational effort is necessary to proceed along such a track.

7.4 Computational Chemistry Software: Gaussian 98W

Gaussian 98W (G98W) (Gaussian Inc., 1998) is a software package that links programs that calculate myriad properties of molecules within the gas phase or solution. The package is a user-friendly tool that can predict properties of an isolated molecule, such as absolute molecular energies, molecular structure, orbital occupancies, atomic charges, multipole moments, polarizabilities, and electrostatic potentials surrounding molecules. Coupled with a closely associated graphical user interface (GUI) called GaussView (Gaussian Inc., 2000), the package serves as a user-friendly tool to implement the theoretical and numerical methods mentioned above.

To conduct a calculation on a molecule, one must become familiar with the format of the input files necessary to direct G98W. For this work, the instructions are separable into five different groupings: the header statements, the route section with associated keywords, the title of the calculation, the Z-matrix, and a filename to end the input file. A full description of a typical input file for G98W is available with the software package and on the Internet, and examples of input files used in this work are offered in Appendix C.

The header statements declare hardware information to G98W. The amount of random-access memory (RAM) available for calculation is declared beginning with the '%mem=' command. For instance, a computer with 512 megabytes (MB) of RAM may have about 400 MB available for a calculation. To utilize this memory, a header

statement should read '%mem=400MB'. A file called a checkpoint file is created to store molecular system information during and at the end of a calculation. A file is created when using the '%chk=' header. For systems with multiple processors, the header '%nproc=' is stated to utilize the hardware.

The route section includes the level of theory (HF, MP2, MP4, B3LYP, etc...), the basis set, and any other secondary calculations to be performed. Each line in the route section must begin with a '#' symbol. For instance, a common calculation is to invoke the HF level of theory with a 6-31+G* basis set. The first statement within the route section reads 'HF/6-31+G*' followed by keywords. An oft-used calculation within this work invokes a restricted MP2 level of theory where electron correlation is applied to all electrons and using the AUG-cc-pVDZ basis set. The keyword line therefore begins with 'MP2(full)/AUG-cc-pVDZ' followed by keywords. A number of keywords most utilized in this work are outlined within the following paragraphs.

A line is devoted to the title of the calculation is usually included for bookkeeping purposes and does not influence the calculation procedure.

The Z-matrix is a set of coordinates that inputs the locations, or initial guess to the locations, of the nuclei in the molecule. The first line of the Z-matrix states the multiplicity (a value of 1 for closed-shell molecules) and the overall charge of the molecule. The following matrix describing nuclear positions can be stated in two forms: Cartesian coordinates or internal coordinates. A description of the Z-matrix in internal coordinates is not given here, since it is somewhat involved and necessitates the use of numerous images and graphs, and since GUIs such as GaussView allow for the construction of the Z-matrix without actual knowledge of how to write one from scratch.

A Z-matrix can be omitted from the input file when a checkpoint file exists for the molecule. The guess of the structure is thereby read from the checkpoint file using a keyword detailed later. Descriptions of how to create Z-matrices exist in the literature (Levine, 2000) as well as on the Internet.

A filename is necessary at the end of the input file to store information requested by the keywords. Of use in this work is the wavefunction resulting from the calculation. Therefore, a filename with the extension '.wfn' ends some of the input files. If this keyword or any other that requests a filename is not used, then the filename is not necessary.

When looking for the ground-state energy, ground-state wavefunction and associated electron density of an isolated molecule (as is done in this work), several steps must be taken to achieve this goal. Firstly, an initial guess to the locations of the nuclei must be determined. This task involves determining rotational conformers of molecules and their relative energies at low theoretical levels, so that the lowest conformer can be used for more intense calculation. Secondly, the geometry of the nuclei must be optimized at a higher theoretical level to achieve representative bond lengths and angles for the low energy conformer. This calculation is called a geometry optimization (OPT) and is performed by G98W in an iterative manner. For this reason, OPTs are typically performed at an intermediate level of theory or with moderately-sized basis sets, or both. Thirdly, the calculation that achieves the energy for this optimized structure is performed with higher theory and larger basis sets. This calculation is called the single point energy (SPE) calculation for the ground-state structure of the molecule. Finally, multiple SPE calculations are performed to determine effect electric fields have on the electron density,

thus determining the polarizability of the molecule. A more detailed algorithm of the methods utilized in this work is offered in Section 8.1.

A list of keywords available within G98W that are related to portions of this work is included in Table 7.1. The levels of theory applied within this work are the 'HF' level and the 'MP2(full)' level, and the basis sets used are the '6-31++G(d,p)' and 'AUG-cc-pVDZ' sets.

The execution of the conformer search and OPT calculations described in Section 8.3 are made easier with use of the GaussView program. The software offers a molecule builder for the user to connect atoms, alter angles and twist dihedral angles to achieve the desired initial guess of a geometry optimization. Given this graphical representation of the molecule, the user is then able to access an input file creator and change the calculation type and keywords. The most popular options are readily available, and users can supplement other keywords in an input box found within the input file creator. Some options used in this work are not included in GaussView, so the software is mostly seen as a convenient way to create the Z-matrices for geometry optimizations.

7.5 Wavefunction Output of Gaussian 98W

Gaussian 98W outputs the wavefunction solution to a molecular system when the 'output=wfn' keyword is used. This command outputs a PROAIMV wavefunction file that is subsequently used within AIM property calculators (discussed in Chapters 9 and 10). Instructions on what the layout of the .wfn file is, how to extract the wavefunction of the molecular system and how to calculate the electron density are given in Appendix D.

Table 7.1 Commonly Used Keywords in Gaussian 98W

keyword	section	description	common options	lines added/omitted
chk	header	checkpoint file declaration	%chk=molecule.chk	
mem	header	system memory available for calculation	%mem=300MB	
nproc	header	number of processors	%nproc=1	
hf or rhf	route	restricted HF method	hf/6-31G	
mp2 or rmp2	route	restricted MP2 method	rmp2/aug-cc-pvdz	
b3lyp	route	B3LYP DFT method	b3lyp/3-21G	
b3pw91	route	B3PW91 DFT method	b3pw91/6-31++G**	
3-21G	route	3-21G basis		
6-31G	route	6-31G basis		
6-31++G(d,p) or 6-31++G**	route	6-31++G** basis		
aug-cc-pvdz	route	AUG-cc-pVDZ basis		
opt	route	geometry optimization		
scf	route	single point energy calculation		
nosymm	route	removes orbital symmetry constraint		
geom	route	source of molecule input	geom=allcheck	no Z-matrix necessary
guess	route	source of initial guess to HF wavefunction	guess=read	
density	route	specifies level of theory to analyze	density=current	
aim	route	AIM analysis on molecule	aim=tight aim=cp	
field	route	applies an electric field	field=x+0.0070	
cube	route	samples space with a cube of points	cube=pot	molecule.cub at end of input file
population	route	orbital population or partial atomic charges	pop=mk pop=chelp	
units	route	controls units in the Z-matrix	units=au	
output	route	requests an output file	output=wfn	molecule.wfn at end of input file

7.6 Summary

The theoretical methods and basis sets described in this chapter offers background on the fundamental nature of the choices made for calculations of this work. Although offering advantages in computational time and accuracy, DFT methods are not used in this work because of the lack of a theoretical basis by which one can increase its effectiveness. Perturbation methods are employed with the goal of devising a series of calculation results that may extrapolate to results offered at greater levels of theory or sizes of basis sets.

The computational chemistry software now available to engineers allows for an easy application of these high levels of theory. Gaussian 98W used on a Microsoft Windows-based personal computer is employed in this work to determine a wide variety of properties for a particular molecule. Some of this information is sufficiently close to experimental results that the predicted results will be used to determine the thermodynamic properties of macroscopic systems. Since the full wavefunction approximation is also made available by this software, AIM properties of the molecule can be calculated as molecular information becomes necessary.

CHAPTER 8

PROPERTIES CALCULATION OF ORGANIC MOLECULES

Computational chemistry in this work is employed to determine information about molecular species without the use of experiment. This information is then used within interaction models and statistical fluid models (presented in subsequent chapters) that attempt to determine macroscopic system behavior from the molecular properties. This begins with the computation of approximations of the molecular wavefunctions available through the techniques reviewed in Chapters 6 and 7.

This chapter details the molecular computations on over 130 molecules analyzed within this work. The levels of theory and basis sets are given. The search for the low energy conformer, important in determining the most probable structure of the molecule, is explained. The method to determine the polarizability of molecules, an important property in intermolecular interaction theory previous found experimentally, is outlined. The computational results of the dipole moment and the polarizability are then compared to experiment, and the utility of several DFT and *ab initio* methods and basis sets at predicting experimental properties is tabulated and compared.

8.1 Calculation Method

Whenever computational chemistry is employed, one must take care in choosing the theoretical method and basis set that will best predict the properties of interest. Of interest in this work is the electron density profile of the molecule and the systematic methods with which to get better representations of the electron density as more

computational power arises. The former, of course, is the goal for all of computational chemistry, and methods are continually being refined to approach this goal. The latter states the need to use perturbation methods (MP2) rather than density functional theory methods. Also, the use of the correlation consistent, polarized valence basis sets (AUG-cc-pVDZ etc...) allows for this by making available the systematic increase of space-filling functions for the approximating wavefunction.

The computational chemistry methods employed in this work are used to determine the ground-state geometry of a molecule, the wavefunction (and thereby the electron density) of the molecular system, and how this wavefunction is affected by uniform electric fields in three perpendicular directions. The algorithm used to determine this information is as follows:

- The initial guess for the ground-state geometry is determined using a rotational conformer search. This is done by performing multiple geometry optimizations (OPTs) at a very low theoretical level and basis set (HF/6-31G) for timely results. More about this procedure and results are presented in Section 8.3.

- An OPT is performed with the initial guess of the ground-state geometry. This is performed at an intermediate level of theory and with a moderately-sized basis set (MP2(full)/6-31++G**). This calculation is meant to achieve a better representation of the bond lengths and angles than the conformer search had provided. An example OPT input file for G98W used in this work is made available in Appendix C.

- A single point energy (SPE) calculation is performed on the molecule with a high level of theory and large basis set (MP2(full)/AUG-cc-pVDZ) to achieve as good

a representation of the electron density as possible. The output from this calculation also includes the wavefunction description (using the 'output=wfn' keyword in G98W) and the critical points within the electron density (using the 'aim=cp' keyword in G98W). The wavefunction is made available in a separate .wfn file, and the critical points are found near the end of the standard G98W output file. Example SPE input and output file for G98W are made available in Appendix C of this work. A description of the .wfn file is available in Appendix D of this work.

- Three more SPE calculations are employed to determine the response of the electron density to electric fields in the x, y, and z directions. These calculations are used to determine the polarizability of the electron density. The level of theory and size of the basis set correspond to that of the SPE calculation of the ground-state wavefunction (MP2(full)/AUG-cc-pVDZ). Similar to the ground-state calculation, the outputs from this calculation include the wavefunction description of this system and the critical points within the electron density.

These four steps yield a total of eight output files that contain information necessary for further steps in this work. These include four .wfn files, one for the ground state and three for the polarized states, as well as four G98W output files that hold the critical points of the electron density. These files and the critical point information are used in AIM integration routines that determine properties for atoms and functional groups within the molecule. Also of importance is the .chk file that is used for the geometry in OPT and SPE calculations. However, the information on the geometry is also included in the G98W output file, so this file is seen as extraneous.

Also of importance to the application of computational chemistry is the level of hardware resources available to the software (here, G98W). The molecular-level calculations for this work are all performed on a single personal computer. This computer runs on a 1.8 Gigahertz Pentium 4 processor, 512 MB of RAM, and runs the Microsoft Windows 2000 Professional operating system. The total hard drive space is 32 GB, of which 20 GB is left free for the G98W program to utilize during calculation. The header line '%mem=400MB' is written to utilize the RAM available on the computer system. The computational methods chosen for the SPE calculations above are the most rigorous *ab initio* methods allowable with the computational resources listed.

8.2 Choice of Molecules

The molecules chosen for analysis in this work have several constraints. Firstly, as noted above, the computational resources available dictate the theoretical method and basis set that are attainable over the range of molecules considered. Also, the methods are affected by the number of electrons within the system; therefore only first-row heavy (non-hydrogen) atoms, specifically C, N, O, and F, are used to build molecules. The available analysis techniques of functional groups also dictate whether certain molecules can be pursued in this work.

The list of molecules calculated for this work is motivated by work that explored use of computational chemistry within models of the excess Gibbs energy (Wu and Sandler, 1991a). The main size constraint is to restrict the number of first-row heavy atoms to six. This size constraint is relaxed in groupings where more molecules are needed for analyzing purposes and where mixture systems of interest contain these larger

molecules. The list of molecular grouping with the molecule names and CAS registry numbers in parenthesis are as follows:

- Alkanes: methane (74-82-8); ethane (74-84-0); propane (74-98-6); butane (106-97-8); isobutane (75-28-5); pentane (109-66-0); isopentane (78-78-4); neopentane (463-82-1); hexane (110-54-3); 2-methylpentane (107-83-5); 3-methylpentane (96-14-0); 2,2-dimethylpentane (75-83-2); 2,3-dimethylbutane (79-29-8)

- Alkenes: ethane (74-85-1); propene (115-07-1); 1-butene (106-98-9); *trans*-2-butene (624-64-6); *cis*-2-butene (590-18-1); 2-methylpropene (115-11-7); 1-pentene (109-67-1); *trans*-2-pentene (646-04-8); *cis*-2-pentene (627-20-3); 2-methyl-2-butene (513-35-9); 3-methyl-1-butene (563-45-1); 1-hexene (592-41-6)

- Amines: methanamine (74-89-5); ethanamine (75-04-7); 1-propanamine (107-10-8); 2-propanamine (75-31-0); 1-butanamine (109-73-9); 2-butanamine (13952-84-6); 2-methyl-1-propanamine (78-81-9); 2-methyl-2-propanamine (75-64-9); 1-pentanamine (100-58-7)

- Diamines: dimethylamine (124-40-3); methylethylamine (624-78-2); diethylamine (109-89-7); methyl-n-propanamine (627-35-0); methylisopropanamine (4747-21-1)

- Triamines: trimethylamine (75-50-3); n,n-dimethyl-ethylamine (598-56-1); methyldiethylamine (616-39-7)

- Nitriles: ethanenitrile (75-05-8); propanenitrile (107-12-0); butanenitrile (109-74-0); 2-methylpropanenitrile (78-82-0)

- Alcohols: methanol (67-56-1); ethanol (64-17-5); 1-propanol (71-23-8); 2-propanol (67-63-0); 1-butanol (71-36-3); 2-butanol (78-92-2); 2-methyl-1-

propanol (78-83-1); 2-methyl-2-propanol (75-65-0); 1-pentanol (71-41-0); 2-methyl-2-butanol (75-85-4); 3-methyl-1-butanol (123-51-3); 1,2-propanediol (57-55-6); 1,3-propanediol (504-63-2); 1-methoxy-2-propanol (107-98-2)

- Ethers: dimethylether (115-10-6); methylethylether (540-67-0); diethylether (60-29-7); methylpropylether (557-17-5); methyl isopropylether (598-53-8); methyl tert-butyl ether (1634-04-4); ethyl tert-butyl ether (637-92-3)

- Aldehydes: methanal (50-00-0); ethanal (75-07-0); propanal (123-38-6); butanal (123-72-8); 2-methylpropanal (78-84-2)

- Ketones: propanone (67-64-1); butanone (78-93-3); 2-pentanone (107-87-9); 3-pentanone (96-22-0); 3-methyl-2-butanone (563-80-4); methyl vinyl ketone (78-94-3)

- Carboxylic Acids: methanoic acid (64-18-6); ethanoic acid (64-19-7); propanoic acid (79-09-4); butanoic acid (107-92-6); 2-methylpropanoic acid (79-31-2)

- Esters: methyl methanoate (107-31-3); methyl ethanoate (79-20-9); ethyl methanoate (109-94-4); methyl propanoate (554-12-1); ethyl ethanoate (141-78-6); propyl methanoate (110-74-7); methyl acrylate (96-33-3); vinyl ethanoate (108-05-4); methyl butanoate (623-42-7); ethyl butanoate (105-37-3); butyl methanoate (592-84-7)

- Fluorides: fluoromethane (593-53-3); fluoroethane (353-36-6); 1-fluoropropane (460-13-9); 2-fluoropropane (420-26-8); 1-fluorobutane (2366-52-1); 2-fluorobutane; 2-methyl-1-fluoropropanel; 2-methyl-2-fluoropropane (353-61-7)

- Amides: methanamide (75-12-7); ethanamide (60-35-5); propanamine (79-05-0); butanamide (541-35-5); 2-methylpropanamide (563-83-7)

- Nitros: nitromethane (75-52-5); nitroethane (79-24-3); 1-nitropropane (108-03-2); 2-nitropropane (979-46-9)

- Inorganics: carbon monoxide (630-08-0); carbon dioxide (124-38-9); hydrogen fluoride (7664-39-3); fluorine (7782-41-4); hydrogen (1333-74-0); water (7732-18-5); ammonia (7664-41-7); nitrogen (7727-37-9); nitrous oxide (10024-97-2); neon (7440-01-9); oxygen (7782-44-7); hydrogen cyanide (74-90-8)

Also, rotational conformers for several linear molecules with electronegative atoms within the terminal functional groups are considered. These are included to analyze the effects of electronegative atoms on the properties of other functional groups within the molecule. The description of these molecules is given in the following section.

8.3 Search for Low Energy Conformer

Molecules exist in constant motion. While a molecule in the fluid phase translates along trajectories within a system, internal vibrations and rotations cause the nuclei to travel in varied directions and speeds allowable by the chemical bonds. This internal motion is not only influenced by intramolecular forces, such as bond strengths and steric effects, but they are also influenced significantly by the molecules around them, especially in the liquid phase.

Every single conformation of a molecule cannot be accounted for within a computational chemistry algorithm meant to predict molecular properties. The description of effects that other molecules have on a molecule is available within computational chemistry, but this involves study into solvation models that average these effects. The first filtering criterion when guessing a conformer of interest is that the

geometry of the molecule must occur within a local energy minimum, where perturbations of any of the nuclei locations would result in a higher energy. These local minima are relatively easy to find, since it is a well-studied subject within organic and physical chemistry. This set of geometries offers a finite number of conformations of the isolated molecule from which one can choose a ground-state geometry.

Figure 8.1 depicts the local minima of the molecule 1-propanol. The local minima can be separated into two major groups: the oxygen atom in a *trans* position relative to the terminal carbon atom, and the oxygen atom in a *gauche* position relative to the terminal carbon atom. Each of these groups has within them another three conformers, of which two are distinct in the *trans* conformer (two are identical) and three exist in the *gauche* conformer. The energy for the conformer (calculated at the HF/6-31G

α-trans
E_{min} + 0.42 kJ/mol

β-trans
E_{min} + 1.49 kJ/mol

α-gauche
E_{min}

β-gauche
E_{min} + 1.31 kJ/mol

γ-gauche
E_{min} + 2.45 kJ/mol
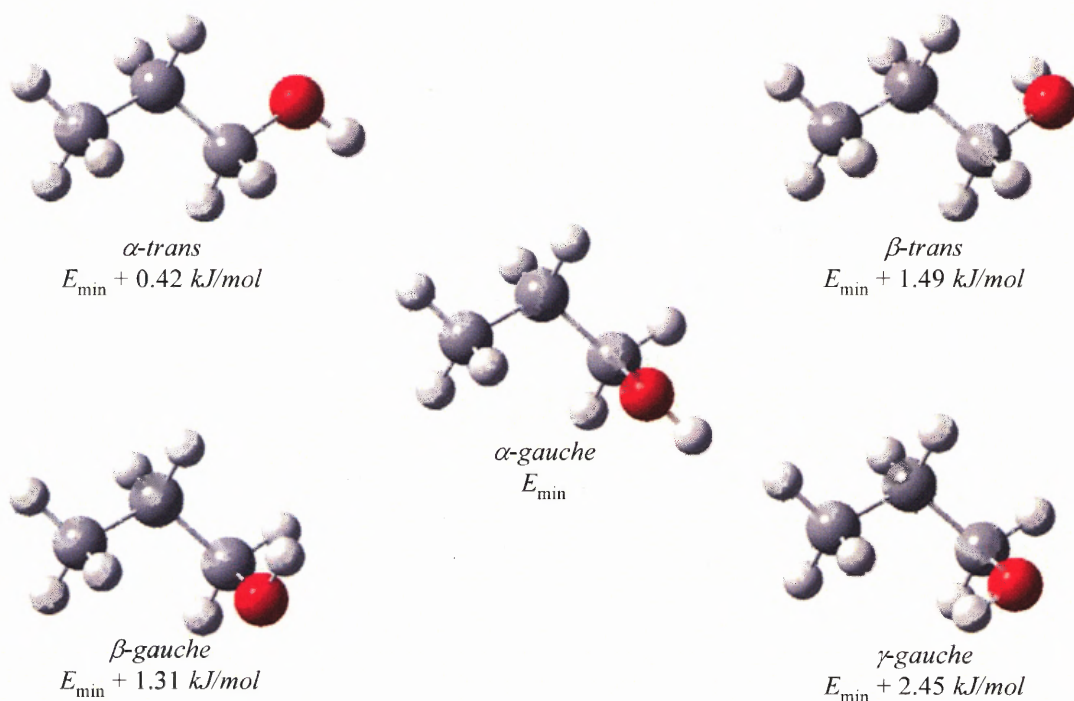
Figure 8.1 Rotational conformers of 1-propanol. All five local minima are shown.

level) is given under the image. The *α-gauche* conformer, where the hydrogen atom is most distant from the remainder of the molecule, is predicted to be the most stable structure. The *α-trans* conformer with the hydrogen most distant from the molecule is found as second most stable. This ordering agrees with experiment, where the *gauche* conformer is found to be more stable than the *trans* conformer by 1.21 ± 0.63 kJ/mol (Abdurahmonov, et al. 1970).

Within linear molecules with four or more heavy atoms, the lowest two conformers found in the conformational analysis are a *gauche* conformer and a *trans* conformer, depicted in general in Figure 8.2. Calculation results that compare the conformer energies are given in Table 8.1.

a)   b)
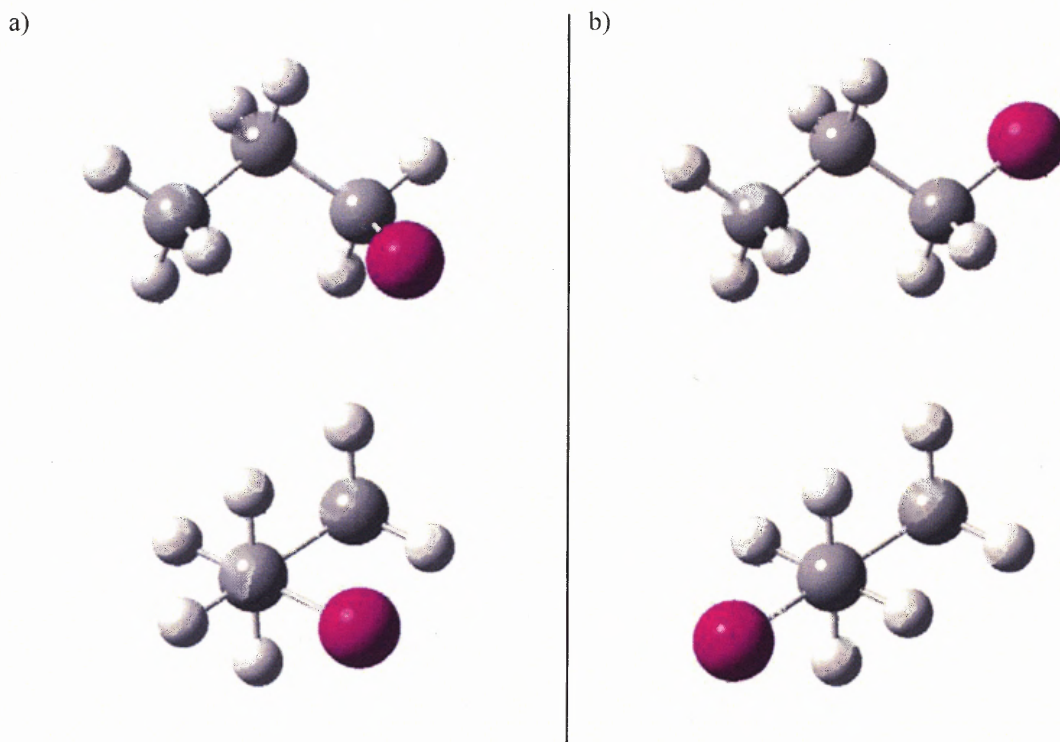
Figure 8.2 Lowest energy *gauche* and *trans* conformers for linear molecules. These are depicted here for a three carbon molecule.

Table 8.1 Absolute Energies of Linear Molecules in Different Conformations

molecule		energy (au)			
		HF 6-31++G**	MP2(full) 6-31++G**	HF AUG-cc-pVDZ	MP2(full)/ AUG-cc-pVDZ
butane	*trans*	**-157.3149**	**-157.9347**	**-157.3151**	**-157.9377**
	gauche	-157.3133	-157.9335	-157.3135	-157.9368
pentane	*trans*	**-196.3529**	**-197.1237**	**-196.3539**	**-197.1285**
	gauche	-196.3511	-197.1226	-196.3522	-197.1277
hexane	*trans*	**-235.3908**	**-236.3127**	**-235.3927**	**-236.3194**
	gauche	-235.3890	-236.3117	-235.3910	-236.3187
1-propanamine	*trans*	**-173.3050**	**-173.9493**	**-173.3112**	-173.9623
	gauche	-173.3045	-173.9492	-173.3105	-173.9623
1-butanamine	*trans*	**-212.3429**	-213.1381	**-212.3499**	-213.1529
	gauche	-212.3423	**-213.1382**	-212.3492	**-213.1530**
1-propanol	*trans*	**-193.1330**	-193.7844	**-193.1443**	-193.8121
	gauche	-193.1329	**-193.7849**	-193.1440	**-193.8124**
1-butanol	*trans*	**-232.1709**	-232.9732	**-232.1830**	-233.0026
	gauche	-232.1707	**-232.9738**	-232.1827	**-233.0031**
1-pentanol	*trans*	**-271.2089**	-272.1623	**-271.2218**	-272.1935
	gauche	-271.2087	**-272.1630**	-271.2215	**-272.1940**
1-fluoropropane	*trans*	-217.1316	-217.7680	-217.1420	-217.8078
	gauche	**-217.1317**	**-217.7685**	**-217.1421**	**-217.8083**
1-fluorobutane	*trans*	-256.1694	-256.9567	-256.1807	-256.9983
	gauche	**-256.1696**	**-256.9573**	**-256.1809**	**-256.9989**

8.4 Calculation of Molecular Polarizability

The molecular polarizability attempts to quantify the response of the electron density to an applied electric field. It is a proportionality factor that describes the linear response of the dipole moment to a small electric field. Greater field strengths induce nonlinear responses, called hyper-polarizabilities. The general formula for the polarizability is given by

$$\Delta\boldsymbol{\mu} = \boldsymbol{\alpha}\cdot\mathbf{E} \tag{8.1}$$

where the dipole vector $\boldsymbol{\mu}$ changes under the influence of the electric field, here expressed as the vector quantity $\mathbf{E}$. The polarizability here is expressed as the tensor $\boldsymbol{\alpha}$, and the elements of the tensor are found through the calculation methods described above.

The polarizability tensor reflects the specific directional responses of the electron density to the electric fields. The density may respond in any of the three perpendicular directions (e.g. the x, y, and z directions) to a field in a particular direction, either the x, y, or z direction. These yield nine responses the electron density may show, therefore the nine elements $\boldsymbol{\alpha}$.

The most significant responses are those in the same direction as the electric field. Therefore one can expect that the diagonal of the polarizability tensor contains the largest contributions. Table 8.2 presents the results of the calculations of the dipole moment vector of 1-propanol in the ground state and with electric fields applied in the x, y, and z directions. The largest changes in the dipole moment vector are in fact in the directions of the electric field. This, in turn, yields polarizabilities an order of magnitude larger on the diagonal than for the other elements. Also note that the polarizability tensor is required to be symmetric. The tensor in Table 8.2 approximately obeys this.

Table 8.2 Response of the Dipole Moment Vector of 1-Propanol to an Electric Field of Strength −0.007 au

μ (Debye)	ground-state	E_x	E_y	E_z
x	0.6311	-0.2334	0.6235	0.624
y	1.0245	1.0164	0.2344	1.0246
z	0.9507	0.9428	0.9507	0.2134
Δμ (Debye) x		-0.8645	-0.0076	-0.0071
Δμ (Debye) y		-0.0081	-0.7901	0.0001
Δμ (Debye) z		-0.0079	0	-0.7373
α (Å³)		7.20	0.06	0.06
		0.07	6.58	0.00
		0.07	0.00	6.14

The polarizabilities of this work are presented as scalar quantities and thus describe the polarizability of the molecule as a spherically symmetric property. To achieve this, the trace of the tensor (the average of the elements of the diagonal) is considered the polarizability of the molecule.

A similar algorithm is followed for the calculation of the polarizabilities of functional groups, presented in Chapter 10. However, there the change in the dipole moment of the group is more detailed, taking into account the charge transfer possible within a molecule influenced by an electric field.

8.5 Results

The results of the calculations and their relation to experimental molecular properties are accomplished in this work by analyzing the difference between the experimental dipole moments and polarizabilities with those predicted by the computational techniques. Comparisons for each individual molecule are listed in Appendix E.

To simplify the analysis between predicted and experimental results, the groupings of molecules are subgrouped by their constituent atoms:

- Hydrocarbons contain Alkanes and Alkenes

- Amines contain molecules with singly-bonded nitrogen atoms

- Alcohols & Ethers contain molecules with singly-bonded oxygen atoms

- Fluorides contain molecules that contain a fluorine atom

- Aldehydes & Ketones contain molecules with a doubly-bonded oxygen atom

- Esters & Acids contain molecules with a carboxylic group ($OC=O$)

- Inorganics contain smaller molecules

- Amides, Nitriles & Nitros contain less-encountered species with a single electronegative atoms or a group of nitrogen and oxygen atoms

The difference between the experimental dipole moments and the calculated results from both the OPT and SPE methods used in this work are presented in Table 8.3. The difference between the calculated polarizabilities from the SPE calculations is considered in Table 8.4. Also within these tables are results from both *ab initio* and DFT methods with various size basis sets. This information is made available from the Computational Chemistry Comparison and Benchmark Database (CCCBDB) online ("Computational Chemistry...", 2004). This selection of methods and basis sets has been made because calculations on a majority of molecules in this work have been completed by the CCCBDB with these methods. The theoretical methods include the DFT methods referenced in Section 7.1, the B3LYP method and the B3PW91 method. Also included for comparison is the MP2 perturbation with correlation only included for the valence electrons, hence the 'fc' (frozen core) notation. The basis sets included from the

CCCBDB is the 6-31++G* set, the larger 6-31G(2df,p) set, and the correlation-consistent sets, AUG-cc-pVDZ and cc-pVTZ sets. Note that the second and the fourth basis sets on this list do not include diffuse functions.

The absolute percent error of calculated dipole moments over the range of methods is presented in Table 8.3. Both the OPT values and the SPE values are represented here, and a large improvement is seen when using the larger basis set. The SPE calculation performs better than those of the DFT methods with smaller basis sets, which is surprising noting that an advantage to DFT methods is that a smaller basis set can be used with comparable accuracy. The method employing the next larger correlation-consistent basis set does incrementally better than the SPE calculation over most of the groupings.

Table 8.3   Absolute Percent Error of Calculated Dipole Moments from Experiment ("Dipole Moments", 2004).

	method and basis set					
	OPT	SPE	MP2(fc)/ cc-pvTZ	B3LYP/ 6-31++G*	B3PW91/ 6-31G(2df,p)	B3LYP/ AUG-cc-pVDZ
Hydrocarbons	13.4%	10.3%	11.1%	17.6%	20.3%	9.8%
Amines	21.6%	7.0%	3.4%	9.9%	8.0%	8.5%
Alcohols & Ethers	17.8%	6.8%	2.1%	8.9%	12.9%	3.5%
Fluorides	22.8%	4.7%	2.7%	12.6%	14.5%	2.8%
Aldehydes & Ketones	25.1%	4.2%	4.8%	10.3%	6.6%	5.3%
Esters & Acids	20.5%	8.1%	3.2%	38.4%	26.3%	37.4%
Amides, Nitriles & Nitros	19.6%	3.8%	2.5%	7.7%	3.0%	13.4%
Inorganics	106.8% (0.357 D)	24.3% (0.036 D)	35.6% (0.165 D)	23.0% (0.078 D)	16.9% (0.083 D)	14.8% (0.041 D)

OPT refers to the MP2(full)/6-31++G** calculations
SPE refers to the MP2(full)/AUG-cc-pVDZ calculations
Results for remaining methods from the CCCBDB ("Computational Chemistry...", 2004)

The DFT method with the same size basis set also performs better than the SPE calculation, although the SPE calculation yields competitive numbers.

A question exists in the performance of the B3LYP/AUG-cc-pvDZ methods for the Esters & Acids grouping. The geometries considered by the CCCBDB may not correspond to those geometries considered in this work.

Also, all methods do poorly in predicting the small molecule dipole moments within the Inorganics grouping. This may be so because the dipole is more influenced by the actual electron density of the system, while the dipole for larger molecules are more influenced by the positions of the more numerous nuclei. Also, the magnitudes of the dipole in this group are smaller, so absolute error may be a better measure on how well the dipoles are predicted. These values are also included within the table, depicting the best results from the SPE method and the DFT method with the large basis set, and reasonable results for the majority of methods.

The absolute percent error of calculated polarizabilities over the range of methods is presented in Table 8.4. One can immediately note the success that basis sets with diffuse functions have over those without. The SPE calculations do a better job than the calculation with the cc-pVTZ basis, mainly due to the lack of the diffuse functions in the latter. This is also seen with the two DFT methods with the 6-31G sets; the calculations with the diffuse functions perform better. Again, when the diffuse basis set is coupled with the DFT method, the numbers are most near that of the experiment.

Table 8.4 Percent Error of Calculated Polarizabilities from Experiment (Miller, 2004)

	SPE	MP2(fc)/ cc-pvTZ	B3LYP/ 6-31++G*	B3PW91/ 6-31G(2df,p)	B3LYP/ AUG-cc-pvDZ
		method and basis set			
Hydrocarbons	4.9%	14.5%	10.8%	18.8%	1.7%
Amines	11.4%	22.2%	18.8%	27.1%	9.1%
Alcohols & Ethers	5.4%	15.5%	12.1%	18.9%	2.7%
Fluorides	16.2%	27.3%	27.3%	29.7%	13.9%
Aldehydes & Ketones	3.6%	10.5%	8.2%	16.2%	1.6%
Esters & Acids	1.0%	17.1%	6.8%	16.4%	2.7%
Amides, Nitriles & Nitros	13.6%	11.4%	15.6%	15.9%	11.2%
Inorganics	10.2%	28.5%	24.6%	30.8%	7.5%

SPE refers to the MP2(full)/AUG-cc-pVDZ calculations
Results for remaining methods from the CCCBDB ("Computational Chemistry…", 2004)

8.6 Conclusions

Computational chemistry offers methods with which one may analyze a large number of molecules without the use of complex and expensive experimental apparatuses. Although no method has ended the search to makes experiments completely obsolete, enough information is available to establish the accuracy and predictive capability of such methods.

The methods utilized in this work allow for a timely prediction of the electrostatic properties of molecules of interest to chemical engineers. The results of these methods can also be improved in a systematic manner as the need arises. The molecular wavefunctions may now be analyzed to determine functional group properties for application within predictive group-contribution methods. The algorithm has been applied to a large number of species, making available molecular-level properties that

may aid in evaluating pure and mixture system properties with techniques described later in the work.

All this information comes from modest computational resources. With the application of more powerful tools, the results in theory will approach those of experimental results. Due to the greater acceptance of such computational techniques, such solution methodologies will become regularly used tools for chemical engineers willing to approach problems from a more fundamental standpoint.

CHAPTER 9

GROUP-CONTRIBUTION METHODS AND
THE CONCEPT OF THE FUNCTIONAL GROUP

Thousands of chemicals are of interest to engineers, and more are added every year. It is impossible to expect that all the relevant properties of these molecules and their mixtures can be readily available to engineers for use in process design. Group-contribution methods (GCMs) have been developed to aid engineers approximate molecular properties through a smaller set of entities called functional groups.

Functional groups are the atoms and collections of atoms that are used to approximate the properties for molecules of interest. The problem of finding all the properties of molecules and their mixtures is reduced to finding the definitions of functional groups, the functional group properties and the models used to assemble molecular properties.

The transferability concept alluded to in Section 5.5 has made the finding of universal and well-defined properties for a group a goal in the advancement of GCMs. The problem is, of course, that the properties of a given group of atoms are affected by what other groups constitute both the molecule and system. This means that the group is almost never transferable (Fresendelund and Prausnitz, 1975; Bader and Becker, 1988; Sandler, 1994), and thus GCMs that assume transferability will have a limited scope of success.

A related problem with current GCMs involves isomers (Fresendelund, et al. 1975). Molecules that consist of geometric rearrangements of the same set of groups are indistinguishable by current models that rely on the assumption of transferability.

133

Atoms in Molecules (AIM) theory (Bader, 1990) offers a rigorous method to find properties of functional groups in molecular environments. Such information is calculable once the numerical approximation to the wavefunction is found using methods described in Chapter 6. Two main bodies of work on this subject exist: by the Bader group, on the theoretical basis of AIM and initial studies of alkanes (Bader, et al. 1987; Bader and Becker, 1988; Bader, 1990; Bader et at, 1992; Bader and Bayles, 2000; Cortés-Guzmán and Bader, 2003); and another by the Mosquera group, on the calculation of properties for a wide range of molecule types (Graña and Mosquera, 1999; Carballo and Mosquera, 2000; Graña and Mosquera, 2000; Vila and Mosquera, 2001; Lorenzo and Mosquera, 2002; Mandado, et al. 2002; Mandado, et al. 2003; Quiñónez, et al. 2003). Since the number of references is numerous, the general body of work from each group within this chapter will not be rewritten, except when a particular piece of work is referenced due to the focus of that work.

This chapter offers an overview of GCMs and the tasks in determining definitions and contributions to the properties of a molecule. Focus is then placed on the role that AIM theory has in alleviating the past problems with definitions and properties. Several numerical routines exist to determine AIM properties; these are considered. Finally, an argument against the full transferability of a functional group definition over molecules is discussed, but on a quantitative scale offered by AIM theory.

9.1 Group-Contribution Methods for Thermophysical Properties

Thermophysical property prediction continues to be a significant research area. The importance of readily available and accurate properties of known and yet unknown species is evident by their use in powerful engineering process simulations. Properties such as heats of formation, boiling points, and heat capacities for all participating species need to be readily available for a robust simulation. Also, costly, time consuming experiments on rare compounds are avoidable if a predictive model of sufficient accuracy can be used within these simulations.

Group-contribution methods are regularly used to predict thermophysical properties. Methods exist that make this information available simply with knowledge of the numbers and types of groups occurring in a molecule. Each functional group contributes a value to the property of interest. In the past, these contributions have been correlated to a large set of data for the property of interest. For simpler schemes, a contribution from a particular group is assumed to be the transferable, the same in whichever molecule the group exists.

To calculate macroscopic properties, GCMs assume that each contribution from a group can be summed linearly to give the property of interest. For a given thermophysical property F_i, the contributions of the functional groups f_j are added to give the result

$$F_i = \sum_j v_{i,j} f_j \tag{9.1}$$

where $v_{i,j}$ is the number of j functional groups in molecule i. Each of the functional group properties is assumed to be independent of the contributions by other groups in the

molecule. Also, these properties are assumed to be transferable, thereby usable for the functional group within any molecular environment.

Recently, this subject matter has been studied within the framework of quantum mechanics and AIM theory. Linear additivity, demonstrated by Equation (9.1), has been shown to exist for structural properties of alkanes (Bader et al. 1987), the polarizability of linear alkanes (Bader et al. 1992), and the HF energies of linear alkanes and oligosilanes (Bader and Bayles, 2000).

Group-contribution methods widely used by engineers have had success assuming linear additivity without a rigorous theoretical framework. A successful additivity scheme (Benson, 1976) establishes a large number of functional groups and their associated properties with which to construct many molecules. The properties made available allow for the calculation of common properties necessary for process design, such as heats of formation, entropies, and heat capacities. The application of this method is rather straightforward, since one only needs to add one contribution from each group.

The complexity of GCMs has grown recently in order to predict larger molecules not previously included in correlation schemes. These methods have multiple contributions from groups of differing scales, usually in a scheme resembling a perturbation expansion

$$F_i = \sum_j v_{i,j}^{[1]} f_j^{[1]} + \lambda \sum_k v_{i,k}^{[2]} f_k^{[2]} + \chi \sum_l v_{i,l}^{[3]} f_l^{[3]} + \dots \qquad (9.2)$$

where the superscripts denote the order of the functional group contribution (first-order, second-order, etc.), and λ and χ are perturbation parameters that assume the values of 0 or 1. A first-order approximation can be applied for a quick estimation of properties. For better accuracy, the higher-order contributions are added to finely adjust the first

result. A scheme has been proposed where a second-order contribution is based on the numbers and types of conjugate structures possible within a molecule (Constantinou and Gani, 1994). These second-order contributions were based, in part, on the enthalpy of reaction of the bond conjugation. Another scheme proposes first-order group contributions on par with the simplicity of Benson, and second-order contributions given by larger groups containing two adjacent, first-order groups (Marrero-Morejon and Pardillo-Fontdevila, 1999). These contributions are then to be summed in global equations containing two correlated parameters. This concept is built on by adding as third-order contribution groups the size of aromatic rings and larger (Marrero and Gani, 2001). This method is geared toward including the prediction of large, polycyclic molecules.

Each of these schemes employs definitions of functional groups that are not necessarily similar to that of prior work. These definitions have been chosen arbitrarily and must be reconsidered when developing or studying any new method.

9.2 Research into Functional-Group Definitions

For a palatable GCM, the definition of a functional group must balance between ease of use and complexity of description. A functional group with fewer atoms/bonds will satisfy the desire for a simple method, but will sacrifice the accuracy necessary in a robust model. A simple functional group will occur in a large variety of molecules, each likely with a different surrounding environment of atoms and electron densities. A group reflects these differences with changes in their electron density profile, and thus, a change in their functional group properties. A method that applies large functional group

definitions serves the purpose of accounting for these environmental effects. Yet these groups tend to occur in fewer molecules, therefore the database of functional groups must be very large to include more unique groups. These larger definitions may also be confusing because molecules are able to be built using different functional groups in this larger database.

The Benson additivity scheme employs a simple group definition that includes a central, polyvalent atom (an atom bonded to two or more other atoms) and its ligands. The notation in that work considers the central atom as a first-order definition and defines further orders with the ligands. For example, ethanamine is defined using four groups stated in the following way: one $C\text{-}(H)_3(C)$, one $C\text{-}(H)_2(C)_2$, one $C\text{-}(H)_2(C)(N)$, and one $N\text{-}(H)_2(C)$. An atom can exist in several group definitions, but only as ligands to other central atoms. Larger cyclic groups, composed of combinations of smaller groups, are defined to allow for small corrections to the values found using the polyvalent atoms. This group definition scheme is likely the simplest that employs information beyond only atoms and bonds.

The work of Marrero and Gani (2001) employs a range of functional group definitions, from simple to complex, to encompass a larger array of molecules. The groups are in general more specific and do not follow as simple a definition as with the Benson method, thereby eliminating the uniqueness of a group definition. For example, it seems plausible to build ethanamine in several different ways under the Marrero and Gani definitions. One way is similar to the definition above: one CH_3, two CH_2, and one NH_2. Another way, which is advised by the authors, is the following: one CH_3, one CH_2, and one CH_2NH_2. When considering the second-order corrections to the result

from the above groups, the second-order groups are allowed to overlap. Although the second-order groups $CH_3CH_2CH_2$ and $CH_2CH_2NH_2$ do not exist in the method, if they did, these would both serve as necessary second-order corrections to the first-order result. Second-order group definitions are allowed to not overlap, partially overlap (as above), and completely overlap (where one definition is a subset of atoms within a larger definition). Third-order group definitions involve rings to incorporate polycyclic compounds into the predictive method and are not related to this work.

Although UNIFAC (Fredenslund, et al. 1975) and modified UNIFAC (Gmehling, et al. 1993) methods do not attempt to predict thermophysical properties, these methods employ a definition as complex as that of Marrero and Gani. Since these methods attempt to correlate VLE data, interaction energies between groups are important. The group definition includes a main grouping and a subgrouping to eliminate fitting parameters to every element of the interaction matrix. Interactions between groups within the same main group vanish. Like the definitions proposed by Marrero and Gani, a similar complication pertaining to the non-unique breakdown of a molecule into functional groups arises.

Attempts have been made to incorporate computational chemistry methods to alleviate some of the concerns of group definitions within the UNIFAC methods (Wu and Sandler, 1991a; Lin and Sandler, 2000), as alluded to in Section 5.5. But these improvements do not eliminate the non-unique definitions for molecules.

Functional groups are incorporated within molecular dynamics simulations to allow for more efficient computations by considering fewer entities within the system space. A functional group definition scheme that has been successful in the correlation

and prediction of alkane system properties (Jorgensen, et al. 1984; Martin and Siepmann, 1998) is called the United-Atom (UA) definition. The definitions for alkanes, alkenes and cyclic hydrocarbons include bonded hydrogen atoms onto the heavy carbon atom, similar to Benson definitions. Further works with alcohols (Chen, et al. 2001) and other oxygen-containing compounds (Stubbs, et al. 2004) include ligands within the group definition, thus returning to the Benson definition of functional groups. Also of note in the UA definitions, hydrogen atoms are separated from oxygen atoms, thus exposing the large partial charge on the hydrogen atom and allowing for the simulated system to represent hydrogen bonding.

The more rigorous routines that find functional group properties, those employing a theoretical framework such as AIM theory, utilize functional group definitions near the simplicity to the Benson and UA schemes: a heavy atom and the attached hydrogen atoms. This is possible since the molecular environment affecting a particular functional group is reflected in the rigorously determined functional group properties. Therefore, subgroupings of the functional group CH_2 are seen purely through the partial charge the group has when attached to electronegative atoms F, O, N, etc...

9.3 Functional Group Properties

In the past, engineers have used correlation to find functional group properties. They would assemble a database of a large number of experiments that measure the system property of interest for a wide array of molecules. Then the database is separated into a training set, used mainly to find the functional group properties, and a validation set, which is used to compare the predicted results of the correlation to the experimental data.

Engineers at this point hope the functional properties can be used in a model that predicts experimental behavior for molecules where data is not available. This makes the finding of universal functional group properties valuable.

The problem of estimating molecular properties reduces to finding the contributions of the functional groups to the molecular properties in relations such as Equations (9.1) and (9.2). The properties of interest to engineers and chemists include the size and shape properties (structural), the contribution to molecular energies (energetics), and partial charges and dipole moments for use in interaction energy schemes (electrostatics). These properties can be found in several different ways: full correlation of experiment on macroscopic systems; correlation of computational chemical properties; approximation from theoretical concepts; or fully rigorous theoretical results. The methods employing AIM theory have not been developed for use in engineering models as have the others, and special focus is placed on these due to their application within this work.

Structural group properties reflect the amount of excluded space a group occupies and the surface area available for interactions with other functional groups. Such properties are important within the lattice-fluid models which account for system entropy of randomly and nonrandomly mixed systems, as described in Chapters 4 and 5. A simple set of formulas exist that approximate the volume and surface area using overlapping spheres with radii that correspond to the van der Waals radii (Bondi, 1964). The UNIQUAC method (Abrams and Prausnitz, 1975), an empirical model, relates the size and area parameters of a molecule, r_i and q_i, to the Bondi size and surface area of a

methylene group in an infinitely long polymer composed of methylene groups. The parameters for a general molecule are found using the formulas

$$r_i = \frac{V_i}{V_{CH_2,\infty}} \text{ and } q_i = \frac{A_i}{A_{CH_2,\infty}} \tag{9.3}$$

where the volume and area for the methylene group in the infinite polymer, $V_{CH_2,\infty}$ and $A_{CH_2,\infty}$, are found to be

$$V_{CH_2,\infty} = 15.17 \text{ cm}^3/\text{mol and } A_{CH_2,\infty} = 2.9 \times 10^9 \text{ cm}^2/\text{mol} \tag{9.4}$$

The COSMO-based methods (Klamt, 1995; Klamt, et al. 1998; Lin and Sandler, 2002; Klamt, et al. 2002) utilize approximately 120% of the van der Waals radius to construct the surface surrounding a molecule in a continuum solvent, although these radii have been correlated in the more recent works. Since functional group volumes are not actually needed within lattice-fluid models (the molecular volume is needed), the functional group volumes may be inferred from liquid volumes (Knox, 1987).

Models that use overlapping spheres formed from van der Waals radii have encountered problems at the intersections of those spheres, where crevices form and create errors in calculations of properties on these surfaces (Klamt, et al. 1998). Fixes have to be implemented to smooth out the crevices.

A concise way to determine structural properties of functional groups without approximations or fixes is achievable using AIM theory. Calculated molecular volumes of alkanes are shown to be additive in the sense of Equation (9.1) (Bader, et al. 1987). This result agrees with the well-accepted notion that system volumes of linear alkanes beyond hexane increase by a constant factor. The surfaces partitioning molecules and an arbitrarily chosen isodensity surface, usually the 0.001 au isodensity surface, enclose a

functional group volume. The area of this isodensity surface also corresponds conceptually to the surface area utilized in lattice-fluid models: the exposed surface area available for interactions with other functional groups. Novel calculations of these surface areas are presented in Chapter 10.

Energetic properties of a functional group reflect the contribution made to the molecule's absolute energy (found in computational chemistry) and heat of formation (mainly found through experiment). The engineering GCMs correlate these functional group properties to experimental data. AIM theory allows for the calculation of the energies of atoms, which are thereby added to get the energies of functional groups. These properties are found to be additive in the sense of Equation (9.1) (Bader and Becker, 1988). A large database of energy values for functional groups within hydrocarbons and oxygen-containing compounds has been made available by the Mosquera group (Graña and Mosquera, 1999; Graña and Mosquera, 2000; Carballo and Mosquera, 2000; Vila and Mosquera, 2001; Lorenzo and Mosquera, 2002; Mandado, et al. 2002; Mandado, et al. 2003; Quiñónez, et al. 2003).

Electrostatic properties, such as partial charges and dipole moments, originally have served as higher-order descriptors that distinguish similar functional group definitions between molecules (Wu and Sandler, 1991a). A large number of methods have been developed to determine the distribution of partial charges amongst atoms within a molecule, mainly due to the importance of such information within molecular dynamics simulations. However, because partial charges are not measurable quantities within the quantum mechanical sense (measurable meaning they can be calculated by an

integral such as Equation (6.35)), partial charges do not exist and must serve as approximations of the electron density profile (Cox and Williams, 1981).

Partial charges have been determined for molecules in numerous ways. An approximation is found by enumerating electrons within orbitals wholly attributable to an atom and equally splitting electrons shared by atoms (Mulliken, 1955). This yields a number reflecting the topological occupation of electrons within the available orbitals. Empirical charge fitting schemes grew out of a motivation to fix some well-documented problems of Mulliken charges: charges are not measurable; the shared electrons are arbitrarily distributed evenly to the nuclei; and in general the Mulliken charges do not appeal to trends that are chemically realistic (Cox and Williams, 1981). A solution is to assign charges to atoms within a molecule by empirically fitting them to the electrostatic potential around the molecule, a measurable quantity with quantum chemistry (Momany, 1978). Since interactions at relatively high energies occur outside the space where electron densities are high, points in this region can be sampled to fit charges on atoms to reflect intermolecular interaction tendencies. Several charge fitting procedures exist within the Gaussian 98W package by evoking the 'pop' keyword: Merz-Kollman (MK) (Besler, et al. 1990); Charges from Electrostatic Potentials (CHelp) (Chirlian and Francl, 1987); and Charges from Electrostatic Potentials using a Grid (CHelpG) (Breneman and Wiberg, 1990). Each of these routines samples the electrostatic potential around the molecule and fits charges accordingly.

While charges fit to the electrostatic potential offer a computationally efficient way to determine partial negative and partial positive fragments of a molecule, charges on nuclei alone cannot fully describe the complex functionality of the electrostatic

potential. Comparisons have been made between the major charge fitting procedures available with quantum chemical calculation software (Sigfreddson and Ryde, 1998). All the well-known charge fitting procedures use a grid of points to sample the electrostatic potential. The procedures available in the G98W software package have been found to be dependent on the orientation of the molecule in the calculations. Since the grid is predetermined within the software, the orientation of the molecule dictates where these points fall. The minimum distance from nuclei at which points are sampled also is an arbitrary choice. The authors suggest that newer methods should employ the higher electrostatic moments for the charge fit or weightings of sampled points.

Another method of determining partial charges on functional groups is to select them according to results within molecular simulations (Martin and Siepmann, 1998; Chen, et al. 2001; Stubbs, et al. 2004). They are attributed to the centers of the UAs, and serve to simplify the Coulombic interactions between polar portions of molecules. According to J. J. Potoff (personal communication, November 7, 2004), the selection of charges at this point is more of an art.

Atoms in Molecules theory, through the partitioning of the electron density of the molecule, allows for the calculation of a partial charge of a functional group, as well as other electrostatic properties such as the dipole moment, quadrupole tensor and the polarizability (Bader, 1990; Bader, et al. 1992). The partial charge of a functional group is given by integrating the electron density within the boundaries of the group found through AIM theory

$$q_\Omega = Z_\Omega - N_\Omega = Z_\Omega - \int \rho \, d\tau_\Omega \qquad (9.5)$$

The dipole moment vector is found by finding the first moment of the electron density

$$\mu_\Omega = -\int \mathbf{r}_\Omega \rho d\tau_\Omega \tag{9.6}$$

where $\mathbf{r}_\Omega$ is the distance with respect to the nucleus. The quadrupole tensor is also given by integration with the appropriate operator, but this quantity is not utilized within this work.

The calculation to determine the polarizability is a somewhat more involved calculation, due to the necessity of molecular calculations in the ground-state configuration as well as on configurations with electric fields being applied. The polarizability of a molecule or a functional group is given by the linear response of the electron density to an electric field. This is expressible through the change in the dipole moment given by Equation (8.1). The change in the dipole moment for an entire molecule can be separated into atomic or functional group quantities (Bader, et al. 1992), which are then used to determine the polarizability for functional groups within Equation (8.1). This relation is give by

$$\Delta\mu = \sum_\Omega \left(\Delta q_\Omega \mathbf{R}_\Omega + \Delta\mu_\Omega \right) \tag{9.7}$$

where $\mathbf{R}_\Omega$ represents the location of the center of atom or functional group Ω.

Determining the latter term within Equation (9.7) involves calculating the dipole moments of a functional group with Equation (9.6) and determining the change in this value after applying electric fields in the x, y, and z directions.

Determining the former contribution within Equation (9.7) involves determining the shift of electrons within a functional group and determining the shift of electrons from the group of interest to and from the remainder of the molecule. For a functional group

with attached hydrogen atoms, such as CH_3, CH_2, and CH, the formula for contribution

to the change in $\Delta\mu$ is given by (Bader, et al. 1992)

$$\Delta q_{CH_n} \mathbf{R}_{CH_n} = \sum_{i=1}^{n_H} \Delta q_{H_i} \left(\mathbf{R}_{H_i} - \mathbf{R}_C \right) + \sum_{i=1}^{n_{BCP}} \Delta q_{R_i} \left(\mathbf{R}_i - \mathbf{R}_C \right) \tag{9.8}$$

where n_H represents the number of hydrogen atoms within the group, n_{BCP} represents the

number of bond critical points the functional group has with other functional groups

within the molecule, $\mathbf{R}_i$ is the location of bond critical point i, and Δq_{R_i} represents the

change in the partial charge of the molecule beyond the bond critical point i. For

functional groups with no attached hydrogen atoms, or for functional group computations

where the bounding surfaces between the carbon atom and the hydrogen atoms are

ignored, only the second term of Equation (9.8) is necessary. The effects of the

intragroup charge transfer in this latter case are quantified in the change of the dipole

moment of the group, $\Delta\mu_\Omega$.

The polarizabilities of atoms and functional groups given by Equations (8.1) and

(9.7) are found to be linearly additive in the sense of Equation (9.1). When computed at

the HF level of theory with correlation corrections, the sum of the group polarizabilities

within linear alkanes are found to be exactly the same as the calculated polarizabilities for

the entire molecule and are about 11% less than experimental values for the molecule

(Bader, et al. 1992).

9.4 The Quantitative Reasoning Against Transferability

Transferability of a functional group over a large number of molecules is considered by most an idealization that has proven difficult to achieve. The properties of a functional group depend on many things, such as group definition, group boundaries and surrounding groups within a molecule. The smaller group definitions, the chances of transferability become less because of the more immediate effect different environments have on its properties. The larger the group definitions, the chances of achieving transferability are better, but mainly because the functional group would occur in fewer molecules. This becomes less useful and complicates the GCMs employing such definitions.

The application of AIM theory to molecular studies has furnished proof of the slight chance of transferability in a GCM. Work on defining the AIM properties of carbon atoms within alkyl groups in aldehydes and ketones (Graña and Mosquera, 2000), in monoethers (Vila and Mosquera, 2001) and in linear alcohols (Mondado, et al. 2002) shows the sensitivity of alkyl groups, the most common functional groups in organic compounds, to an electronegative atom like the oxygen atom.

The number of transferable group definitions within the studies above suggests that a chemist's viewpoint of transferability is extremely sensitive and demands a large number of subgroupings to be of use. Table 9.1 shows the number of subgroupings within the molecules mentioned above. First note that the number of different carbon atoms within these groups of molecules number 47. The main deciding factor within separating functional group definitions is found to be the distance from the electronegative atom. The range of partial charges within the ethers and alcohols is due

Table 9.1 Analysis of Transferability of Atoms within Alkyl Groups (Graña and Mosquera, 2000; Vila and Mosquera, 2001; Mondado, et al. 2002).

molecules	number of molecules considered	transferable atom	number of distinct subgroupings	range in partial charge
aldehydes and ketones	12	H bonded to CO	3	0.020
	42	C	21	0.100
monoethers	33	methyl C	5	0.625
		methylene C	11	0.589
linear alkanes and alcohols	19	C	10	0.576

to the difference between a carbon atom bonded directly to an oxygen atom, which makes the partial charge on carbon become largely negative, and a carbon atom far away from the electronegative atom, which makes the partial charge only slightly negative.

Such a demonstration on the way AIM charges are affected by electronegative atoms within relatively simple molecules eliminates the validity of the transferability assumption from any engineering model that uses AIM properties. There is a possibility that transferability exists approximately, as shown in alkyl groups in alkanes (Cortés-Guzmán and Bader, 2003), or shown by the groups in linear alcohols approaching the properties of the alkyl groups in alkanes (Mandado, et al. 2002). Such conclusions are used in limited cases within the AIM properties of functional groups calculated for this work. These are presented in the next chapter.

9.5 AIM Property Calculators

Three software packages are available to calculate AIM properties: PROAIMV (Biegler-König, et al. 1982); a subroutine within Gaussian 98W evoked by the 'aim' keyword (Stefanov and Cioslowski, 1995); and AIM2000 (Biegler-König, et al. 2000). The most

computationally intensive portion of an AIM calculation is determining the partitioning surfaces. The first two routines do this, albeit in different manners, while the third, AIM2000, attempts to proceed integrating without such information. The user interfaces, the necessary user input and the routine outputs are described.

The earliest and probably most widely used software is the PROAIMV routine available within the AIMPAC suite. This suite is made available by the Bader research group. The partitioning surface is found by walking along gradient paths that begin at a bond critical point. An array of points on the surface is determined, and triangles are used to connect the points and represent the surface. The program is run on a PC from a DOS prompt after compiling in a FORTRAN 77 compiler. The user input for this routine is the most involved of all the routines; the user needs the .wfn file made available within G98W and the location of the bond critical points. Also necessary for calculation are the numbers of bond, cage and critical points (the latter two not discussed in this work), the number of planes and radial points used for the three dimensional integration within the atomic space, and an array of values that allow for customized calculations. The output of the routine is a formatted text file with all the calculable properties within AIM theory, except the external surface area and polarizability. Since the interface is a DOS prompt, it is possible to write scripts that allow for multiple integrations to be run without user intervention. Also, if one or more bond critical points are omitted from the input file, the routine will integrate as though the surface is not there. This is helpful for functional group property calculation, where the user may omit the bond critical points between the hydrogen atoms and the carbon atom in a methyl group to achieve a single integration of the alkyl group properties.

A method that attempts finding of the partitioning surface more elegantly is made available within G98W by using the 'aim' keyword (Stefanov and Cioslowski, 1995). An analytic function is used to fit the partitioning surfaces between two atoms. This program is easily executed from a G98W input file and needs no further input from the user. The AIM properties for all the atoms are calculated when the routine is executed. The output of this routine is a section within the formatted output regularly available after a G98W calculation. The properties calculated include the partial charges, the dipole moments relative to the attractor in the basin (the maximum of the electron density), the quadrupole moment, and a variety of atomic energies and forces. An example of the output for this routine is given in Appendix C.

A method that attempts to avoid calculating the partitioning surface altogether is with the free software called AIM2000, Version 1.0. This package numerically integrates by finding the points within the atomic space using a technique within differential equations. Each of these points is found by beginning at the attractor within an atomic basin and walking downhill along a gradient path. By theory, all these points on the gradient path belong to the attractor on which the path begins. The interface of this program is a Windows-based GUI with a three dimensional representation of the molecule, the bonds, and the bond critical points. The inputs from the user involve mostly integration parameters and are set within menus available in the software. Calculation routines with AIM2000 are utilized to determine the attractor centers and bond critical points. The output of the calculations includes all the information available from PROAIMV, but within a table that is inaccessible to outside software. No text-based output is available with AIM2000.

Two different integration techniques are made available within AIM2000. The first choice is similar to the PROAIMV, where a beta sphere around the atom center is formed. A two step integration is then performed: one inside the beta sphere and one outside the beta sphere. The second choice is called a radial integration, which is a one part integration of the entire atomic space using the atom center as the starting point.

The coarseness of the partitioning surface finder within PROAIMV motivates the creation of the routine that finds the analytical representation of the surface (Stefanov and Cioslowski, 1995). Although elegant, this routine is prone to crashing when attempting to find surfaces with sharp curvatures (Biegler-König, 2000). Since the AIM properties of all the molecules are found together using the 'aim' keyword, if one surface fails in the process, the entire calculation fails and no output of AIM properties for any atom is offered. The AIM2000 routine is also elegant, but its methods have been criticized as numerically expensive and too time consuming (Stefanov and Cioslowski, 1995).

Due to the nature of numerical calculation, the values determined by the software are subject to some variability. A measure of the numerical error by the AIM integrators, $L(\Omega)$, is proportional to the Laplacian of the electron density integrated within the volume. If the properties of the atoms and the partitioning surfaces are determined exactly, this quantity theoretically vanishes for each atom (Bader, 1990; Cortés-Guzmán, and Bader, 2003). To achieve an error of less than $0.0022e$ in the number of electrons integrated within a basin, the value of $L(\Omega)$ is to be less than 1.0×10^{-3} au and 1.0×10^{-4} au for a hydrogen atom and carbon atom, respectively (Bader, et al. 1992).

When calculating the magnitude of the dipole moment vector, Gaussian 98W and AIM2000 use different conventions than the PROAIMV routine. In AIM2000, the dipole

moment needs to be multiplied by a factor of negative one to alleviate the difference in sign convention. In Gaussian 98W, the centers of the attractors, maxima of the electron density at or near the location of the nuclei, are the centers of positive charge used when determining the dipole moments.

9.6 Summary

Group-contribution methods have a successful record at modeling many thermodynamic properties of fluid systems. However, with so many models in existence, there are as many definitions for functional groups as there are methods. New definitions have been created to best correlate the data for that particular model.

The emergence of computational power and a theoretical method, AIM theory, now may eliminate the arbitrary practice of defining functional groups. Instead of striving for transferable group properties, deemed impossible in its classical conceptualization, group definitions can be given less attention as long as the accompanying properties are defined. It is these functional group properties that are physically significant and can be related directly to macroscopic system behavior. Information on the structural, electrostatic and energetic properties of functional groups had previously been unavailable, and new techniques to describe thermochemistry and thermodynamics can utilize this information for more fundamentally sound GCMs. With the availability of software to calculate AIM properties, such information can be disseminated as more molecular calculations are performed.

This work progresses along this path by taking the molecular calculation results and analyzing them using AIM theory. This information becomes the basis for newly conceived physical parameters within interaction and statistical models.

CHAPTER 10

FUNCTIONAL GROUP PROPERTIES USING AIM THEORY

The statistical models that depict macroscopic system thermodynamics have a number of parameters that are related to the structure and energetic behavior at the molecular level. In the past, engineers used approximations or correlation techniques to determine these properties. These correlations, especially those within quasi-chemical models, yield parameters with physical significance.

This work uses the computational results from Chapter 8 to determine functional group properties using Atoms in Molecules (AIM) theory. To simplify the often confusing range of definitions of functional groups, a straightforward definition is used within this work. The transferability assumption will not be assumed, eliminating the need for complicated rules-based methods that have governed past definitions. This also allows for unique definitions of molecules. The AIM properties serve as the higher-order distinguishing characteristics, thus eliminating the problem with isomers.

The algorithm to calculate functional group properties results from computational chemistry methods is outlined. The calculation method determines the exposed surface area of a group, an important property within quasi-chemical group-contribution methods (GCMs), is presented. Tables of AIM properties for molecules of interest in this work are also presented, and these are referenced back to macroscopic properties in the attempt to validate the properties. Since most of the properties are additive, the comparison of the functional group property results to experimental data are found within the comparisons to experimental data in Chapter 8 and Appendix E.

10.1 Working Definition of a Functional Group

The wealth of information available within the AIM property calculations allows for very simple definitions of functional groups. The scheme employed in this work is motivated by the success of the United-Atom (UA) schemes in molecular dynamics and by the ease of application of the Benson method for thermophysical properties.

The molecules in this work contain functional groups that are classified in one of eight definitions: CH_3; CH_2; CH; C; H; N; O; and F. The only qualifier in this list is that the H functional group must be attached to a non-carbon heavy atom, namely an N, O or F. This reasoning follows that of the UA approaches, where the H atom has a relatively large positive charge ($q > 0.1$) as compared to the H atoms within the alkyl functional groups ($q < 0.1$).

The simplicity of this scheme becomes apparent when one considers the suggestion by past GCMs that employ quantum mechanics: to attach alkyl groups or H atoms to adjacent groups that contain electronegative atoms. If this suggestion were employed for this work, the list of groups would be very large. For a start, the following basic functional group definitions make the list: CH_3; CH_2; CH; C; NH_2; NH; OH; O; and F. Attaching the adjacent alkyl group to the electronegative group would add 24 more definitions (take one from the first four groups and combine with one from the last six groups). This would bring the total up to 34 definitions. If one were to add other combinations, such as the NO_2 group for nitro-containing molecules and COOH for carboxylic acids, as well as other qualifiers such as doubly-bonded atoms and aromatic groups, the number grows such that there is nearly an order of magnitude more definitions than the original eight proposed for this work.

To distinguish the myriad environments within which the eight functional groups can exist, the AIM properties for the functional group serve as descriptors. These properties include one energetic parameter (the energy of the group E), three structural parameters (the volume V, the exposed surface area A, and the average distance r_{avg} to the 0.001 au isodensity surface), and three electrostatic parameters (the partial charge q, the magnitude and direction of the dipole moment μ, and the polarizability α). Such rigorously calculable functional group properties have not been used within engineering group-contribution methods in the past, as they are in this work.

10.2 The Lack of Transferability

If the transferability assumption were to hold for such group definitions as above, then one must expect that all the information available for the transferable groups must be identical. This includes not only the topological definition, but also the seven properties in the AIM analysis. Past experience has shown this to be unachievable. The view that transferability is unattainable is absolutely acceptable now because of quantitative evidence presented in Section 9.4 and the results of this work.

As noted in Section 9.4, there are a few instances where transferability strictly exists. From an engineer's point of view, many of the differences that a chemist may cite are too small to eliminate transferability as a first approximation. Such schemes that employ a looser definition of transferability can be created, albeit with arbitrary bounds of what is considered transferable.

For the majority of this work, the transferability assumption is not made. There are a few instances when, attempting to calculate the AIM properties of functional groups,

problems within the computational routine arise and give inaccurate results. Past experience is used to determine these functional group properties, specifically through the consideration of instances of transferability from prior calculated results. Details of this are explained in a later section. But for the majority of the cases, the computational power and availability of AIM properties allows one to disregard the transferability assumption, which was made to simplify and make more efficient the estimation of thermophysical properties when such computational power did not exist.

10.3 Calculation of Group Properties

The algorithm for the calculation of functional group properties is an extension to the algorithm in Section 8.1 that obtains the molecular wavefunction from *ab initio* methods. Integrations for a particular functional group must be run four times: once for the ground-state molecular wavefunctions, and once each for the wavefunction under the influence of the three electric fields. To perform an AIM integration using the PROAIMV routine, the critical points within the G98W output file and the .wfn files are needed. Also needed are input files and the executable file that runs the PROAIMV code.

The algorithm to achieve the seven functional group properties follows. Example input files and executable calls referenced within each step are presented in Appendix F.

- The PROAIMV routine is called to determine the energies, volumes, partial charges and dipole moments for each of the functional groups within the ground state of the molecule. The critical points from the G98W output file are used within the input files of PROAIMV, and the ground-state wavefunction is utilized in the routine. Integration options must be set within the PROAIMV input file.

- The PROAIMV routine is called to determine the partial charges and dipole moments for each group within the molecule under the influence of the three electric fields. The bond critical points have to be rewritten into new input files, since they are different for the molecule under the electric field. The wavefunctions from the molecular calculations with the applied electric fields are utilized in this step. To balance the computational time, the integration parameters for this step are changed so to offer a slightly rougher integration.

- A series of routines has been created in this work to determine the exposed surface area, the average distance to the exposed surface area, and the volume of the group using a Monte Carlo calculation. An input file similar to that of the PROAIMV routines is necessary here, with a few added values to set more numerical parameters. The critical points of the functional groups are also necessary for this step, as well as the ground-state wavefunction of the molecule. More details on the routines that determine structural parameters are offered in the next section.

- The results of the calculations on all the functional groups are tabulated. The energy, integrated volume, partial charge and dipole moment are found directly from the integrations of the ground-state wavefunction. The polarizability must be calculated from the integration results for the ground-state system and the systems influenced by the electric field. Equation (8.1) is used to determine the polarizability of the group, where Equation (9.7) is used to determine the change in the dipole moment. The exposed surface area and the distance to the isodensity surface are accepted as successful computations upon comparison of the Monte

Carlo volume with the integrated volume. Examples of successful and unsuccessful computations are detailed in the next section.

The computations of AIM properties of functional groups are the most time-intensive task in this work. The integration routines usually involve determining the bounding zero-flux surfaces that partition a molecule. This task is approached in several different ways, each referenced in Section 9.4, and none of the routines researched offers a combination of elegance, correctness and timeliness. Also, there are numerous such integrations necessary to extract the information for this work, particularly the polarizability. For each functional group, there are four AIM integrations and one routine to determine the external surface area. This task can made somewhat easier to accomplish by creating a batch file to run multiple tasks in series from a DOS prompt.

10.4 Exposed Surface Area of a Functional Group

To determine AIM properties, the integrators referenced above perform three-dimensional integrations on the space that constitutes the functional group. In directions where no zero-flux boundary surface (and thusly no adjacent functional group) exists, the integration ray extends out to points where the electron density is deemed negligible. However, for structural properties, such as the volume, an isodensity surface must be defined to enclose the space. This isodensity surface can therefore be considered the exposed surface area of the functional group, and its size aids in the application of AIM properties to engineering models of fluid behavior.

Models that employ functional group interactions attempt to approximate the external surface area of functional groups. This property, in turn, approximates the

number of intermolecular interactions available to that portion of the molecule. Models such as UNIFAC use the van der Waals radius of atoms to determine how large this surface is. The COSMO-based models use a value slightly larger than the van der Waals radius that is correlated to experimental data. This work employs the 0.001 au isodensity surface, due to its significance in small molecule interactions. The various surfaces surrounding a *trans*-1-propanol molecule and how they relate to the electron topology of this molecule are shown in Figure 10.1. As one can see, the surfaces corresponding to the van der Waals radii or some multiple of it do not correspond to any one isodensity surface.
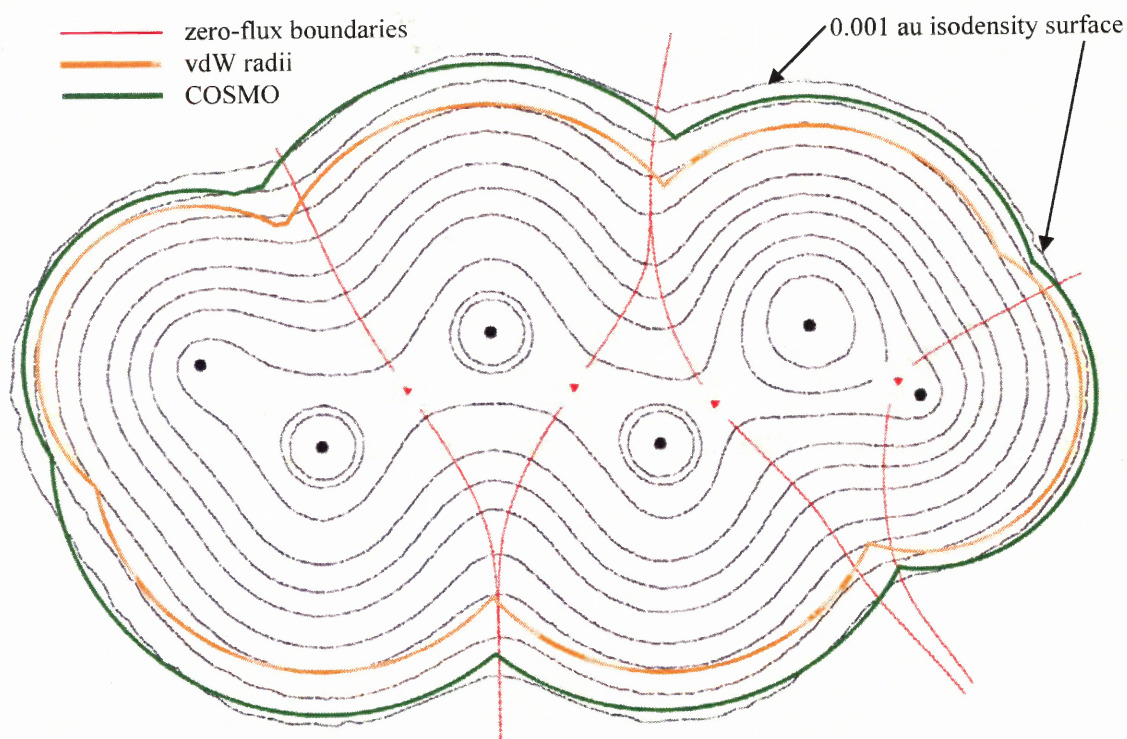
Figure 10.1 Electron density topology of *trans*-1-propanol.

The use of a consistent surface across all molecules, whether it be the 0.001 au surface or the 0.002 au surface (another important surface in the study of AIM theory) allows for a non-arbitrary standard across all molecules. The surface determined within the COSMO-based schemes here roughly corresponds to the 0.001 au isodensity surface in this plane, but this surface contains crevices that have needed to be fixed in subsequent works.

The calculation of the size of the 0.001 au isodensity surface and its average distance from the functional group center motivates the creation of numerical routines that employ the PROAIMV zero-flux boundary method. Three separate steps are taken for the determination of the exposed surface area: determination of the set of points (r,θ,ϕ) that bound the integrated volume in the PROAIMV routine; determination of the points that correspond to $\rho(r',\theta,\phi) = 0.001$ au; and the summation of the area that is created by the points with $\rho(r',\theta,\phi) = 0.001$ au. The three steps correspond to the images within Figure 10.2.

The first step in the routines collects the extent of the integration, r, for each θ and ϕ, where the origin of the coordinate system exists on the nucleus. For the rays that end on the zero-flux boundaries partitioning the molecule (depicted in Figure 10.2 as the red surfaces), the points are given as the location of zero-flux surfaces (these points are not depicted in the figure to avoid confusion). The important points for the exposed surface area calculation are those that extend toward infinity beyond the 0.001 au isodensity surface (in gray). These points are shown in Figure 10.2a, and their coordinates are tabulated.
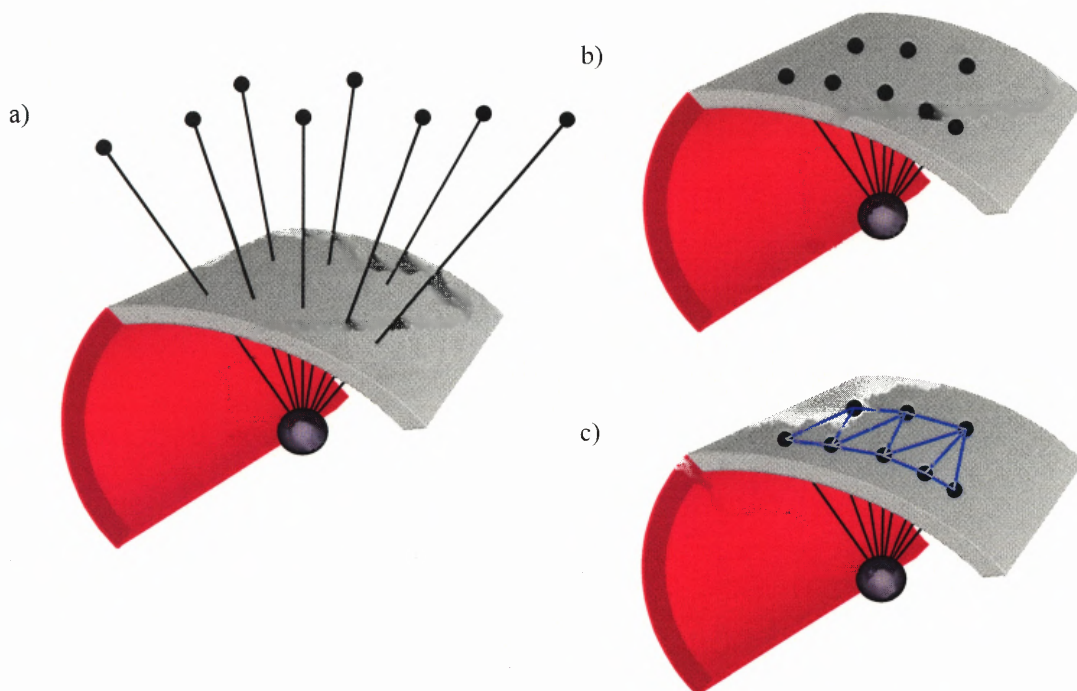
Figure 10.2 Three step routine to determine the location and size of the exposed surface area of an atom.

The second step in the routines is to make sure the points found in the first step either lies on a zero-flux surface or the 0.001 au isodensity surface, depicted in Figure 10.2b. The integration points that extend beyond the surface are now assigned a new r' on the same θ and ϕ so that $\rho(r',\theta,\phi) = 0.001$ au . A depiction of the set of points that enclose a hydrogen atom and an oxygen atom within a water molecule are shown in Figure 10.3. Once these points are tabulated, the distances are averaged to determine the distance of the exposed isodensity surface from the central nuclei of the functional group.

The third step in the routines is to determine the exposed surface area from the points found in the second step. As depicted in Figures 10.2c and 10.4a, triangles are drawn between adjacent ϕ (longitudinal) lines at each intersecting θ (latitudinal) line. In
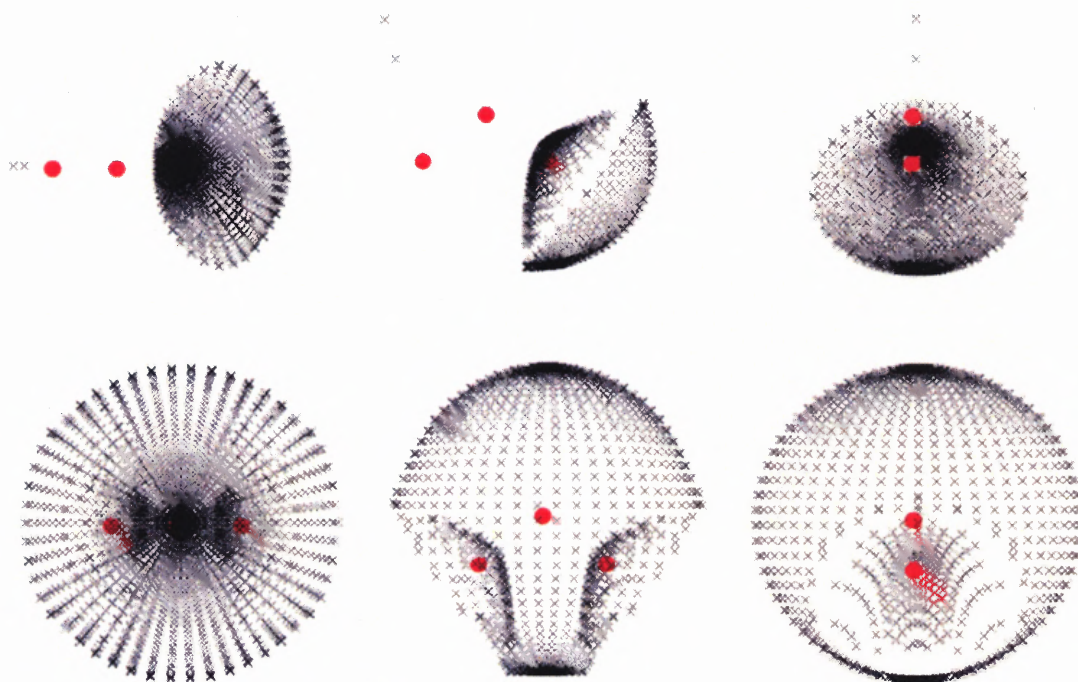
Figure 10.3 Points on the exposed and zero-flux surfaces of water. Structure of a hydrogen atom (above) and an oxygen atom (below) is shown from three different perspectives.

the cases where not all the θ exist for a given ϕ-line, triangles such as those drawn in Figure 10.4b approximate the surface area in that vicinity. The areas of all the triangles are thereby added to give the exposed surface area of the functional group.

In conjunction with the calculated surface area of the third step is a Monte Carlo calculation of the functional group volume. It is sometimes the case that the reduction in length of the rays (the step from Figure 10.2a to Figure 10.2b) fails to place the points on to the 0.001 au isodensity surface. The MC volume is used to verify that the surface area summed from the triangles offers the surface area of the group. Table 10.1 shows various
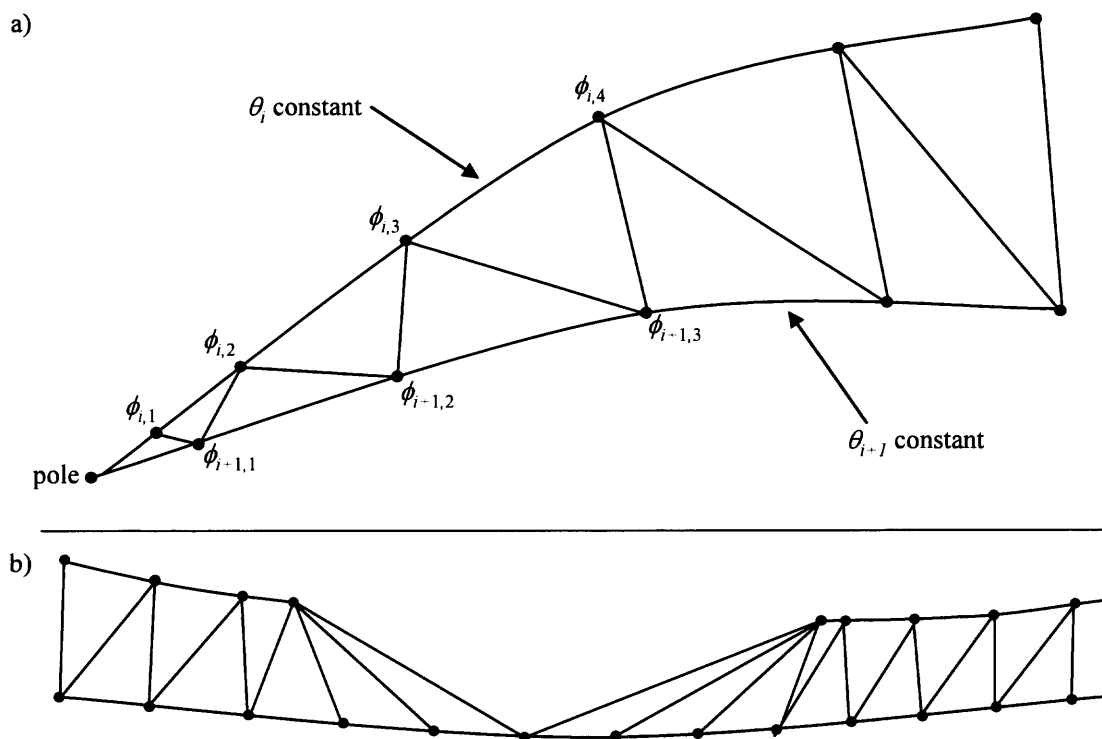
Figure 10.4 Depiction of the method to quantify the exposed surface area through the summation of triangle areas.

calculations of the integrated volumes, MC volumes, errors attributable to the MC volumes, and surface area of the functional groups. Note here that two factors can trigger suspicion in the calculated area. First, if the MC volume does not correspond to the integrated volume, as in the methylene group within 1-butanol, the points amassed by the above routines do not correspond with the isodensity surface. Also, if the MC volume is close to the integrated volume, but the error in the MC volume is larger than 1.0 au, as in the methylene group in 1-propanol, then the area calculations are off the actual values. The demonstration also shows what can happen to groups with very few exposed surface points, as in the carbon atom in various molecules. A very small error in the routines above can lead to a gross error in the calculation of a small surface area.

Table 10.1 A Comparison of Integrated Functional Group Volumes to Monte Carlo (MC) Functional Group Volumes and Results of Exposed Surface Area Calculations

group	molecule	integrated volume	MC volume	error in MC calculation	A
CH$_2$	3-methyl-1-butanol	148.38	148.26	0.636	80.84
	gauche-1-butanol	150.87	201.15	15.472	115.71
	gauche-1-propanol	151.35	151.86	5.977	216.42
C	neopentane	40.84	40.64	0.373	0.002
	2-methyl-2-propanamine	37.54	2154.29	202.874	547.27

10.5 Results

The properties of functional groups within most of the molecules listed in Section 8.2 are presented in Appendix G. The molecule name, CAS registry number and an image of the conformer is included to ease reference to other works that have calculated AIM properties. The functional group properties included in the tables are the following: group name; the number of protons in the group; the partial charge; the energy of the group; the numerical error associated with the PROAIMV integration; the magnitude and direction of the dipole moment; the polarizability; the volume; the exposed surface area; and the average distance from the nucleus denoting the group center to the exposed surface area. More details on how to read and use the tables are offered in Appendix G.

The tables include group definitions that have been used within past GCMs. Properties for the full hydroxyl group, OH, and the amine groups, NH$_2$ and NH, have been calculated for those who wish to devise a GCM based around these group definitions. Also, these calculations serve as checks to the more refined calculations on the separate atoms in the group.

The rigorous nature of the partitioning of the electron density into functional groups allows for an analysis of a particular functional group over a series of molecules. As generally seen within the tables in Appendix G, the alkyl groups are those that are most influenced by the environment within the molecule, while the properties of the electronegative groups tend to fluctuate less.

The extent of how the electron density is affected by an electronegative functional group is seen in the effects on the electron density topology in Figure 10.5. Over the series of molecules CH_3X, where $X=CH_3, NH_2, OH, F$, the bond critical point tends more toward the methyl group as the electronegativity of the bonded group increases (shown in Figure 10.5a). The proximity of the bonded group shows its influence directly on the electron density profile of the methyl group, namely by occupying more electron density situated between the groups. If the electronegative group exists on the other end of a long molecule, as in the series $CH_3(CH_2)_n X$, where $n \geq 3$ and $X = CH_3$, NH_2, OH, F, the effect on the methyl group electron density is nearly negligible however electronegative the opposite terminal end is (shown in Figure 10.5b). This is also seen when the partial charge of the methyl group is plotted against the distance from the terminal electronegative group, shown in Figure 10.6. The charges for groups within amines, alcohols and fluorides all approach the charge of the alkane, $-0.015e$, as the chain length increases. Also, a terminal oxygen or fluorine atom affects the charge of the methyl group when it is two groups away from the atoms, while the effect of the nitrogen atom approaches that of a carbon atom. The effect of the terminal electronegative group is nearly negligible at a distance of three groups and greater.
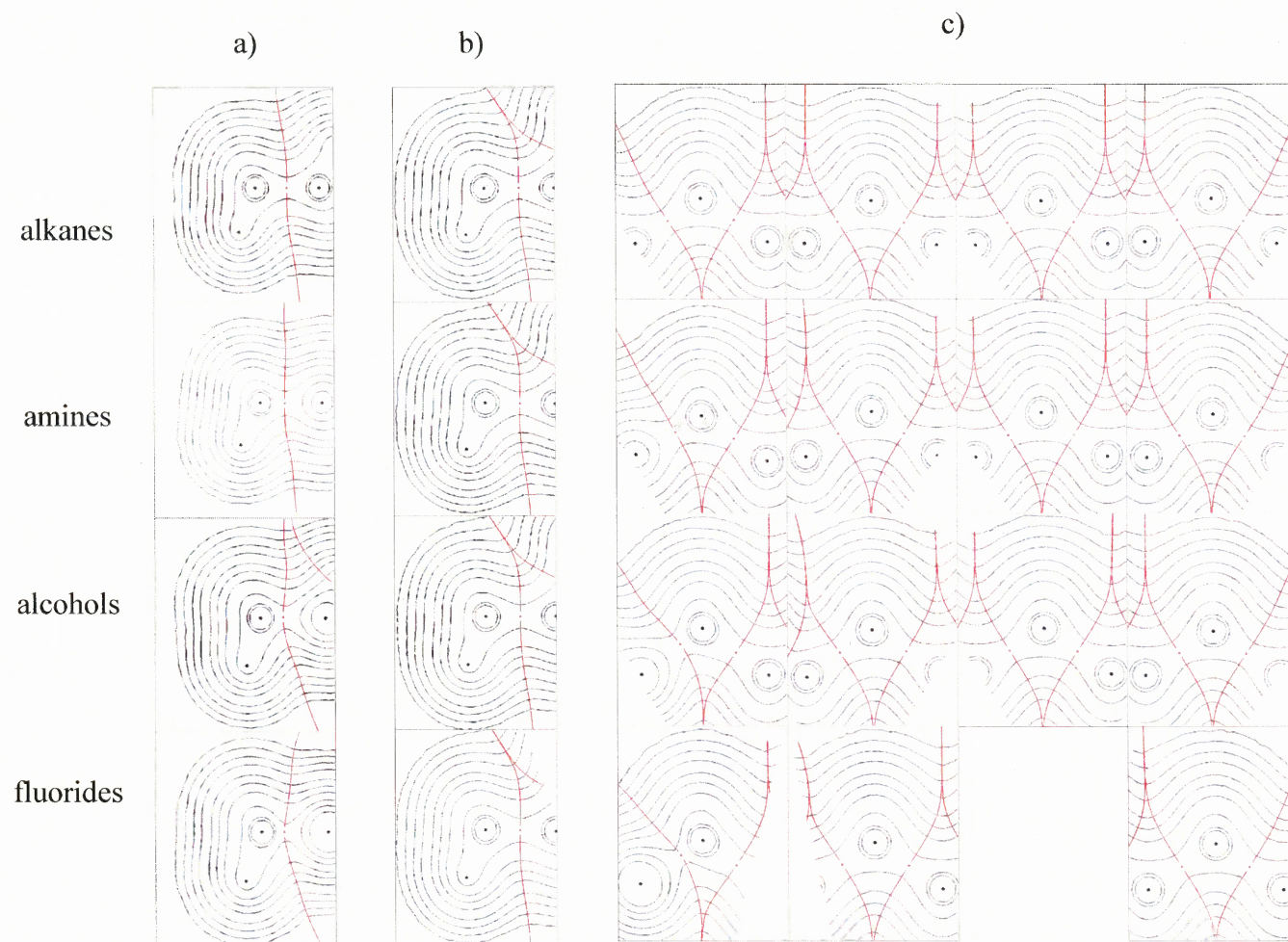
Figure 10.5 Topological Depictions of Methyl and Methylene Groups along Linear Organic Molecules
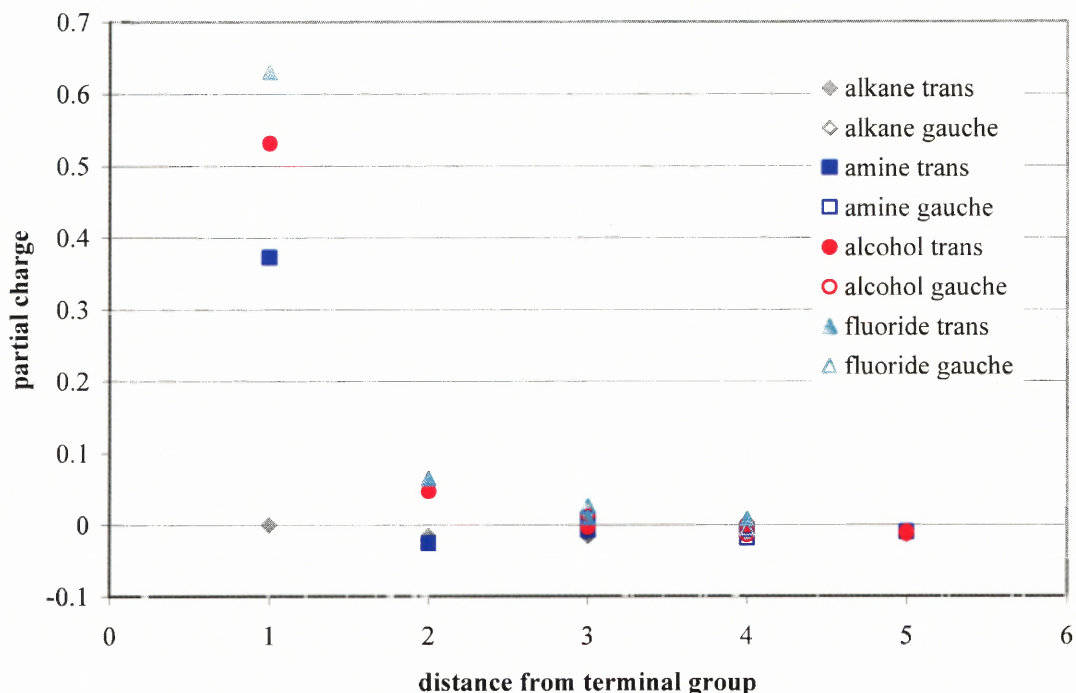
Figure 10.6 Partial charge of a methyl group as a function of the distance from a terminal group with an electronegative atom.

The effect of the terminal electronegative group along the length of the molecule is seen when inspecting the topology of methylene groups in linear molecules, presented in Figure 10.5c. The methylene closest to the electronegative group, shown in the left column, responds to the increased electronegativity of the terminal group (C, N, O, F going top to bottom) by yielding more space and electron density to the terminal group. A large positive partial charge results, as expressed in the tables in Appendix G. As the methylene groups tend farther away from the terminal group, the influence diminishes. The farthest methylene group in these linear molecules, all attached to the terminal methyl group, all are very much similar to one another. These differ slightly from the other methylene groups by the boundary they make with the methyl group in this plane.

The effect on the partial charge of the methylene groups is shown more quantitatively in Figure 10.7. Here, one can note the direct influence on the partial charge that the adjacent electronegative group has, while this effect becomes negligible (assumes the same value as in an alkane) as one proceeds further away. The more electronegative groups do have an effect on the second-bonded methylene, as seen in the topology of these groups in Figure 10.5c. In Figure 10.7, the partial charges of the second-bonded methylenes are affected by the oxygen atom and the fluorine atom. Table 10.2 shows the difference of these partial charges over all of the molecules considered; there is a distinct difference between the second-bonded methylenes in alkanes and amines and those within alcohols and fluorides.
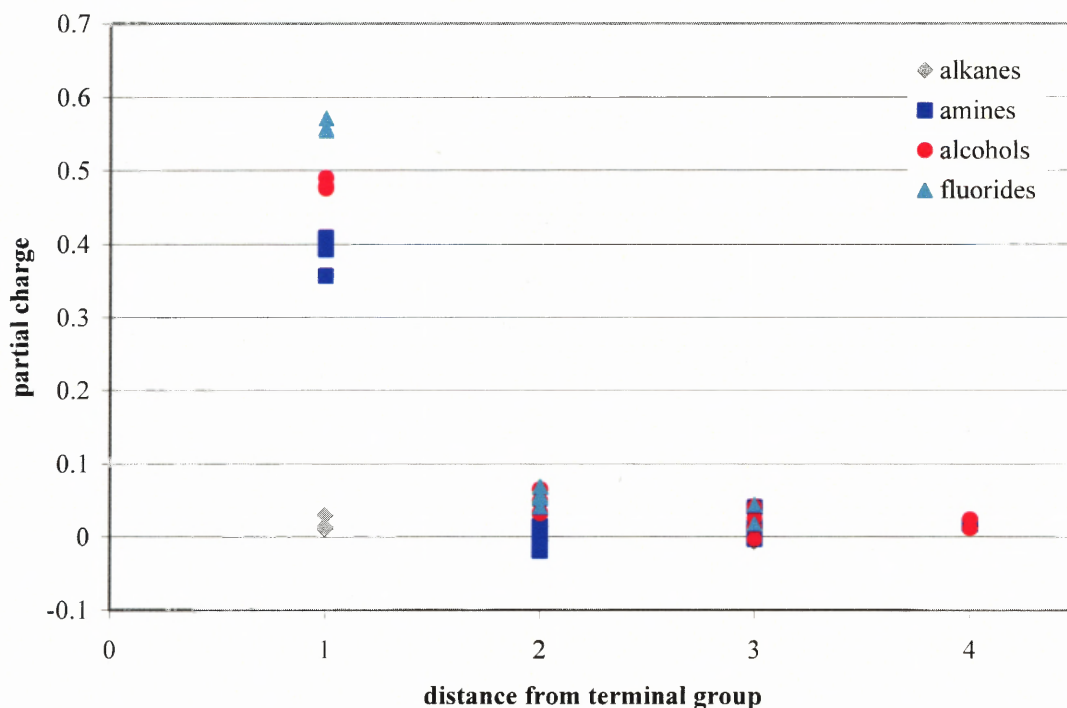
Figure 10.7 Partial charge of a methylene group as a function of the distance from a terminal group with an electronegative atom.

Also of note within the tables in Appendix G is the distribution of charges for hydrogen, oxygen and nitrogen atoms. These quantities are presented in Table 10.3, as is the molecule within which these groups occur in and, for the hydrogen atoms, what the atom is attached to. The distribution for the hydrogen atom is large: from showing a slightly negative charge in methane to being stripped of electron density by atoms of increasing electronegativity. The charge on the oxygen atoms in this study is a more stable quantity since it is mostly attached to carbon atoms (singly- or doubly-bonded) or hydrogen atoms. The charges mainly fall between $-1.0e$ and about $-1.2e$. A large deviation from this is seen in 1-nitropropane, where the oxygen atoms are directly bonded to a nitrogen atom. The nitrogen atoms within this study also show similar variability, mainly between $-1.0e$ and about $-1.3e$. Again, a difference is seen when the atom is attached to another highly electronegative atom, where oxygen atoms tend to pull charge away from the nitrogen atoms.

Table 10.2 Average Partial Charge of Second-Bonded Methylene Groups

terminal group	average partial charge
CH_3	0.002
NH_2	-0.005
OH	0.046
F	0.055

As noted earlier in Section 9.4, the numbers within the AIM property calculations are subject to error associated with the numerical nature of the integrations. The majority of the calculations within this study fell within the error tolerance stated earlier, where an error of $|L(\Omega)| < 0.001$ au assures a maximum error in the electron density of $0.0022e$.

Table 10.3 AIM Partial Charges of Hydrogen, Oxygen, and Nitrogen Atoms within a Variety of Molecules, from This Work

H			O		N		
attached	q	molecule	q	molecule	q	molecule
CH_3	-0.034	methane	-1.202	water	-1.330	propanamide
C	0.209	hydrogen cyanide	-1.187	propanamide	-1.232	ammonia
N	0.377	methylisopropanamine	-1.187	butanoic acid	-1.197	butanenitrile
N	0.391	ethanamine	-1.184	propyl ethanoate	-1.165	2-propanamine
N	0.411	ammonia	-1.154	butanoic acid	-1.112	diethylamine
N	0.468	butanamide	-1.132	1-methoxy-2-propanol	-1.063	trimethylamine
O	0.579	2-propanol	-1.130	3-pentanone	0.094	nitrous oxide
O	0.601	water	-1.123	1-propanol	0.272	nitrous oxide
O	0.619	ethanoic acid	-1.110	butanal	0.425	nitropropane
F	0.720	hydrogen fluoride	-1.089	propyl ethanoate		
			-1.067	1-methoxy-2-propanol		
			-1.062	methyl propyl ether		
			-0.486	1-nitropropane		
			-0.477	1-nitropropane		

The calculations on some larger molecules, such as methyldiethylamine and 3-methylpentane, have larger errors in the integrations. Therefore, the group properties of these molecules have been left out of the database for the time being. A different type of error existed for *cis*-2-butene, where the electron density shows a local minimum between the two terminal methyl groups. This results in what AIM theory calls a ring point, a local minimum that occurs between non-bonded atoms or groups within a molecule. Studies into this molecule, as for other molecules with ring points (such as benzene or cycloalkanes) have not been conducted due to the existence of ring points and due to the inability of the surface area routine to account for such points.

10.6 Comparisons with Prior Calculations and Experiment

The calculations of the functional groups properties within this work offers a wealth of new information not made available previously in such a large and varied scale. However, to test the validity of these quantities, one must relate these results to measurable quantities found through experiment. Unfortunately, no experiments at this point can measure the properties of a portion of a molecule. Also, the functional group definition, however theoretical, must be viewed at a point in the motion of a molecule that only serves as an average depiction of its nature.

Some AIM properties can be assembled in such ways as to return full molecular properties found in the original G98W computations or even experiment. As noted in the previous sections on AIM theory, the functional group energies, partial charges, polarizabilities, volumes and areas must add up directly to the appropriate molecular quantities.

The dipole moments presented for the functional groups are magnitudes and general directions of dipoles and do not obey additivity as the other quantities. The sign of the number only depicts whether the dipole is pointing outward toward the center of the exposed surface area ($\mu > 0$), or inward away from that point ($\mu < 0$). One can combine the vector quantities and return back to the molecular, and possibly experimental, dipole moments when one assembles the locations of the positive and negative charge centers of the atoms. The water molecule arrangement and the distribution of partial charges within the space of the molecule are presented in Figure 10.8. The partial charges are calculated from the integrations through spaces of the atoms, and the locations of the centers of negative charge are found by using the calculated
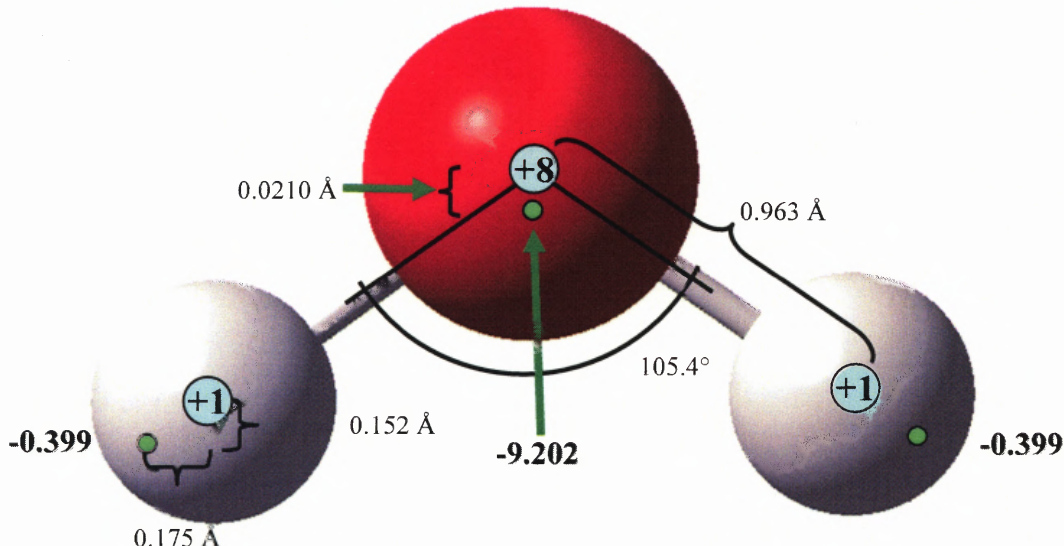
Figure 10.8 Distribution of point charges determined by AIM partial charge and dipole moment calculations on water. The green points are the locations of the centers of negative charge.

dipole moment vectors relative to the atom center. By using the formula for the dipole moment in a given direction, e.g. the z-direction,

$$\mu_z = \sum_i \left(q_i^+ z_i^+ + q_i^- z_i^- \right) \tag{10.1}$$

the dipole for the entire molecule is found. Here, q represents the partial charge at a charge center, z is the distance relative to the origin of the system (which can be any point in a neutral molecule, unlike a non-neutral system), and the superscripts + and - denote the positive and negative centers, respectively, in atom i. The calculation of the molecular dipole moment for water using the quantities in Figure 10.8 is $1.854\,D$,

whereas the full molecule calculation gives, from Table E.16, 1.857 D and the experimental value is 1.855 D.

Comparisons of structural properties are able to be made with values of past engineering GCMs, namely the Bondi scheme for UNIFAC and the COSMO-based surfaces. The volumes and surface areas from these schemes are compared with results from this work in Table 10.4. Note that the documented values do not include the single

Table 10.4 Structural Properties for Functional Groups using Published Values, Calculated Values using the Methods of Bondi, and Calculated Valued from This Work (Bondi, 1964, Klamt, et al. 1998)

	attached to	molecule found in	V (cm^3/mol)				A (nm^3)			
			Bondi lit	Bondi calc	COSMO	calc	Bondi lit	Bondi calc	COSMO	calc
CH$_3$	C	ethane	13.67	13.51	17.51	19.84	2.12	2.10	2.35	2.51
CH$_2$	C,C	propane	10.23	10.05	12.26	14.32	1.35	1.33	1.32	1.54
CH	C,C,C	2-methyl-propane	6.78	6.60	7.01	8.87	0.57	0.55	0.30	0.69
C	C,C,C,C	neopentane	3.33	3.15	1.76	3.64	0	0.00	0.00	0.00
H	C	methane		1.84	1.19	4.67		0.49	0.38	0.75
H	N	ammonia		2.26	1.78	2.77		0.56	0.48	0.50
H	O	water		2.34	2.21	1.93		0.57	0.55	0.38
H	F	hydrogen fluoride		2.38	1.81	1.39		0.58	0.49	0.31
N	H,H,H	ammonia		6.90	13.06	13.45		0.82	1.49	1.72
N	C,H,H	methanamine		5.48	9.27	11.25		0.57	0.92	1.19
N	C,C,H	dimethylamine		4.07	5.48	9.37		0.33	0.36	0.79
N	C,C,C	trimethylamine	4.33	2.66	1.69	7.82	0.23	0.08	0.00	0.52
O	H,H	water		6.75	10.89	13.53		1.02	1.48	2.00
O	C,H	methanol		5.38	7.37	11.48		0.79	0.97	1.46
O	C,C	dimethylether	3.7	4.02	3.85	9.49	0.6	0.56	0.45	0.99
F	H	hydrogen fluoride		7.10	12.99	11.66		1.29	2.00	1.97
F	C	fluoromethane	5.72	5.76	9.15	10.13	1.1	1.06	1.43	1.44

atoms, such as the nitrogen atom within an amine group. For groups of atoms, similar trends can be seen throughout the methods, but those used in UNIFAC and COSMO-based models tend to attribute more volume to the carbon atom than the AIM integrations. This is predictable because the separation of space within the former schemes is arbitrarily equal. The AIM treatment reflects the electronegativity by yielding more space to the more electronegative atoms (as seen in the analysis of alkyl groups in linear molecules in Figure 10.5). It must be noted that COSMO-based methods are not sensitive to which functional group a piece of surface area belongs to, as is the case with UNIFAC and this work.

As noted earlier, group polarizabilities are linearly additive and result in the value of the molecular polarizability calculated for the molecules in this work. The comparison of the molecular polarizabilities, and thusly the group polarizabilities, with that from experiment is given in Section 8.5. Prior calculations of the polarizability for methyl and methylene groups (Bader, et al. 1992) are comparable to those from alkyl methyl and methylene groups from this work. The past results (approximately $\alpha_{CH_3} = 13.5$ au and $\alpha_{CH_2} = 11.3$ au) correspond with the calculated results from this work (approximately $\alpha_{CH_3} = 14.4$ au and $\alpha_{CH_2} = 11.8$ au). The difference is likely due to the larger basis set used in this work.

10.7 The Spherical Gaussian Approximation

A property of interest within the quasi-chemical equations is the exposed surface area, since this depicts the number of interactions a functional group may participate in. As suggested by AIM theory and by of COSMO-based interactions, molecules within the

liquid phase tend to interact by having these exposed surface areas overlapping tangentially. Thus, the calculation of the quantity r_{avg}, the average distance from the center of the functional group to the exposed surface area, is conducted to give information on the interaction distances between functional groups within such models.

Within this work, r_{avg} is found using the electron density of a functional group at the 0.001 au isodensity surface. It is assumed that, at these interaction distances, the sum of the r_{avg} between the two groups, enough electron density overlap occurs to develop a significant repulsion effect. A spherically symmetric wavefunction is used to approximate the electron density around r_{avg}, thus allowing for a doubly-occupied orbital representation of the electron density of the functional group at these distances. Observations on the functional group interactions using this information are presented in the next two chapters.

Two types of spherically-symmetric functions may be used to approximate an atomic orbital: a Slater-type orbital (STO) and a Gaussian-type function (GTF). The STO is given by Equation (7.18), where a normalized $1s$-type function is given by

$$s_{100} = \frac{\zeta^{3/2}}{\sqrt{\pi}} e^{-\zeta r} \tag{10.2}$$

The GTF is given by Equation (7.19), where a normalized $1s$-type function is given by

$$g_{000} = \left(\frac{2\xi}{\pi}\right)^{3/4} e^{-\xi r^2} \tag{10.3}$$

These functions, as depicted in Figure 7.1, offer different functionalities as they approach the center of the orbital, but offer similar nature near the tail of the wavefunctions.

Both of these functions contain orbital exponents, ζ and ξ, that determine the extent of the space within which an electron exists. In this work, these exponents are used to reproduce the extent of the electron density of a functional group to correspond to an electron density of 0.001 au at the distance of r_{avg}. To accomplish this, it is assumed that doubly-occupied orbital represent the functional groups. For STOs, the exponent ζ is found by solving

$$0.001 = 2s_{100}^2 = 2\left(\frac{\zeta^3}{\pi}\right)e^{-2\zeta r_{avg}} \qquad (10.4)$$

while for GTFs, the exponent ξ is given by

$$0.001 = 2g_{000}^2 = 2\left(\frac{2\xi}{\pi}\right)^{3/2} e^{-2\xi r_{avg}^2} \qquad (10.5)$$

Figure 10.9 depicts the functionality of the exponents with respect to r_{avg}. One may notice how smaller exponents are necessary to depict more diffuse electron densities.

Of particular interest in this work is how GTFs reproduce the electron densities of small molecules and functional groups. This is important because these functions, called spherical Gaussian approximations (SGA) to the electron density profile, are used within a binary interaction function in Chapter 12 which, with other AIM properties, determines interaction energies and distances. For small molecules, the SGA serves as the spherical average of the electron density, while for a functional group, the SGA serves as the extent to the electron density in directions important in the development of a repulsive interaction. Figure 10.10 depicts the SGA against the electron density profiles of a methane molecule and a nitrogen molecule. It is shown how the 0.001 au isodensity surface of the SGA, depicted by the blue line, corresponds to this surface on the
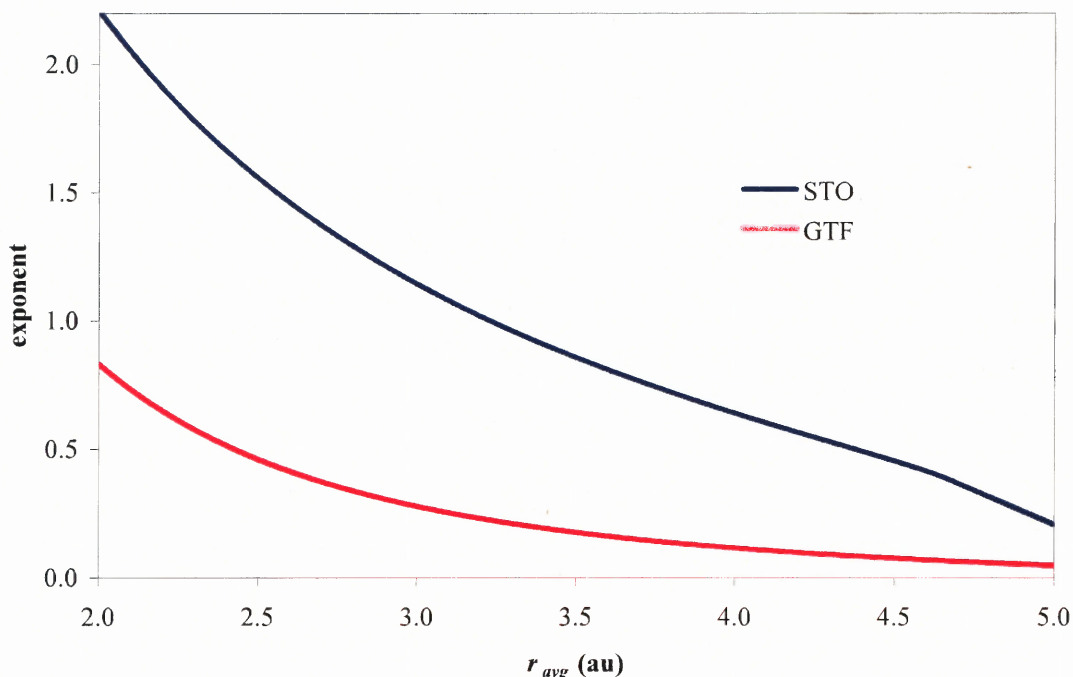
Figure 10.9 Spherically-symmetric orbital exponents as functions of r_{avg}.

molecules, depicted by the outermost isodensity line. As these molecules vibrate and rotate in the gas phase, the SGA depicts approximately how these molecules behave, since the rotational and vibrational motions are so fast at ambient temperatures that the electron densities of these molecules may be assumed spherically symmetric.

Figure 10.11 depicts the electron density profiles of the SGA against three rays from the center of a methane molecule. Ray 1 samples the electron density along a C-H bond and continues beyond the H-atom. Ray 2 bisects two C-H bonds. Ray 3 trisects three C-H bonds. As expected, the magnitude of the electron densities along the rays always shows ray 1 greatest, followed by ray 2 and ray 3. The SGA predicts that the spherical average at the 0.001 au isodensity surface falls nearly on ray 2, and between all three rays. The density of the SGA tends to not rise as quickly as it approaches the center
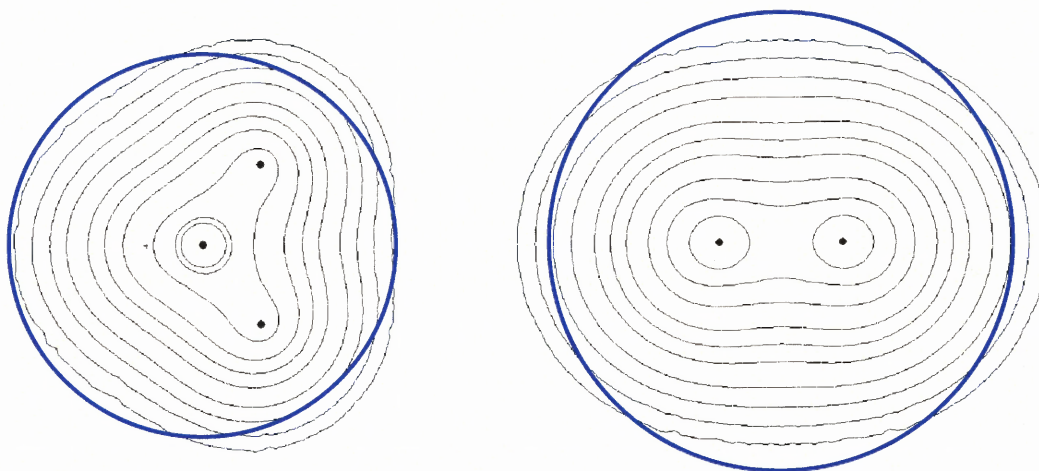
Figure 10.10 Spherical Gaussian approximations to methane and nitrogen.
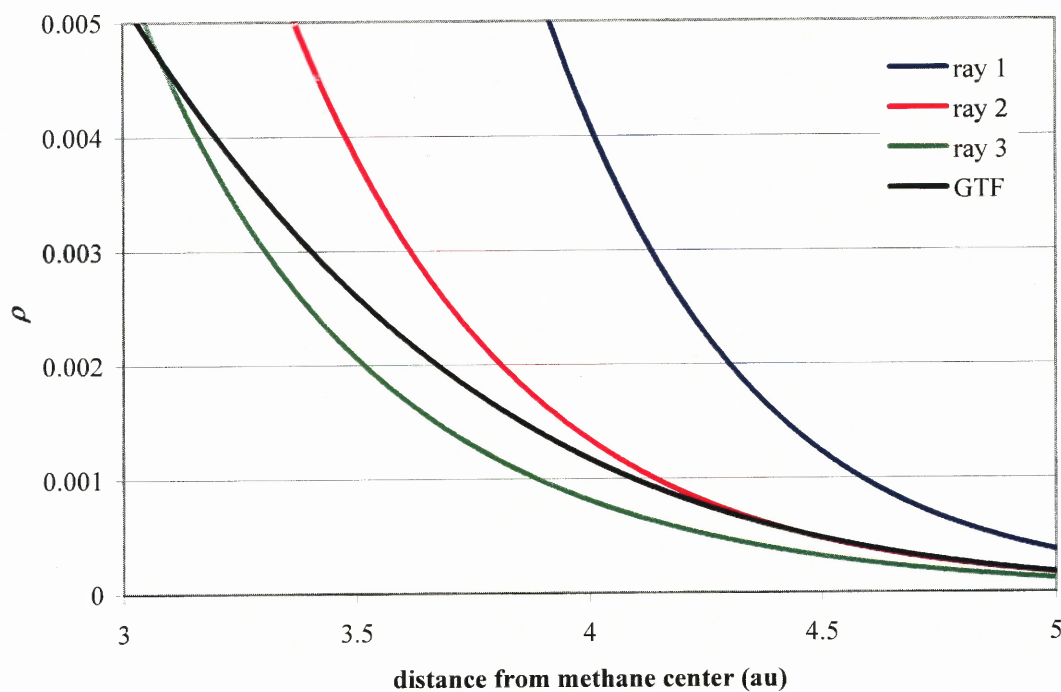
Figure 10.11 Electron density profile of the spherical Gaussian approximation compared to three rays originating from the center of a methane molecule.

of the molecule. This is true because the orbital depicted by the GTF only contains two electrons, while the electron density within the molecule accounts for 10 electrons.

Figure 10.12 shows how the SGA reproduces the external surface area of the alkyl functional groups CH_3 and CH_2. In both graphs, the SGA, depicted by the blue line, reproduces the exposed surface areas well (for the latter, the surface near the top for the CH_2 group). The expected bulging of the group electron density exists near the hydrogen atoms and the sunken portions exist between hydrogen atoms.

Figure 10.13 relates the electron density profile of the SGA against three rays within the methyl group. The rays here sample density on similar paths as those in Figure 10.11. A similar functionality exists between the SGA and the electron densities at the tail. Again, the SGA predicts the location of the 0.001 au surface in the middle of the three rays.

The importance of Figures 10.11 and 10.13 is that the SGA approximates the functionality of the electron density tails rather well. This is important due to the overlap effects that these portions of the electron densities experience in intermolecular interactions. The need to use STOs to reproduce the electron density may give better behavior as one approaches the centers of the molecules and the functional group, but use of these results in a much more complicated interaction function. The use of GTFs to approximate the electron densities is adequate for this work.
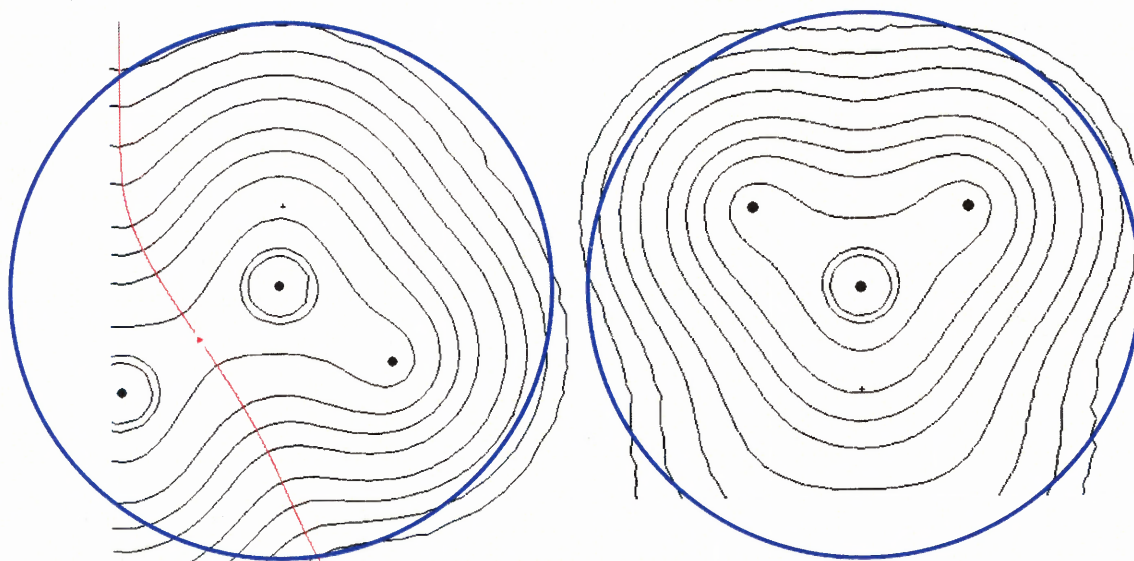
Figure 10.12 Spherical Gaussian approximation of a methyl group and a methylene group in propane.
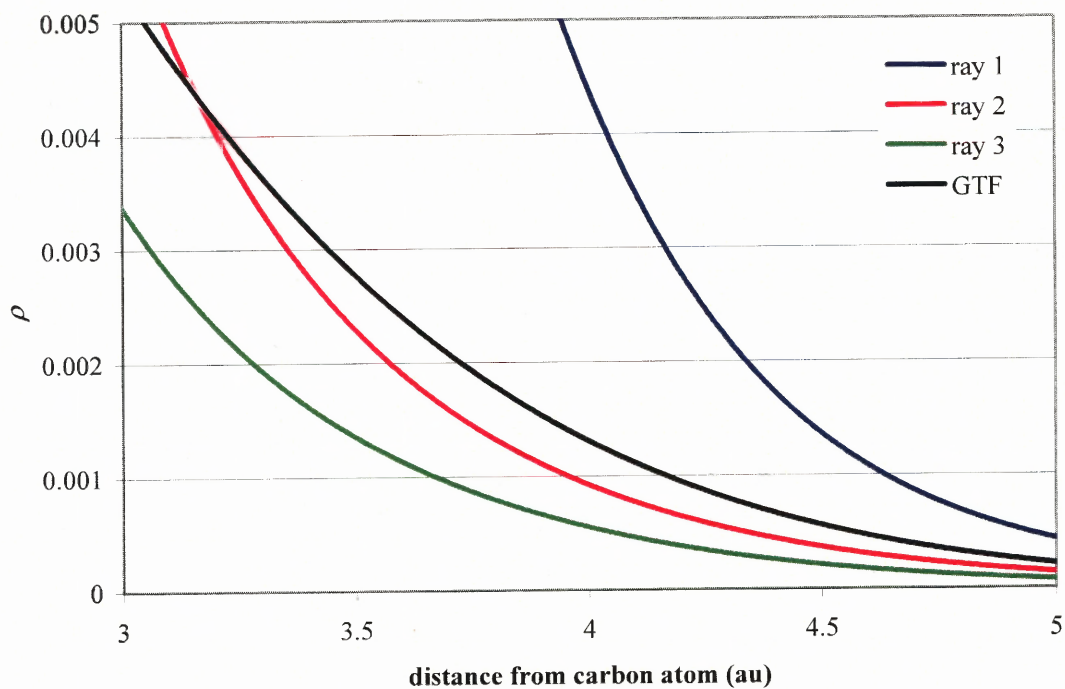
Figure 10.13 Electron density profile of the spherical Gaussian approximation compared to three rays originating from the carbon atom of a methyl group.

10.8 Conclusions

Models used by chemists and engineers have in the past either approximated or correlated functional group properties. These quantities have contributed to useful group-contribution methods for the thermochemistry of macroscopic systems. However, with the development of functional property determination methods within the last two chapters, such correlations should be unnecessary in the future.

The ability to calculate rigorous functional group properties allows for a significant amount of simplicity when building a group-contribution method. Functional group definitions are allowed to be smaller due to the use of rigorous properties as higher-order descriptors. Transferability can be eliminated completely as an assumption, except for use to estimate functional group properties in those cases when it may be quantitatively applicable. The structural and electrostatic properties presented in this chapter serve as physically significant characteristics that can serve as the basis for future modeling efforts. With further computations into functional group properties, especially at higher levels of *ab initio* theory, a database of properties can be formed and powerful models can be created to take advantage of such information.

This work utilizes electrostatics and the SGA for interaction energies, and structural parameters for statistical modeling to determine macroscopic system behavior from first principles. With the structural properties having been calculated, intermolecular interaction theory is explored to determine how the electrostatic properties of a functional group can reflect the nature of an interaction.

CHAPTER 11

INTERACTIONS BETWEEN CLOSED SHELL MOLECULES

The study of the interactions of closed-shell molecules continues to be a hurdle in the development of theoretical treatments of fluids in the liquid phase. In the gas phase, molecules exist at large distances from one another, allowing for descriptions using classical methods. The functional forms presented in Section 3.4 combined with a simple repulsive term, such as a hard-sphere repulsion, offers an adequate model of molecular behavior of real gases near the ideal gas limit. Such a simplification does not exist in the liquid phase, where molecules are always close together and where the form of the repulsive term of any interaction model takes on importance.

The goal of this research is to find an intermolecular potential energy function within which the functional group properties presented in Chapter 10 may be applied. This will serve as a simplification of an interaction between two molecules, where it will be assumed that a functional group interaction is the significant contribution. It is presumed that functional groups offer the defining properties that cause a favorable or unfavorable interaction, an assumption made in all group-contribution methods employing group interactions (Fredenslund, et al. 1975; Kehiaian, et al. 1978; Knox, 1987). Three contributions are found to be important within an interaction potential energy function for functional groups: short-range repulsion; long-range attraction; Coulombic interaction for partial charges. These terms are explored at the molecular level.

This chapter reviews the current state of intermolecular interaction theory, including the use of perturbation theory on the Hamiltonian of the system, the application of computational chemistry on binary and cluster systems, and the functional forms most utilized in empirical potentials. Molecular-level concepts are explored as possible simplifications for use in a functional group interaction scheme. Classic work of the interaction of doubly-occupied orbitals is presented. Applications of contemporary techniques are presented for hydrogen-bonding systems, as well as for calculation of macroscopic system properties. Here, the use of effective functional group interactions within molecular dynamics simulations is seen to yield promising results.

11.1 Results from Perturbation Treatments

When one considers a quantum system of two interacting molecules, the Hamiltonian of the system must be defined. An approach to solving the problem is thereby similar to that of an isolated molecule; one writes the Hamiltonian of the system as in Equation (6.31) and attempts to solve for the wavefunction as in Equation (6.2). The Hamiltonian of the system includes all the interactions between electrons of molecule A and electrons of molecule B, nuclei of molecule A and electrons of molecule B, nuclei of molecule B and electrons of molecule A, nuclei of molecule A and nuclei of molecule B, as well as the interactions within the two isolated molecular systems. This is shown to be

$$\hat{H} = \sum_i \sum_j \frac{1}{r_{ij}} - \sum_i \sum_\beta \frac{Z_\beta}{r_{i\beta}} - \sum_j \sum_\alpha \frac{Z_\alpha}{r_{j\alpha}} + \sum_\alpha \sum_\beta \frac{Z_\alpha Z_\beta}{r_{\alpha\beta}} + \hat{H}_A + \hat{H}_B \qquad (11.1)$$

where i and α cycle through the electrons and nuclei of molecule A, respectively, and where j and β cycle through the electrons and nuclei of molecule B, respectively. The

energy of interest within the study of interactions between molecules is the difference between the total system and the isolated molecular systems

$$\varepsilon_{AB} = E_{AB} - E_A - E_B \tag{11.2}$$

This interaction energy is found by applying Equation (11.1) and subtracting the isolated molecular Hamiltonians from the system Hamiltonian

$$\left(\hat{H} - \hat{H}_A - \hat{H}_B\right)\Psi = \varepsilon_{AB}\Psi \tag{11.3}$$

The Hamiltonian of interest is therefore now

$$\hat{H}_{int} = \sum_i \sum_j \frac{1}{r_{ij}} - \sum_i \sum_\beta \frac{Z_\beta}{r_{i\beta}} - \sum_j \sum_\alpha \frac{Z_\alpha}{r_{j\alpha}} + \sum_\alpha \sum_\beta \frac{Z_\alpha Z_\beta}{r_{\alpha\beta}} \tag{11.4}$$

This problem is as difficult as that given by the molecular Hamiltonian, although some points can be made about the nature of the solution. A binary interaction energy is a very small change from the absolute molecular energies found in quantum chemistry, so ε_{AB} is considered a perturbation. Also, in a number of solution methods, the total wavefunction of the system is assumed to be a perturbation of the product of the isolated molecular wavefunctions. This latter treatment of the problem is called the Heitler-London (HL) solution method.

For interactions between molecules at long-range, it is correct to assume that the unperturbed portion of the wavefunction is a simple product of the isolated molecular wavefunctions. Regular Rayleigh-Schrödinger (RS) perturbation theory (Buckingham, 1967) is employed using Equation (11.4) as the perturbing Hamiltonian. The resulting intermolecular potential, if one truncates after the second-order terms, is summarized as three notable effects: electrostatic; induction; and dispersion (Engkvist, et al. 2000)

$$u_{AB} = u_{electrostatic} + u_{induction}\left(B \text{ by } A\right) + u_{induction}\left(A \text{ by } B\right) + u_{dispersion} \tag{11.5}$$

where the Coulombic interactions are included in the electrostatic term, and where the last three terms usually stabilize the system. This treatment offers rigor to the attractive long-range interaction terms offered in Section 3.4.

The only difference between Equations (11.5) and (3.54) is the exchange-repulsion contribution. Because RS perturbation theory applies only at long-range, the electron densities of the two molecules do not significantly overlap. Therefore, the densities do not have a chance to experience the repulsive effects due to Pauli exclusion or an exchange-attraction at short distances. In fact, the RS theory does not work at short distances because of this omission. Several short-range perturbation theories have been created (Hayes and Stone, 1984; Jeziorski, et al. 1994) with such a contribution in mind. The main differences are that the unperturbed wavefunction is an asymmetric product of the molecular wavefunctions (a change that accounts for the exchange contribution), and that more care is taken to determine the perturbing portion of the Hamiltonian. Perturbation theories that use the asymmetric wavefunction, called Symmetry-Adapted Perturbation Theories (SAPT), have a similar exchange-repulsion contribution to the interaction energy (Stone, 1996). The exchange contribution is given by

$$u_{exchange} = -\frac{1}{2}\sum_{i}\sum_{j}\left(a_i b_j \middle| b_j a_i\right)$$

(11.6)

where the orbitals here are spin-orbitals, and where a and b denote the orbitals on molecule A and B, respectively. The repulsion term is given by

$$u_{repulsion} = \sum_{i}\sum_{j}\left\langle \psi_i \middle| \frac{-Z_A}{r_A} + \frac{-Z_B}{r_B} + \frac{1}{2}\nabla^2 \middle| \psi_j \right\rangle\left(T_{ij} - \delta_{ij}\right)$$
$$+ \frac{1}{2}\sum_{i}\sum_{j}\sum_{k}\sum_{l}\left(\psi_i\psi_j \middle| \psi_k\psi_l\right)\left(T_{ij}T_{kl} - \delta_{ij}\delta_{kl} - T_{il}T_{kj} + \delta_{il}\delta_{kj}\right)$$

(11.7)

where, again, the orbitals are spin-orbitals and the summations cycle through all the occupied spin-orbitals in the system, whether they are on molecule A or molecule B. Here, δ_{ij} represents the Kronecker delta

$$\delta_{ij} = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases} \tag{11.8}$$

and T represents the inverse of the matrix where the element in row i, column j is the overlap integral S_{ij}

$$T_{ij} = \left(S^{-1}\right)_{ij} \tag{11.9}$$

Although predominantly affecting the interaction at long-range, induction and dispersion have corresponding short-range contributions. Damping functions are used to modify these effects for systems involving small molecules (Knowles and Meath, 1986a; Knowles and Meath, 1986b; Knowles and Meath, 1987). These functions approach unity at long-range, while they tend toward zero at short-range.

11.2 Interactions Using Quantum Calculations

Interaction energies and orientations are concepts that can be studied within computational chemistry. As opposed to an isolated molecule calculation, two or more molecules can be placed within close proximity for analysis. The software packages available for geometry optimizations and single point energy calculations, one being Gaussian 98W, are used to determine favorable orientations for interactions between molecules and the quantitative energy change given that orientation.

A calculation involving multiple molecules within the proximity of one another is sometimes called a supermolecule calculation. Calculations of this type have become

extremely common, as they have been geared toward solving specific problems involving a small number of species. A review of the use of computational chemistry for non-bonded interaction energies for small molecules outlines the methods (*ab inito* and DFT) and challenges of using such techniques (Rappé and Bernstein, 2000).

Supermolecule calculations have been conducted from systems containing small molecules, such as argon and formic acid (Wawrzyniak, et al. 2004), to systems considering the aqueous effects on biomolecules (Yeganegi, et al. 2003) to systems dominated by hydrogen-bonding effects (Estrin, et al. 1996; Bartha, et al. 2003; Kozmutza, et al. 2003; Kryachko and Scheiner, 2004). Each of these studies considers the possible interaction arrangements, calculated interaction energies and how they compare to past work.

In the liquid phase and within the gas phase of strongly associating molecules, clusters of two or more molecules form and exist within the system for a time longer than a normal intermolecular interaction. Calculations on clusters of small alcohols (Wu and Sandler, 2000) reveal possible orientations for up to six molecules at close range. The binding energies of these clusters are calculated at the HF and MP2 levels of theory, and these energies are shown not to behave linearly, i.e. the binding energy of a trimer of methanol molecules does not equal three times the binding energy of a binary methanol system. This offers quantitative evidence of the need to include ternary interaction potentials within models for higher density molecular system, such as those describable using the third virial coefficients in Equation (3.41).

A non-intuitive problem with supermolecule calculations is the evolution of a stabilization energy from what is called basis set superposition error (BSSE). A naïve

way of calculating the interaction energy using the supermolecule methods involves the calculation of each isolated molecule with its basis sets (molecule A with basis A and molecule B with basis B), and the supermolecule with the combined basis sets (supermolecule AB with basis set AB). This interaction energy is given by

$$\varepsilon_{AB} = E_{AB}\left(\text{basis } AB\right) - E_A\left(\text{basis } A\right) - E_B\left(\text{basis } B\right) \tag{11.10}$$

This leads to an interaction energy that, for an orientation that yields a favorable interaction energy, is too negative. The electrons within the supermolecule have a more complete basis set, which usually leads to a lower absolute energy, than they have in the respective isolated molecule calculations. To alleviate this, it is assumed that as long as the same basis set is used in all three calculations in Equation (11.10), then the interaction energy for the system will be properly represented. This is accomplished by making the supermolecule basis set AB available to the isolated molecules. The interaction energy is now represented by

$$\varepsilon_{AB} = E_{AB}\left(\text{basis } AB\right) - E_A\left(\text{basis } AB\right) - E_B\left(\text{basis } AB\right) \tag{11.11}$$

This correction is usually called the counterpoise correction, and the process involved to alleviate BSSE is given within the user manual of Gaussian 98W.

It has been noted that the problems with BSSE may have more to do with the incomplete basis sets used in supermolecule calculations (Stone, 1996), and that calculations with a large BSSE have more fault in the original choice of basis set. This viewpoint is supported by the reduction of BSSE seen in supermolecule calculations applying systematically larger basis sets (Rappé and Bernstein, 2000).

11.3 Empirical Potentials

Several intermolecular potentials serve mainly as functional forms that represent the gross repulsion and attraction effects within an interaction. The utility of such expressions is mainly found within computer simulations of fluids, where numerous evaluations of the potential functions are necessary to determine trajectories and positions of molecules at the next time step. Potentials with a simple functional form are inexpensive computationally and are favored in large-scale calculations. Since somewhat accurate results arise when using the simpler empirical equations, they continue to be used.

Such functions tend to consist of terms that account for major effects in an interaction and fitting parameters to model the specific molecules of interest. The parameters within the expression have been found in several ways: fitted to experimental data, such as second virial coefficient or viscosity data; approximated using theoretical concepts, such as use of the van der Waals radius as a size parameter; or fitted to reproduce the potentials from the more rigorous perturbation or supermolecule methods.

Empirical potentials vary in their functional forms and the number of parameters. All try to combine a repulsive interaction at short distances and an attractive interaction at intermediate and long distances. The equations presented in the following paragraphs have similar features, and the notation is shared. Let σ be the distance at which the energy of interaction vanishes. For single-well potentials, this point is on the repulsive wall and represents the effective radius of the molecule. Let ε represent the magnitude of the potential well. This quantity reflects the strength of the attraction between the two

molecules at its most probable interaction distance. This position reflects the most probable distance of interaction.

A primitive interaction potential is the hard sphere potential. This potential takes the form

$$u^{HS}(r) = \begin{cases} \infty & \text{when } r \le \sigma \\ 0 & \text{when } r > \sigma \end{cases} \tag{11.12}$$

The hard sphere potential mimics the behavior of non-polar species in the gas phase, and viscosity data is sometimes used to correlate σ. The stepwise nature of this potential is very easy to implement in molecular simulation routines.

A modification to the hard sphere potential is the square well potential. Given by

$$u^{SW}(r) = \begin{cases} \infty & \text{when } r \le \sigma \\ -\varepsilon & \text{when } \sigma > r \ge \sigma' \\ 0 & \text{when } r > \sigma' \end{cases} \tag{11.13}$$

the square well potential adds an attractive interaction to the infinite repulsion. Here, σ' is a parameter that represents the length of the square well. Again a stepwise function, this form is easy to implement within molecular simulations. Modifications of this functional form have appeared recently as a transferable potential that contains multiple steps (Unlu, et al. 2004).

Probably the most widely used potential functional forms arose more than 70 years ago. This function is called Mie's potential and takes the form

$$u^{Mie} = \frac{C_{rep}}{r^n} - \frac{C_{att}}{r^m} \tag{11.14}$$

where $n > m$. This form is a continuous function through all interaction distances, and as long as $n > m$, the function has the characteristic interaction energy well. The exponent

within the attraction term usually takes on a value of $m = 6$, due to the attraction terms derived in classical work, Equations (3.48), (3.51) and (3.53). When $n = 12$, Equation (11.14) is called the Lennard-Jones (LJ) potential. This function is best known with σ and ε explicit

$$u^{LJ}(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right] \tag{11.15}$$

This function, although not as easy to implement into computer routines as the stepwise functions, is likely the easiest functional form that is altogether continuous, differentiable and physically meaningful.

A more flexible potential than the LJ potential with similar computational simplicity is available (Kihara, 1978). This form is given by

$$u^{Kihara}(r) = \begin{cases} \infty \text{ when } r \leq 2a \\ 4\varepsilon \left[\left(\frac{\sigma - 2a}{r - 2a} \right)^{12} - \left(\frac{\sigma - 2a}{r - 2a} \right)^{6} \right] \text{ when } r > 2a \end{cases} \tag{11.16}$$

Here, a is called the hard-core radius and offers extra flexibility with the added parameter.

For those critical of the arbitrary repulsion term, a modification is offered that reflects the exponential nature of the repulsion due to electron overlap (Born and Mayer, 1932)

$$u^{exp/6} = C_{rep} e^{-C_{exp} r} - \frac{C_{att}}{r^{6}} \tag{11.17}$$

This expression has also been modified to include a C/r^{8} term to include higher-order attractive effects (Buckingham and Corner, 1947). These expressions have the correct functional form for the repulsive wall, yet the LJ potential has become the more popular

in time. This may be so because of the difficulty of computing the exponential function before the time of fast microprocessors (Stone, 1996). An issue with the exponential-6 potential is that it does not yield an infinitely positive result at interaction distances near zero. It has been noted that such interaction schemes with an exponential repulsion term may include a r^{-21} term to avoid such inconsistencies (Brdarski and Karlström, 1998).

11.4 Approximations to Short-Range Interactions

Long-range interactions are well-studied phenomena and have as solutions the expressions given in Section 3.4, the RS Perturbation theory and a powerful technique not presented in this work, the distributed multiple analysis (Stone, 1996). As stated earlier, these expressions break down as the electron densities begin to overlap at small interaction distances. The energy developed at such a short range is of interest not only to those in the theoretical community, but to those in molecular dynamics who desire a rigorous reasoning behind the empirical potentials in Equations (11.15) through (11.17).

One of the simpler and more descriptive interaction systems in quantum chemistry is the interaction between a pair of s-type orbitals. When these orbitals are doubly-occupied, the resulting non-bonded interaction is that experienced by a pair of helium atoms. The HL interaction energy for molecules with s-type orbitals outside inner closed-shells has been studied (Rosen, 1931). The HL energy is calculated by assuming the system wavefunction is an unperturbed, asymmetric product of the isolated molecular wavefunctions. The study includes the interaction between two singly-occupied orbitals, a singly- and doubly-occupied orbital, and two doubly-occupied orbitals, with long-range effects being excluded from the study. For molecules A and B

each with a doubly-occupied s-type orbital as their outer shell, the short-range interaction energy is given by

$$u = \frac{Z_A Z_B}{r} + \frac{4}{1-S^2}\left(J_{ab} + \frac{1}{2}\left\langle b\left|\frac{-Z_A}{r_A}\right|b\right\rangle + \frac{1}{2}\left\langle a\left|\frac{-Z_B}{r_B}\right|a\right\rangle \right)$$
$$- \frac{4}{1-S^2}\left[K_{ab} - S\left(\frac{1}{2}\left\langle a\left|\frac{-Z_A}{r_A}\right|b\right\rangle + \frac{1}{2}\left\langle a\left|\frac{-Z_B}{r_B}\right|b\right\rangle \right) \right] \quad (11.18)$$
$$+ \frac{1}{\left(1-S^2\right)^2}\left[S^2 J_{ab} + K_{ab} - S\left(L_{ab} + L_{ba}\right) \right]$$

where a and b are the orbitals for molecules on A and B, respectively. This readily gives the interaction energy between two helium atoms by setting $Z_A = Z_B = 2$.

A more recent derivation of the HL energy of a system of interacting molecules offers a simplified expression of the total overlap repulsion energy at short-range (Hazma and Mayer, 2001a). The simplification is due to the application of Löwdin's pairing theorem (Amos and Hall, 1961; Löwdin, 1962; Mayer, 1997). Application of this theorem allows for a transformation of the orbitals within the two interacting molecules A and B that retains the orthogonal nature of the orbitals

$$\left\langle a_i \middle| a_j \right\rangle = \left\langle b_i \middle| b_j \right\rangle = \delta_{ij} \quad (11.19)$$

while pairing one orbital in A with one orbital in B

$$\left\langle a_i \middle| b_j \right\rangle = S_j \delta_{ij} \quad (11.20)$$

where a and b within Equations (11.19) and (11.20) represent transformed orbitals, and S represents the overlap integral, similar to Equation (7.23). This allows for a great simplification in the computation of the energy of interaction. In an interaction where the

molecules each have 10 orbitals, one may only consider effects depending on the overlap of 10 paired orbitals instead of the original $10^2 = 100$ pairings.

These transformed wavefunctions are then applied to an expression yielding the HL energy. The resulting interaction energy is separated into physically meaningful terms

$$u = u_{electrostatic} + u_{exchange} + u_S \tag{11.21}$$

where the first two terms are the classical electrostatic term and the exchange term, and the final term holds all the effects evolved from the overlap of the wavefunctions. The overlap effects can by further separated

$$u_S = u_{S,basis} + u_{S,intra} + u_{S,el/ex} + u_{S,direct} \tag{11.22}$$

where the first term arises from a description using an incomplete basis set (similar to BSSE), the second term is the overlap effects on the intramolecular energy, the third term holds the effect that electron overlap has on the electrostatic and exchange terms, and the final and most significant term hold the so-called direct overlap effects that dominate the short-range repulsion.

A work that compares the above interaction energy terms to computations of hydrogen-bonding systems shows that the interaction of one pair of orbitals, call them orbital pair k, contributes to the interaction energy an order of magnitude more than the other orbital pairs (Hazma and Mayer, 2001b). Therefore, it is shown that the overlap of the kth pair of orbitals is the only significant overlap. The expressions presented in the original paper are modified mathematically by applying

$$S_k = S_i \delta_{ik} \tag{11.23}$$

thus further simplifying the expressions. Within the following two-electron integral statements, an orbital name without a subscript is assumed to be the kth orbital, and all overlap integrals $S = S_k$ in (11.23). The finite basis contribution is given by

$$u_{S,basis} = \frac{2S^2}{1-S^2}\left(\left\langle a\left|\hat{f}_A\right|a\right\rangle + \left\langle b\left|\hat{f}_B\right|b\right\rangle\right) - \frac{2S}{1-S^2}\left(\left\langle b\left|\hat{f}_A\right|a\right\rangle + \left\langle a\left|\hat{f}_B\right|b\right\rangle\right) \qquad (11.24)$$

where $\hat{f}_i$ is the Fockian of molecule i. The intramolecular effect is given by

$$u_{S,intra} = \frac{S^4}{1-S^2}\left[\left(aa|aa\right) + \left(bb|bb\right)\right] \qquad (11.25)$$

The effect of small intermolecular overlap on the original wavefunctions is said to be insignificant due to the S^4 dependence of Equation (11.25). The overlap effects on the exchange and the electrostatic contribution are given by

$$u_{S,el/ex} = \frac{2S^2}{1-S^2}\left(\sum_{i\in A}\left\langle b\left|-Z_i/r_i\right|b\right\rangle + \sum_{j\in B}\left\langle a\left|-Z_j/r_j\right|a\right\rangle\right)$$
$$+ \frac{2S^2}{\left(1-S^2\right)^2}\left(2J_{ab} - K_{ab}\right) \qquad (11.26)$$
$$+ \frac{2S^2}{1-S^2}\left[\begin{array}{c} \sum\limits_{i=1}^{n_e(\text{molecule }A)}\left(2J_{a_ib_k} - K_{a_ib_k}\right) \\ + \sum\limits_{i=1}^{n_e(\text{molecule }B)}\left(2J_{a_kb_i} - K_{a_kb_i}\right) \end{array}\right]$$

The simplification of the direct overlap contribution is given by

$$u_{S,direct} = \frac{2S}{1-S^2}\left(\sum_{i\in A}\langle a|-Z_i/r_i|b\rangle + \sum_{j\in B}\langle b|-Z_j/r_j|a\rangle\right)$$

$$+\frac{2S^2}{(1-S^2)^2}(3K_{ab}-J_{ab})$$

$$-\frac{2S}{1-S^2}\left\{\begin{array}{l}\displaystyle\sum_{i=1}^{n_e(\text{molecule }A)}\left[2(a_ib_k|a_ia_k)-(a_ib_k|a_ka_i)\right]\\[4mm]\displaystyle+\sum_{i=1}^{n_e(\text{molecule }B)}\left[2(b_ib_k|b_ia_k)-(b_ib_k|a_kb_i)\right]\end{array}\right\}$$

$$-\frac{4S^3}{(1-S^2)^2}(L_{ab}+L_{ba})$$

(11.27)

The calculations assuming the significant contribution from one overlapping pair give an interaction energy that is different from the original interaction energy by less than one millihartree over all interaction distances.

At near ambient temperatures, the interaction energy between closed-shell molecules results from a relatively small overlap in the electron density, as encountered in the numerous theoretical and computational studies of intermolecular interactions. Also, at distances where this overlap occurs, the electron density begins to resemble a spherical wavefunction that decays exponentially with distance. These concepts have been combined to assume that intermolecular interaction overlaps are describable by the interaction of two spherically symmetric Gaussian-type wavefunctions (Jensen, 1996), similar to the SGA wavefunctions of section 10.7. The spherical Gaussian overlap (SGO) approximation considers the interaction of two $1s$ GTF orbitals, both with the same orbital exponent ξ. The overlap of these two orbitals is given by Equation (7.23) and is simplified when using equal orbital exponents and the normalization constant

$$S_{ij} = e^{\frac{-\xi}{2}r_{ij}^2}$$

(11.28)

The orbital exponent is therefore found by rearranging this expression

$$\xi = -\frac{2}{r_{ij}^2} \ln S_{ij} \tag{11.29}$$

This effective orbital exponent varies with distance and with the actual exponents of the orbitals, implicit within S_{ij}. For instance, if the exchange integral is calculated for spherical Gaussians with the same exponent, Equation (7.26) is used

$$K_{ij} = 2\sqrt{\frac{\xi}{\pi}} e^{-\xi r_{ij}^2} \tag{11.30}$$

When the SGO approximation is applied, inserting Equation (11.29) into Equation (11.30) yields

$$K_{ij} = 2\sqrt{\frac{-2\ln S_{ij}}{\pi}} \frac{S_{ij}^2}{r_{ij}} \tag{11.31}$$

This approximation yields two-electron integrals as functions of electron density overlap, which can be found within supermolecule calculations. The SGO approximation has also been applied to electron-electron and electron-nucleus interactions, resulting in similar simplifications of the one- and two-electron integrals (Kairys and Jensen, 1999).

11.5 Hydrogen Bonding

Hydrogen-bonding interactions are long-lived, attractive interactions between a highly electronegative atom in one molecule and an exposed hydrogen atom within the other. The interaction energies are usually an order of magnitude greater than non-bonded interactions between neutral molecules. The interaction distances are small, and complexes of many molecules hydrogen-bonded to one another have a profound effect on the macroscopic system thermodynamics of a fluid.

The role of hydrogen bonding within biological systems is very important. Deoxyribonucleic Acid (DNA), the fundamental molecule holding the genetic code of biological species, is held together by hydrogen bonding. Biomolecule-receptor complexes are important for the transfer of important material into and out of cells. Multiple hydrogen-bonding sites on molecules, as well as their relative positioning, yield a lock and key relationship to the cell in this case.

Hydrogen bonding in the past has been determined experimentally (Curtiss, et al. 1979), while more recently theoretical and computational studies have been undertaken (Estrin, et al. 1996; Kairys and Jensen, 1999; Hazma and Mayer, 2001b; Bartha, et al. 2003; Kozmutza, et al. 2003; Kryachko and Scheiner, 2004). Classically, such interactions have included an electronegative atom interacting with a hydrogen atom that has a large positive partial charge, while more recently studies have included systems where the hydrogen atom has a near neutral charge, as in methane.

Comparisons have been made between the experimental hydrogen-bonding interaction energy ($\varepsilon = -2740 \pm 350$ K) and oxygen-hydrogen interaction distance ($r_{OH} \simeq 2.01$ Å) (Curtiss, et al. 1979) with *ab initio* and DFT methods for the water dimer (Estrin, et al. 1996). This interaction is depicted within Figure 11.1. The computational results show that theoretical methods predict interaction energies within experimental error, with the *ab initio* methods tending to be more accurate. Interaction distances are predicted better with the DFT methods.

Simplifications have been attempted using the theoretical approximations at short-range (Hazma and Mayer, 2000b). Within that work, it is assumed that the short-range interaction is dominated by one orbital on each of the water molecules. The topological
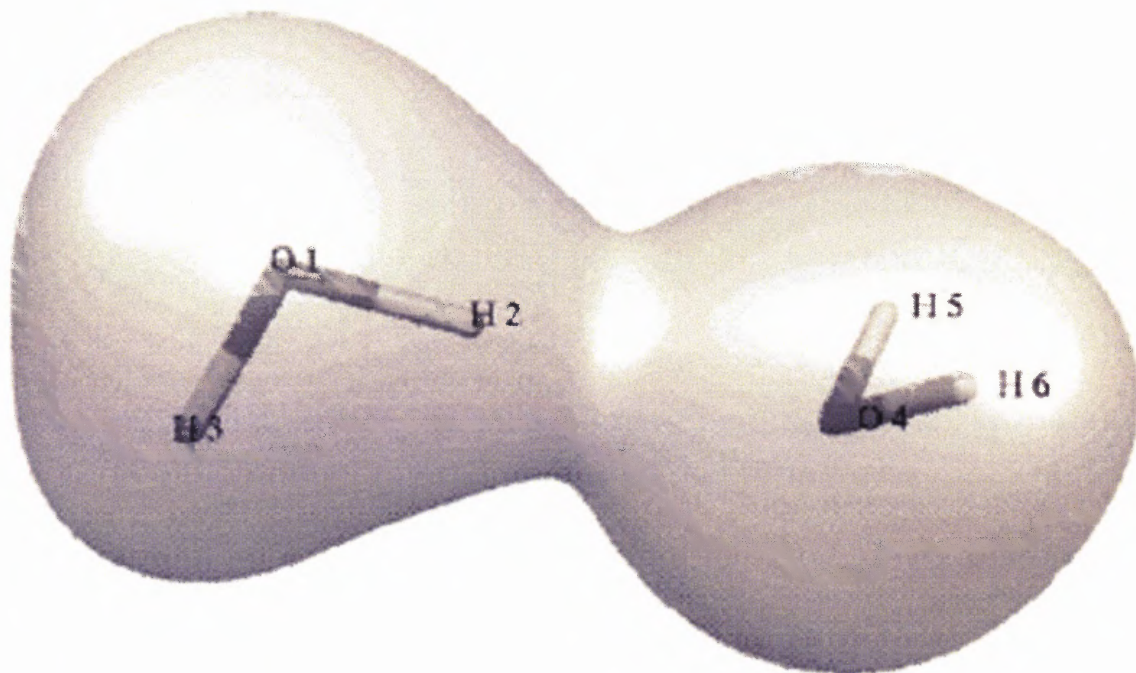
Figure 11.1   Depiction of a hydrogen-bonding interaction between water molecules (Bartha, et al. 2003). The surface depicted is the 0.01 au isodensity surface.

depiction of this assumption is shown in Figure 11.2. The result using this method predicts a hydrogen-bonding interaction energy of $\varepsilon = -2054$ K at an oxygen atom-hydrogen atom interaction distance of $r_{OH} = 2.2$ Å .

11.6 Applications of Potential Functions

The energies of interaction that molecules experience have a direct impact on the thermodynamics of the macroscopic system. If it were the case that molecules did not interact, or only interacted as hard spheres, then all matter would exist as ideal gases with no attractive forces to condense groups of molecules into liquid and solid phases. The existence of these attractive forces is the main cause of this condensation. The magnitude
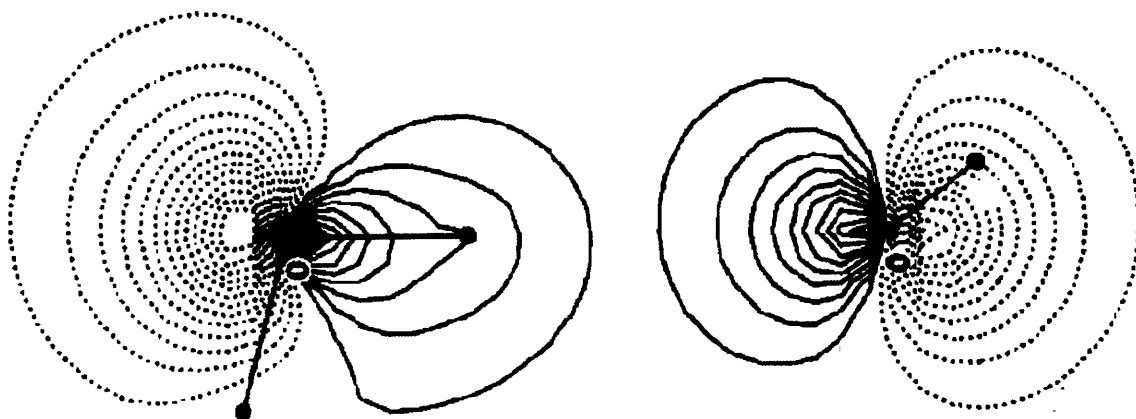
Figure 11.2 Orbitals within water molecules after application of Löwdin's pairing theorem (Hazma and Mayer, 2001b). The outermost solid line is the 0.05 au isodensity surface.

of these forces at the molecular level, as well as orientation effects, causes species to behave very differently in condensed phases. These differences grow out of the difference within the electronic structure at the molecular level.

A direct measure of how multiple-body interaction effects at the molecular level dictate the macroscopic thermodynamics at the system level is shown in the virial equation of state, Equation (3.38). If binary interaction potentials presented above are inserted into Equation (3.39), the theoretical result can be compared to second virial coefficient data of real fluids.

The study of potential energy functions has grown as the need for them in molecular simulation studies has grown. Specifically, site-site potentials are used to

approximate the whole interaction between molecules as a set of interactions between the sites of the different molecules. Not only does the interaction function need to be of proper form, but they should be computationally efficient for use in large-scale computational methods.

A vast majority of the contemporary force fields used in molecular dynamics and Monte Carlo simulations, such as Chemistry at Harvard Molecular Mechanics (CHARMM) (Brooks, et al. 1983), Optimized Potentials for Liquid Simulations (OPLS) (Jorgensen, et al. 1984), Transferable Potentials for Phase Equilibrium (TraPPE) (Martin and Siepmann, 1998), and Anisotropic United-Atoms (AUA) (Ungerer, et al. 2000), utilize the LJ potential of Equation (11.15) with the inclusion of a Coulombic interaction term

$$u^{LJ}\left(r_{ij}\right) = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{r_{ij}} \tag{11.32}$$

where this interaction energy corresponds to when sites i and j are a distance r_{ij} apart. The interaction parameters σ_{ij} and ε_{ij} are found from the pure parameters by the use of combining rules, the most common of which are the arithmetic mean

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \tag{11.33}$$

and the geometric mean

$$\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}} \tag{11.34}$$

sometimes referred to jointly as the Lorentz-Berthelot combining rules. The partial charges are found by one of the theoretical or semi-empirical routines considered in Section 7.3. Most of the major, accessible molecular dynamics programs ask for the site-

site potential parameters for the functional form in Equation (11.32). It is a computationally efficient function, although it exists only as an effective potential that considers the balance between the repulsive short-range effects, the attractive van der Waals effects and the Coulombic interaction from the effective partial charges.

Up to this point, there is no straightforward way of determining parameters for Equation (11.32) in molecular dynamics simulations. It is common to set parameters for groups within linear alkanes before extending the model parameters to groups with heavy atoms other than carbon. Within OPLS, the alkyl group parameters are optimized to reproduce the densities and heats of vaporizations of 15 liquids. Within TraPPE, these parameters are found to reproduce critical temperatures and saturated liquid densities for alkanes between methane and dodecane. For AUA, which contains another parameter that moves the interaction center off the nucleus of the heavy atom, the parameters are optimized by reproducing the vapor pressures, heats of vaporization and liquid densities of ethane, pentane and dodecane. Further parameters for non-alkyl functional groups are found in a trial-and-error manner to reproduce system properties like those above. Table 11.1 lists the alkyl group LJ parameters for the three methods described above, and expresses how different fitting criteria can alter the functional group properties. According to J. I. Siepmann (personal communication, August 25, 2004), methods that employ statistical thermodynamics would aid in this process since they may offer initial guesses to the parameters.

Some criticism of the LJ functional form has arisen. As stated above, the repulsive contribution is better represented as a decaying exponential, rather than the arbitrary r^{-12} contribution (Knowles and Meath, 1986a). This reasoning has lead to the

Table 11.1 Lennard-Jones Parameters for Alkyl Groups within Molecular Dynamics Force Fields (Jorgensen, et al. 1984; Martin and Siepmann, 1998; Ungerer, et al. 2000)

	CH$_3$			CH$_2$			CH		C	
	OPLS*	AUA	TraPPE	OPLS	AUA	TraPPE	OPLS	TraPPE	OPLS	TraPPE
ε/k_B (K)	88.1	120.15	98	59.4	86.291	46	40.3	10	25.2	0.5
σ (Å)	3.905	3.6072	3.75	3.905	3.4612	3.95	3.85	4.68	3.8	6.4

*denotes the use of the CH$_3$ functional group within n-butane

use of the exponential-6 functional form, Equation (11.17). Also noted is that the attractive contribution may not be well represented as a r^{-6} function (Hart and Rappé, 1992). That study considers the overlap of singly-occupied s-type STOs, as in the work of Rosen (1931), and creates a relatively simple potential. It contains a repulsive term similar to the exponential-6 potential, and an attraction term that contains polynomials in interaction distance multiplied by damping functions. Although not nearly as simple a functional form as the empirical potentials, the theoretically-motivated model is shown to be more capable in describing the interactions between small diatomic molecules.

11.7 Summary

Several approaches exist that attempt to model the interactions of closed-shell molecules. These methods range in complexity, from the full consideration of molecular orbitals within perturbation treatments, to a consideration of selected orbitals with significant overlap, to empirical expressions where the details of the interaction are found within correlation parameters. Very specific results can be attained through the use of computational chemistry software. The method one chooses depends entirely on how accurate one wishes to be with predicting interaction energies and distances.

Several key points within this research are carried through to the interaction model developed in this work. Due to the theoretical examinations of the hydrogen-bonding schemes, considering the overlap of one pair of orbitals that obey the SGO approximation simplifies the description of functional group interactions. Also, since the helium dimer interaction model fits best with the information available from AIM theory, namely the SGA and the extent of the electron density, this becomes the basis for the short-range repulsions within this work. This model will not be as simple as the empirical functions used in molecular simulation, but it may be useful as a force field with parameters that are entirely definable using first-principles methods.

CHAPTER 12

MODELING SMALL MOLECULE AND FUNCTIONAL GROUP INTERACTIONS

Interaction theory lays a descriptive reasoning for the nature of intermolecular interactions. The most rigorous solutions involve theoretical frameworks within perturbation theory, usually resulting in expressions that account for overlaps between all the orbitals in one molecule with all the orbitals in the other. Simpler interaction models exist, either by reducing the number of significant overlaps considered in a short-range interaction, or by assuming an empirical form to the potential and fitting parameters to known experimental data. Within the range of expressions considered, an engineering model can be developed with predictive capabilities of the more complicated expressions while wielding relatively simple concepts.

This chapter describes novel work towards an interaction potential energy model usable for small molecules and functional groups. This work begins by applying the concept of the Spherical Gaussian Overlap (SGO) approximation in Section 11.4 to molecules and functional groups whose wavefunctions are approximated by the Spherical Gaussian Approximation (SGA) in Section 10.7. These orbitals are then applied to the helium dimer expression in an attempt to describe short-range interactions. The long-range interactions are described using the Atoms in Molecules (AIM) properties with classical electrostatic interaction expressions from Section 3.4.

Small molecule intermolecular potentials are predicted and compared to second virial coefficients through theory presented in Section 3.3. Interactions between larger molecules are considered as a set of possible functional group interactions and compared

207

to United-Atom (UA) interaction potentials, namely TraPPE. A study is conducted to account for the long-range effects of highly electronegative atoms and groups in these functional group interaction models. Finally, an interaction matrix is assembled toward the prediction of fluid properties using the statistical models offered in Section 5.4.

12.1 Overlap of Spherical Gaussian Wavefunctions

Electron density overlap using the SGO approximation (Jensen, 1996; Kairys and Jensen, 1999) as described in Section 11.4 allows for a simplified representation of the overlap of electron density between two closed-shell molecules at close-range. It has been argued that this is accurate due to the large repulsive energy generated by a relatively small overlap between the electron densities.

Also, as demonstrated by the work of Hazma and Mayer (2001a, 2001b), the interaction between two closed-shell molecules may be rigorously represented by the overlap of one pair of orbitals, given that the proper transformation of the molecular orbitals is conducted.

These works suggest that molecular interactions may be approximated by the overlap of two spherical GTFs, like those presented Section 10.7. Since a small overlap results in a large repulsion, the GTFs need only represent the tails of the electron densities.

Figure 12.1 depicts the approximations made in this work toward a simplification of the interactions between two closed-shell molecules. On the left, the p-type orbitals within the neon atom and methanol molecule are explicit and overlap with the p-type orbital of another neon atom or the s-type orbital of the hydrogen atom with another
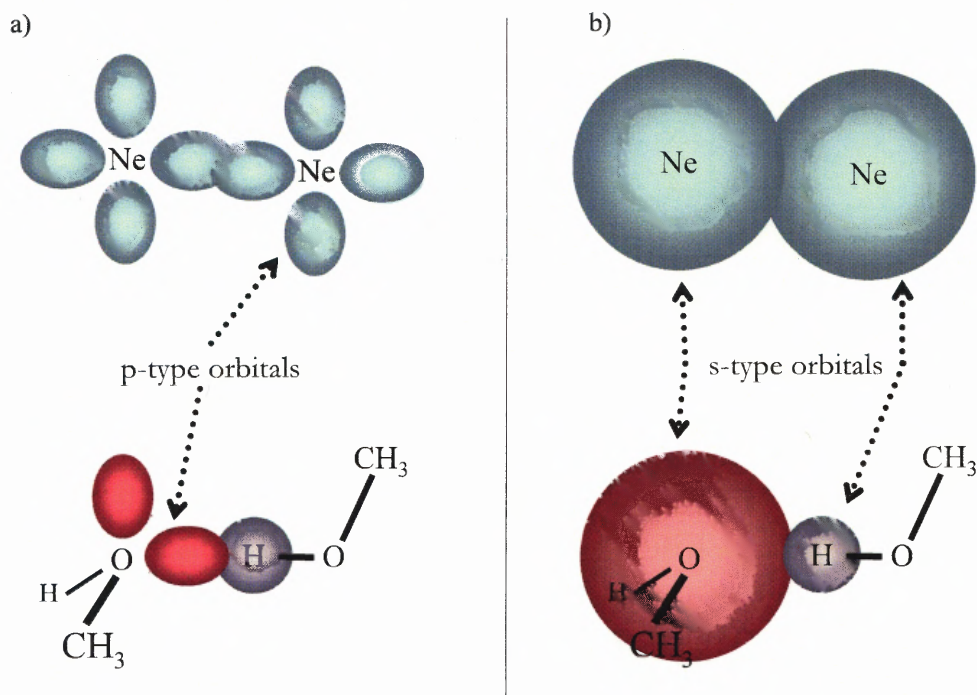
Figure 12.1 Binary interactions between neon atoms and methanol molecules, depicted by real orbital overlap and spherical Gaussian overlap.

methanol molecule. With the SGO approximation, these overlaps can be approximated with those of spherical GTFs, depicted on the right. Of note is the loss of information on the angular nature of the hydrogen-bonding scheme depicted between the two methanol molecules. Considering partial charges placed on the nuclei, an orientation unlike that depicted in Figure 12.1b may be more favorable energetically.

12.2 Full Expression of Dimer Interaction

With SGAs of the interacting functional groups, the interaction model is considered. Within Sections 11.1 and 11.4, several expressions that describe the short-range interaction energy have been reviewed. Within the theoretically rigorous expressions given by perturbation methods (Hayes and Stone, 1984) in Equations (11.6) and (11.7)

and by the application of Löwdin's pairing theorem (Hazma and Mayer, 2001b) in Equations (11.21) through (11.27), information on all the orbitals within the molecule is necessary for an accurate calculation. Given that this work attempts to determine interaction potentials with only one GTF on each interacting entity, the above expressions are not used and may be revisited in future work.

The expression of the helium dimer in Equation (11.18) by Rosen (1931) offers a concise description of the short-range interaction effects evolved from a pair of doubly-occupied orbitals. Also an advantage is the lack of a long-range contribution to the potential, which is to be considered using the classical expressions within Section 3.4.

The integrals involved in Equation (11.18) are given in general, thereby STOs or GTFs may be used. If the latter are used, the quantum integrals result in relatively simple expressions that combine the orbital exponents, ξ_a and ξ_b, and the interaction distance r. Assuming the use of normalized GTFs of the form

$$g_i = \left(\frac{2\xi_i}{\pi} \right)^{3/4} e^{-\xi_i r_i^2} \tag{12.1}$$

The integrals within Equation (11.18) are given as the following:

$$S = e^{\frac{-\xi_a \xi_b}{\xi_a + \xi_b} r^2} \tag{12.2}$$

$$J_{ab} = \frac{1}{r} erf\left(r \sqrt{\frac{2\xi_a \xi_b}{\xi_a + \xi_b}} \right) \tag{12.3}$$

$$K_{ab} = S^2 \sqrt{\frac{2}{\pi}(\xi_a + \xi_b)} \tag{12.4}$$

$$L_{ab} = \frac{S}{r} \left(\frac{\xi_a + \xi_b}{\xi_b} \right) erf\left(r\xi_b \sqrt{\frac{2\xi_a}{(3\xi_a + \xi_b)(\xi_a + \xi_b)}} \right) \tag{12.5}$$

$$L_{ba} = \frac{S}{r}\left(\frac{\xi_a + \xi_b}{\xi_a}\right) erf\left(r\xi_a\sqrt{\frac{2\xi_b}{(\xi_a + 3\xi_b)(\xi_a + \xi_b)}}\right) \tag{12.6}$$

$$\left\langle b \left| \frac{-Z_A}{r_A} \right| b \right\rangle = -Z_A \frac{erf\left(r\sqrt{2\xi_b}\right)}{r} \tag{12.7}$$

$$\left\langle a \left| \frac{-Z_B}{r_B} \right| a \right\rangle = -Z_B \frac{erf\left(r\sqrt{2\xi_a}\right)}{r} \tag{12.8}$$

$$\left\langle a \left| \frac{-Z_A}{r_A} \right| b \right\rangle = Z_A \frac{S}{r}\left(\frac{\xi_a + \xi_b}{\xi_b}\right) erf\left(\frac{\xi_b r}{\sqrt{\xi_a + \xi_b}}\right) \tag{12.9}$$

$$\left\langle a \left| \frac{-Z_B}{r_B} \right| b \right\rangle = Z_B \frac{S}{r}\left(\frac{\xi_a + \xi_b}{\xi_a}\right) erf\left(\frac{\xi_a r}{\sqrt{\xi_a + \xi_b}}\right) \tag{12.10}$$

where *erf* is the error function. Slater-type orbitals may also be used to determine the integrals necessary for Equation (11.18), but they are not used in this work due to the complexity of the resulting two-electron integrals.

The helium dimer interaction model includes the Coulombic effects between nuclei and electrons. The nucleus-nucleus interaction is explicit in Equation (11.18), while the electron-electron interaction is implicit within Equation (12.3) and the nucleus-electron interactions are implicit in Equations (12.7) and (12.8). In this work, this effect is considered through the interactions between partial charges and dipole moments found in AIM theory. Therefore, these effects are removed, and the expression describing short-ranged interactions for this work is given by

$$u_{\text{He}_2} = \frac{4}{1-S^2} \left(J_{ab} + \frac{1}{2} \left\langle b \left| \frac{-Z_A}{r_A} \right| b \right\rangle + \frac{1}{2} \left\langle a \left| \frac{-Z_B}{r_B} \right| a \right\rangle \right)$$

$$- \frac{4}{1-S^2} \left[K_{ab} - S \left(\frac{1}{2} \left\langle a \left| \frac{-Z_A}{r_A} \right| b \right\rangle + \frac{1}{2} \left\langle a \left| \frac{-Z_B}{r_B} \right| b \right\rangle \right) \right]$$

$$+ \frac{1}{\left(1-S^2\right)^2} \left[S^2 J_{ab} + K_{ab} - S \left(L_{ab} + L_{ba} \right) \right]$$

$$- \left(\frac{4}{r} - \frac{2Z_A + 2Z_B}{r} \right) \tag{12.11}$$

This expression serves as the short-range exchange-repulsion term in interactions between closed-shell molecules. Its effectiveness is considered in the intermolecular potential function between small molecules and molecules where hydrogen bonding plays an important role.

12.3 Small Molecule Interactions

The statistical mechanics presented within Chapter 3, and more specifically the virial equation of state within Section 3.3, offers a window between the behavior of interacting molecules and the experimentally determinable virial coefficients. Any model that attempts to predict small molecule interactions can be assessed by how well it predicts second virial coefficients using Equation (3.39).

The model to describe interactions between small molecules is given by Equation (3.54), where the exchange-repulsion contribution is considered through the helium dimer expression in Equation (12.11), and where the long-ranged electrostatic effects are given by the averaged expressions in Section 3.4. In total, the small molecule interaction model is given by

$$u = u_{\text{He}_2} + \overline{u}_{\mu_1 \mu_2} + \left(\overline{u}_{\mu_1 \alpha_2} + \overline{u}_{\mu_2 \alpha_1} \right) + u_{\alpha_1 \alpha_2} \tag{12.12}$$

No pure Coulombic interaction is necessary here, since the molecules are neutral and no partial charge distributions are considered explicitly. To make sure the validity of the helium dimer interaction is the focus of the analysis, the experimental values for the dipole moment μ and the polarizability α are used instead of the calculated values (presented in Appendix E).

To use the helium dimer interaction, the molecule must be approximated as the SGA around a nuclear center. Figure 12.2 outlines the process for the nitrogen molecule, a spherically asymmetric molecule. First, the AIM volume is calculated for the entire molecule, which is given by the sum of its constituent functional groups within Appendix G and is depicted in Figure 12.2a. Next, this volume is used to create an SGA, where the 0.001 au surface of the SGA and the bounding surface of the molecule create the same
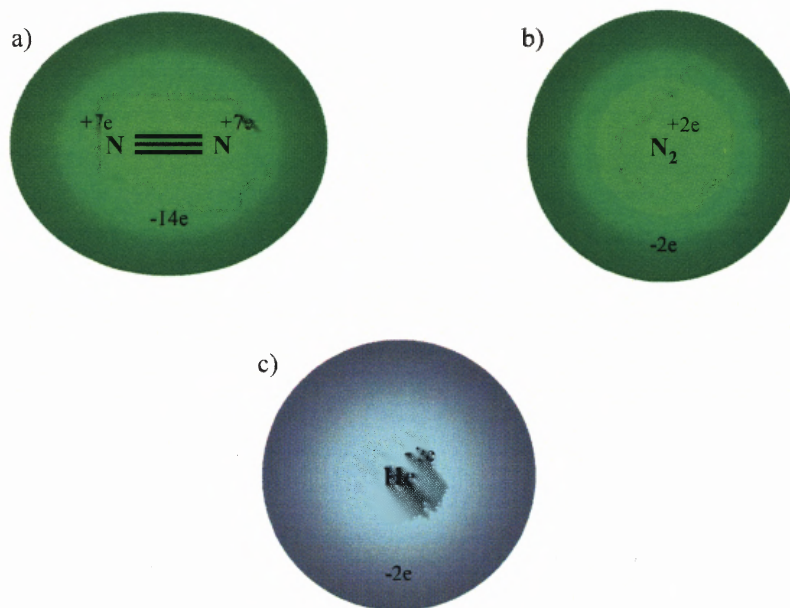
Figure 12.2 Algorithm followed to allow for a nonspherical small molecule to be used within the helium dimer interaction model.

volume, which is shown in Figure 12.2b. Then this orbital is assumed to be surrounding a helium nucleus, thereby allowing for the use of Equation (12.11) with nuclear centers of charge $+2$. How the SGA corresponds to the bounding surface of methane and nitrogen can be seen in Figures 10.10 and 10.11.

Figure 12.3 shows the resulting interaction potential model between two methane molecules after applying the procedure described above. Both a curve with the experimental μ and α (in blue) and with calculated μ and α (in red) are shown to reveal slightly different interaction wells. They also show the characteristic repulsive walls at short distances, the interaction well where energies are favorable, and the tail where the interaction energy vanishes as the interaction distance increases. Not shown in this figure is the interaction energy function at distances around 1 Å. There,
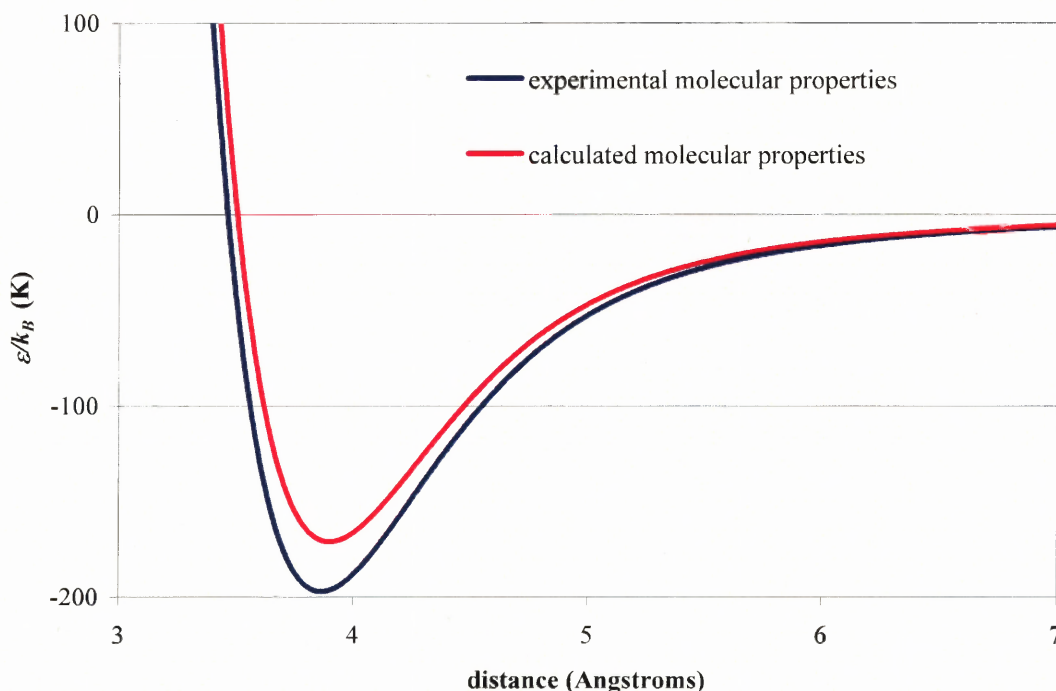
Figure 12.3 Binary interaction energy for the methane dimer system.

the SGO approximation breaks down, as an unrealistic negative interaction energy results. Since these interaction distances are unattainable at the thermodynamic conditions considered in this work, the results at these extremely short ranges are ignored. Fixes applied to the exponential-6 interaction potential presented at the end of Section 11.3 are needed here if this interaction potential were to be applied numerically.

With analytical interaction curves such as that presented for methane, the second virial coefficient is able to be calculated for any nearly spherically symmetric molecule that has undergone an AIM analysis. Equation (3.39) is used for this calculation, with a modification made to alleviate the problem of the interaction model at extremely short interaction distances. Therefore, the modified formula to determine the virial coefficient is given by

$$B(T) = -2\pi \left[-\frac{r_{\text{cutoff}}^3}{3} + \int_{r_{\text{cutoff}}}^{\infty} \left(e^{-u(r_{12},T)/k_B T} - 1 \right) r_{12}^2 dr_{12} \right]$$

(12.13)

where r_{cutoff} is chosen to correspond to a very large repulsive interaction energy. Also, since the statistical average for the dipole-dipole interaction is utilized, the temperature dependence in this term is considered and must be included within the functionality of u.

The calculations of the second virial coefficients for seven small molecules and how they compare to experiment is presented in Figure 12.4. The upper figure presents the species, namely methane, nitrogen and fluorine, whose intermolecular potential is represented well enough to give a reasonable description of the second virial coefficient. The lower figure shows poor predictions for the species carbon dioxide, nitrous oxide, carbon monoxide and oxygen, although the model performs somewhat well for oxygen.
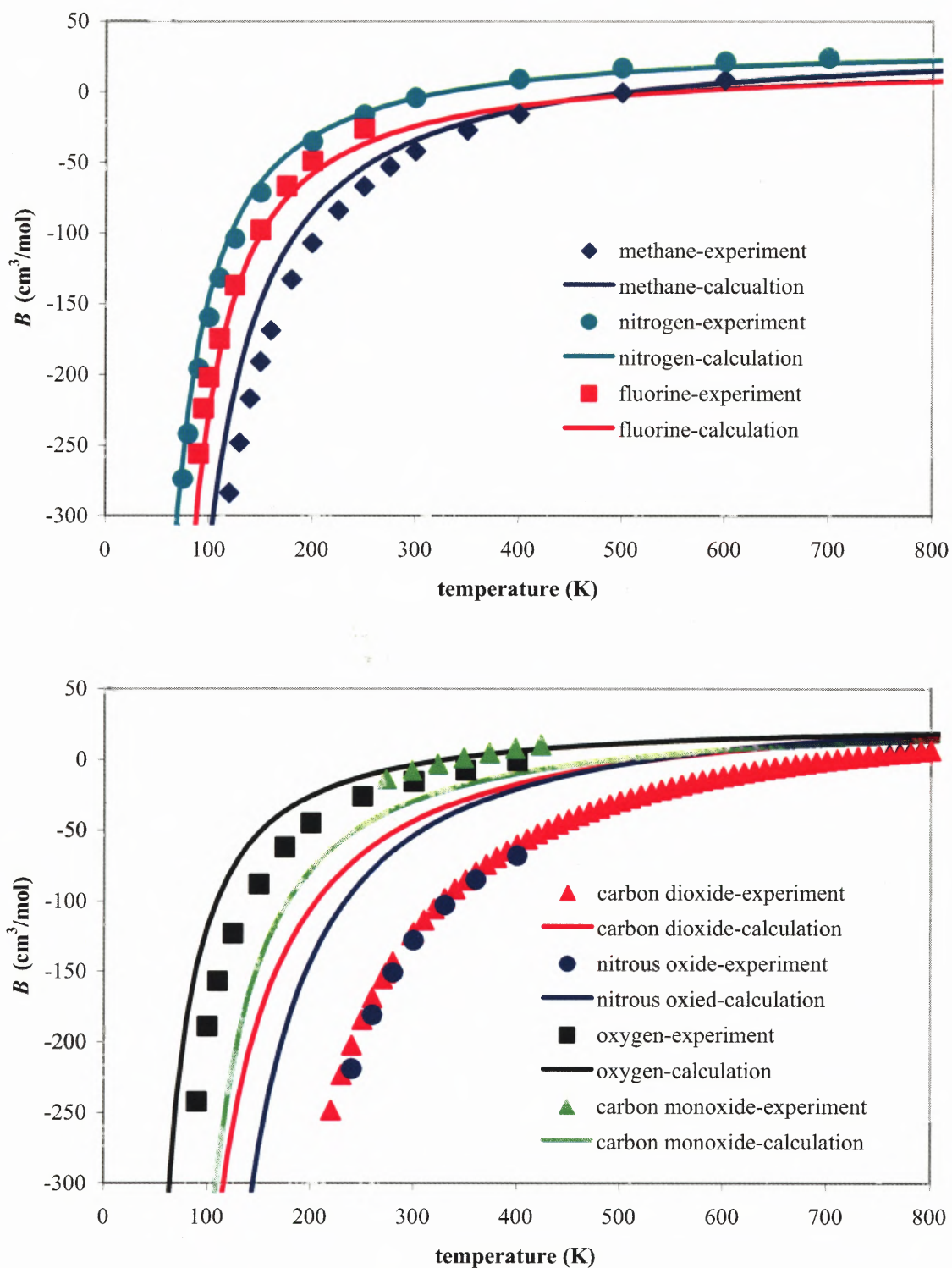
Figure 12.4 Calculated second virial coefficients for small molecules compared to experiment.

A predictive model such as the one outlined in this work simplifies the description of intermolecular interactions by eliminating assumptions about cross interactions. In presently used empirical potentials, arbitrary combining rules such as those within Equations (11.33) and (11.34) are used to determine model parameters for interactions between unlike species. Figure 12.5 shows a close up of the predicted interaction curves involving methane and nitrogen. As suggested by the arithmetic rule of Equation (11.33), the point at which the cross interaction crosses the x-axis (corresponding to σ_{ij}) should lie exactly between the points for the homogeneous interactions. In the predicted model, this is not the case. Also, the geometric rule for ε suggests a deeper interaction well for the cross interaction than is predicted. A model such as this offers a simpler description by eliminating the need for combining rules.
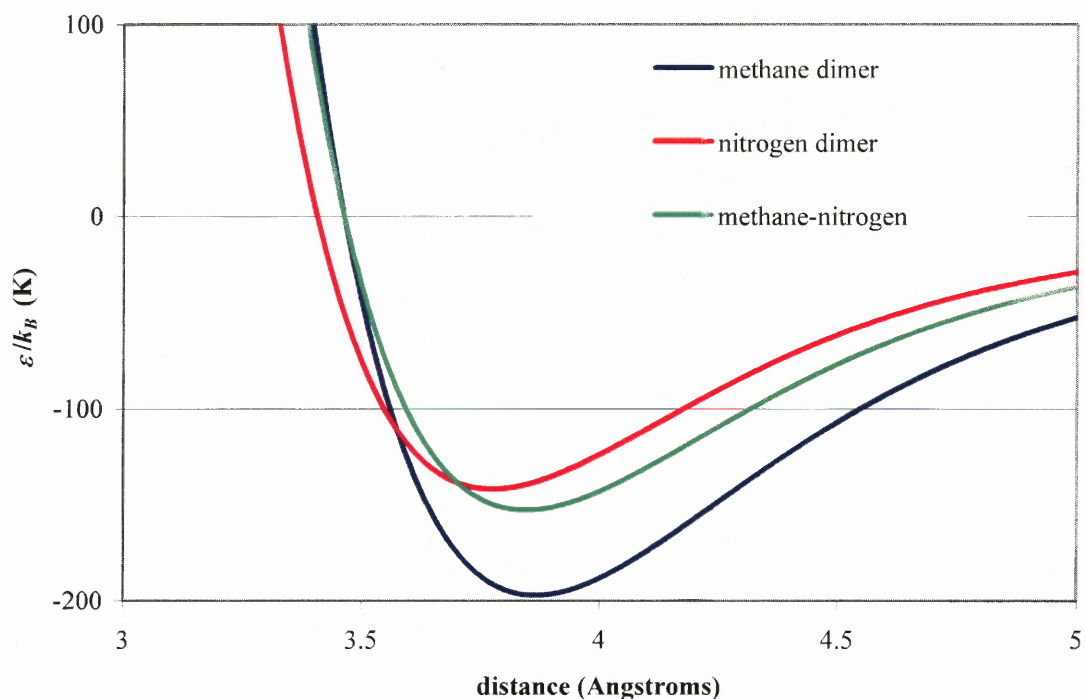
Figure 12.5 Predicted interaction energies for methane and nitrogen.

The sporadic results in the predictions of second virial coefficients suggest that the interaction model within Equation (12.12) offers too simple of a description of the nature of interactions between small molecules. Higher-order attraction effects may be necessary, such as the quadrupole moments or effects that lead to terms that are not dependent on r^{-6} (Hart and Rappé, 1992). Also, the use of STOs and p-type functions within Equation (12.11) would offer a description of short-range effects that resemble real systems more closely.

12.4 Large Molecule Interactions through Functional Group Interactions

The interactions between large molecules can not be analyzed in as straightforward a manner as the interactions between small molecules. The approximation that the molecule is spherically symmetric is not a good assumption for non-cyclic molecules with more than three heavy atoms. The electronegativities of particular atoms within the larger molecules may result in a distribution of electron density that, for modeling interactions, is best describable through partial charges rather than the overall molecular dipole moment. Steric effects play a role in interactions also, since not all portions of the molecule are available to electron density overlap.

The role of functional groups within the analysis of interactions between large molecules is to offer an entity the size of a small molecule that serves as the interaction site. The functional group becomes the interacting entity, and its group properties are the main properties considered in the interaction model. The remainder of the molecule and its set of group properties also contribute to the interaction, but in a way that is secondary to the interacting functional groups. Steric effects are considered in functional group

interactions, since a group is available to interact only if it has the available external surface area. Through functional groups, the interaction between two large molecules can be reduced into a set of functional group interactions.

The functional group interaction model is similar to that of the small molecule interaction model in Equation (12.12), except that the non-averaged electrostatic expressions are utilized. The functional groups within a molecule do not have the freedom to rotate as entire molecules do in the gas phase, therefore the direction of the dipole moment (given by the sign of μ in Appendix G) is also important to the kinds of interactions functional groups have. The interaction model between functional groups is describable through the following expression

$$
\begin{aligned}
u = u_{\text{He}_2} + u_{qq} + \left(u_{q_1\mu_2} + u_{q_2\mu_1} \right) + u_{\mu_1\mu_2} \\
+ \left(u_{q_1\alpha_2} + u_{q_2\alpha_1} \right) + \left(u_{\mu_1\alpha_2} + u_{\mu_2\alpha_1} \right) + u_{\alpha_1\alpha_2}
\end{aligned}
\tag{12.14}
$$

Again, the helium dimer interaction serves as the short-range repulsion, while the Coulombic interaction term is included so as to characterize the partial charge distribution of the molecule for that particular orientation. As is seen shortly, the Coulombic effects from surrounding functional groups play an important part in determining the interaction characteristics between the molecules at the specified orientation.

Like the analysis of the small molecule interaction, assumptions must be made to achieve a spherically symmetric description of the electron density of a functional group. Figure 12.6 demonstrates the algorithm followed to accomplish this. Initially, the functional group, a methyl group in Figure 12.6a, is reduced to the UA in Figure 12.6b. Then, the SGA is used to reproduce the distance between the center of the UA and the
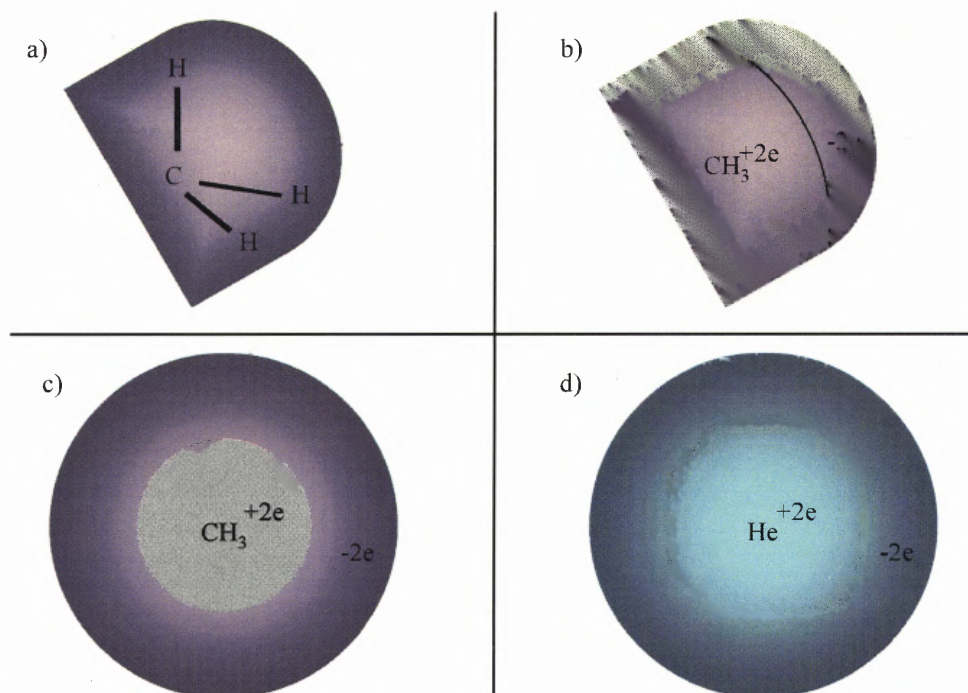
Figure 12.6 Algorithm followed to allow for a nonspherical functional group to be used within the helium dimer interaction model.

bounding surface. This spherical orbital, depicted in Figure 12.6c, reproduces this distance instead of reproducing the volume of the functional group. As before, this spherical functional group then takes on the nucleus of a helium atom, so as to use a nuclear charge of +2 in Equation (12.11). Figures 10.12 and 10.13 demonstrate how the SGA reproduces the bounding surface of alkyl functional groups.

The first set of interactions considered in this work is that between two propane molecules. The functional groups within propane molecules have relatively small partial charges, sizeable dipole moments pointing inward, and few enough functional groups that orientations between the two molecules can be considered concisely. For this study, three types of interactions exist: methyl-methyl; methyl-methylene; methylene-methylene. Figure 12.7 presents the three interactions with four models that attempt to describe the

interaction energy: a model based on GCM concepts, where only the properties of the functional group contribute to the interaction; a GCM model including the partial charges on the remainder of the molecule; a GCM model including partial charges and dipole moments on the remainder of the molecule; and the TraPPE model, described in Equation (11.32), where the LJ parameters are correlated to reproduce phase equilibrium, and where no partial charges are attributed to the groups.

Figure 12.7 suggests that the most general description of the functional group interactions cannot be reproduced appropriately by considering simple models that ignore the complete Coulombic contribution from surrounding functional groups. These long-range Coulombic effects allow for more favorable interactions, appropriate considering propane is a non-polar molecule that should interact somewhat favorably due to van der Waals forces. The AIM partial charges on the functional groups, however small, and the dipole moments cause the molecules to have unfavorable or neutral interactions if only the GCM model is considered. The inclusion of other partial charges and opposite-pointing dipole moments in the molecule causes a modification in this original interaction, thereby resulting in the more favorable interaction regimes.

The interactions given by the TraPPE potential assumes an attraction regime in all the group interactions in propane through the use of LJ potentials and no partial charges. The TraPPE potential would not be able to produce a cross interaction like the full model in this work, since the combining rule for the well depths must result in a cross interaction depth that is between the pure interaction depths. The cross interaction well depth of the model in this work is deeper than either of the interactions, mainly due to the
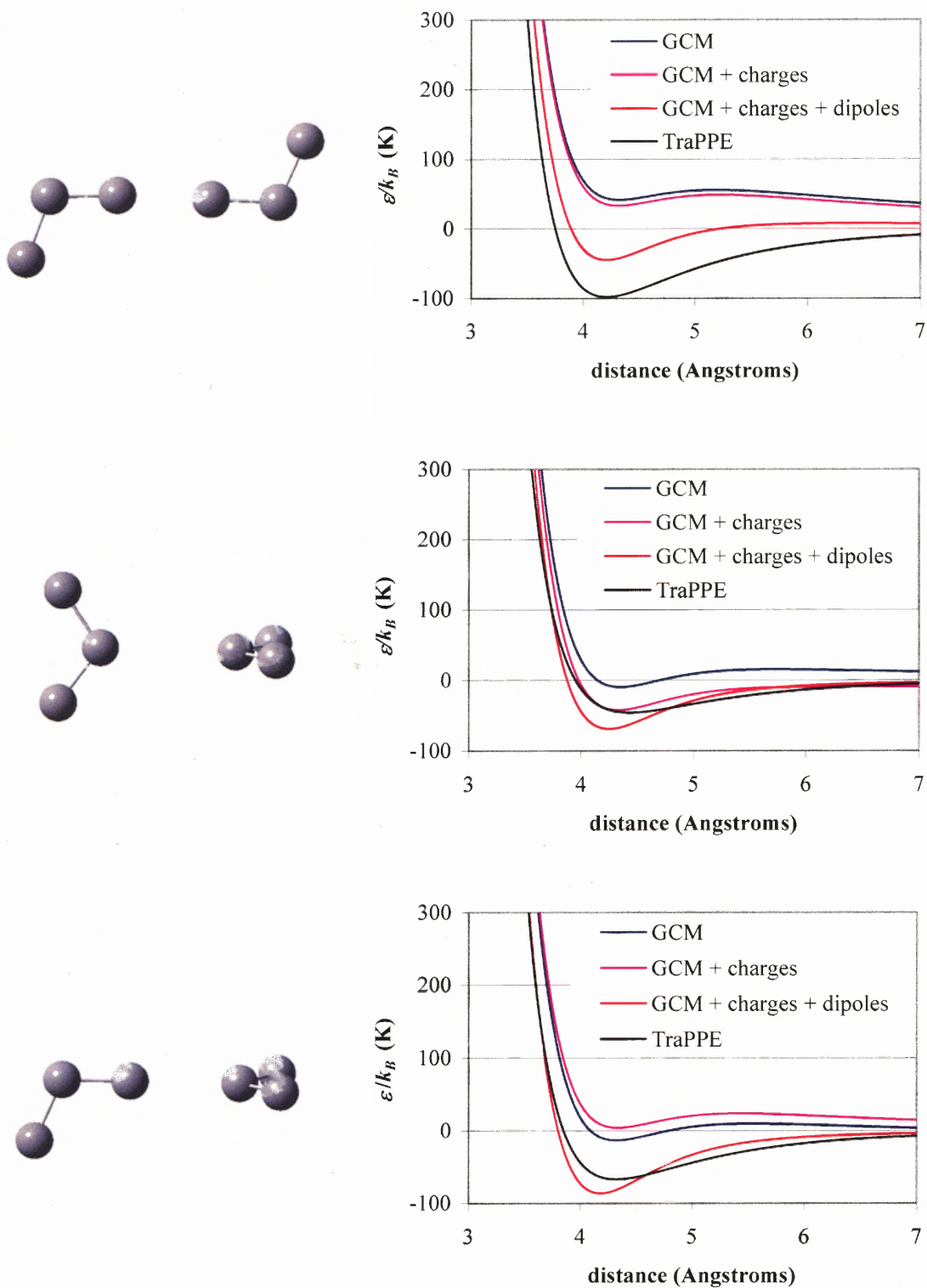
Figure 12.7 Alkyl group interactions between propane molecules (methyl-methyl, methylene-methylene, and methyl-methylene).

opposing partial charges of the functional groups, and thus a strongly favorable interaction scheme from the Coulombic effect.

Interactions between water molecules are well-studied, as presented in Section 11.5, and the hydrogen-bonding scheme between two water molecules serves as a good measure at how well an interaction model performs. Figure 12.8 shows the interaction potentials between the hydrogen atom and oxygen atom using AIM properties and the interaction model presented in Equation (12.14). Here, it is quite clear that the partial charges and the dipoles within the remainder of the molecule are necessary to predict an attraction between the two molecules. This is plausible, given that all the information about the electrostatics within the water molecule is necessary to predict the molecular dipole moment (discussed in Section 10.6). The interaction model used by the TraPPE force field is also presented, assuming that the partial charge on the hydrogen atom
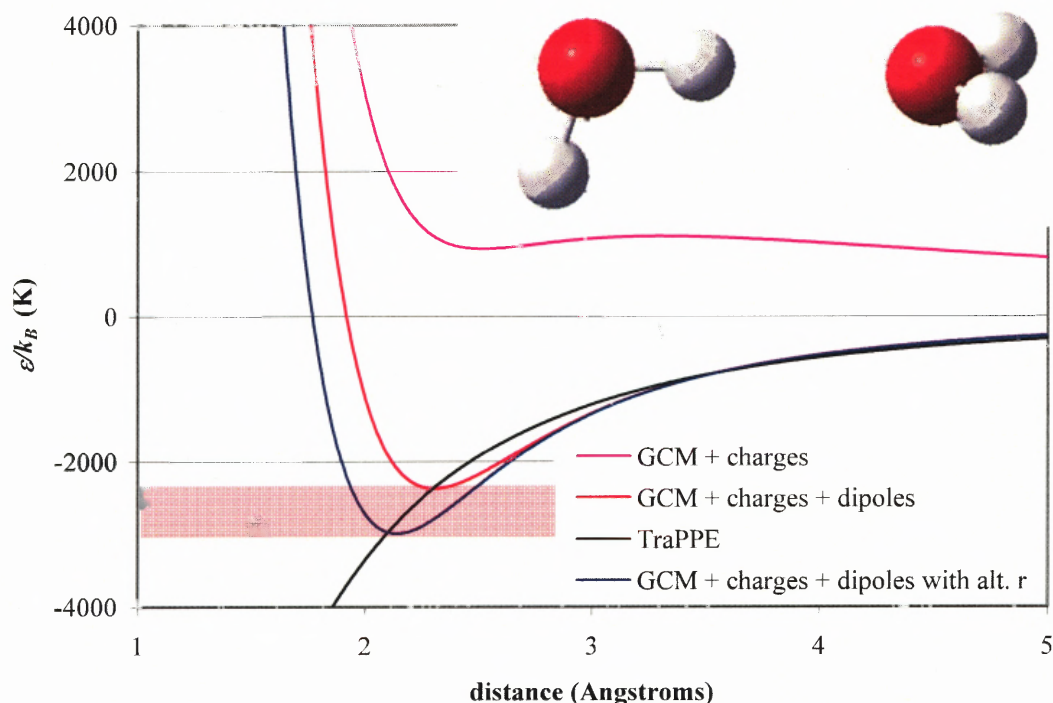
Figure 12.8 Hydrogen-bonding interaction energies between water molecules.

within an alcohol is the same as the partial charge within a water molecule. The lack of a repulsive wall is an artifact of the model description for hydrogen.

The sensitivity of the hydrogen-bonding scheme between water molecules is also presented in Figure 12.8. Here, instead of the Gaussian coefficient corresponding to r_{avg} in the AIM property table, a Gaussian coefficient is used that corresponds to the actual extent of the electron density along the interaction axis. These new values $(r(H) = 2.32 \text{ au}, r(O) = 3.38 \text{ au})$ correspond to slightly different exponents on the GTF $(\xi(H) = 0.563, \xi(O) = 0.196)$. These alternate values cause a slightly less repulsive short-range interaction and result in a hydrogen-bonding distance closer to experimental values. The interaction well is also deeper by about 25% of the value predicted using r_{avg}.

12.5 Interaction Matrices

To account for the possible interactions between two molecules, one may begin to reduce the problem by finding all the possible interactions between the functional groups in one and the functional groups in the other. This would result in a matrix of interaction curves that describe a finite number of possible intermolecular interactions. This is exactly the matrix of quantities that engineering models in Section 5.2 use in excess Gibbs relations in the mixture of fluids. The functional group interactions of this work are applicable within such engineering models.

Problems arise in the creation of such a matrix from the interaction model of this work. The analysis of the interaction model for functional groups, given by Equation (12.14), suggests that to use AIM properties of functional groups to describe interactions between molecules, one must consider all the functional groups in each molecule and

their influence on the interaction. Although this conclusion is directly contrary to the classical view within GCMs, where only the interacting groups influence the energy, the GCM approach is much more computationally efficient than the approach derived in this work. Due to the necessity of including the long-range Coulombic effects of non-interacting functional groups, creating a matrix of values for any molecule with more than three functional groups involves accounting for the locations of all functional groups. The computations of such interaction curves are quite involved and necessitate the creation of specific software that can handle minimizing interaction energies while not altering interior angles and bond lengths. No such software has been created to calculate such curves for the interaction model of this work.

A similarly conceived interaction model to that of this work is the UA model, TraPPE. This model considers effective potentials in the form of a LJ interaction equation and a Coulombic interaction term. The Coulombic interaction between functional groups is greatly simplified since partial charges are placed only on selected UAs and since no dipole moments of functional groups are considered. Table 12.1 presents the parameters from the TraPPE potential that are used to determine the interaction energies between functional groups of interest in this work.

Motivated by the study of mixtures between alkanes, alcohols and ethers, an interaction matrix considering the possible functional group interactions in these systems is presented in Table 12.2 and Figure 12.9. The values given within the table represent the energies of interaction either at the minimum of the interaction curve (for favorable interactions), at a distance corresponding to the sum of the r_{avg} values (for repulsive interactions), or at a

Table 12.1 Parameters from the TraPPE Potential for use in Functional Group Interaction Matrix

LJ parameters	σ (Å)	ε / k_B (K)	q
CH_3 alkyl[a]	3.75	98	
CH_3 alcohol[a]	3.75	98	0.265
CH_3 ether[b]	3.75	98	0.250
CH_2 alkyl[a]	3.95	46	
CH_2 alcohol[a]	3.95	46	0.265
O alcohol[a]	3.02	93	-0.700
O ether[b]	2.8	55	-0.500
H alcohol[a]			0.435

bond lengths	(Å)
CH_x-CH_y[a]	1.54
CH_x-OH[a]	1.43
CH_x-O[b]	1.41
O-H[a]	0.945

bond angle	(deg)
CH_x-CH_2-CH_x[a]	114
CH_x-O-H[a]	108.5
CH_x-O-CH_y[b]	112

[a] Chen et al, 2001.
[b] Stubbs et al, 2004.

Table 12.2 Interaction Matrix for Alkane, Alcohol and Ether Systems (energies in K)

	CH_3 alkyl	CH_3 alcohol	CH_3 ether	CH_2 alkyl	CH_2 alcohol	O alcohol	O ether	H alcohol
CH_3 alkyl	-98.00	-98.00	-98.00	-67.14	-67.14	-95.47	-73.42	0.00
CH_3 alcohol	-98.00	-8.15	-8.84	-67.14	117.3	-627.3	-571.6	232.7
CH_3 ether	-98.00	-8.84	-11.53	-67.14	93.79	-574.5	-522.3	222.1
CH_2 alkyl	-67.14	-67.14	-67.14	-46.00	-46.00	-65.41	-50.30	0.00
CH_2 alcohol	-67.14	117.3	93.79	-46.00	202.3	-715.9	-655.5	353.1
O alcohol	-95.47	-627.3	-574.5	-65.41	-715.9	1031	876.4	-2896
O ether	-73.42	-571.6	-522.3	-50.30	-655.5	876.4	776.4	-2402
H alcohol	0.00	232.7	222.1	0.00	353.1	-2896	-2402	690.5
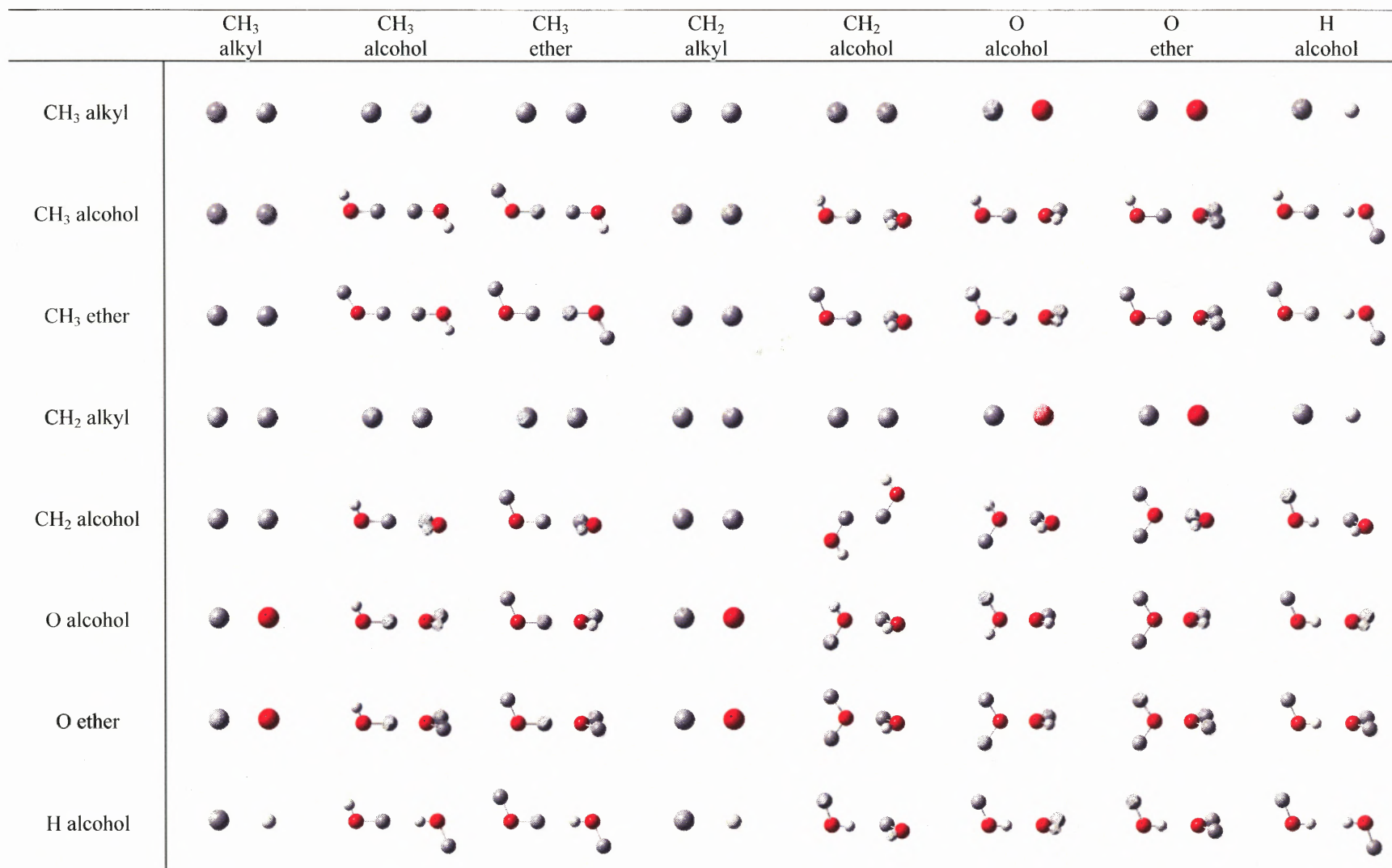
Figure 12.9 Interaction matrix for alkane, alcohol and ether systems: orientation.s

distance corresponding to a hydrogen-bonding interaction ($r = 2.01$ Å for oxygen atom-hydrogen atom interactions). Coulombic effects between interacting and neighboring functional groups are considered only when both interacting functional groups have partial charges. This condition is achieved in all interactions except those including a CH_3 alkyl or CH_2 alkyl functional group. For all others, the partial charges of the nearby groups are considered in the orientations given by the images within Figure 12.9. An attempt was made to choose orientations that would result in the most favorable interactions.

12.6 Conclusions

Information calculated using AIM theory offers rigorous information about molecules and portions of molecules. This information, in turn, is usable within first-principles interaction energy functions between closed-shell molecules. In combination with the helium dimer short-range interaction term and the classical long-range interaction terms, the AIM properties are usable in approximating second virial coefficients of small molecules and functional group interactions within larger molecules. Of particular importance is the prediction of the interaction energy in the water dimer system that is comparable to experimentally measured energies.

The simplicity of the GCM models and the TraPPE potentials make them advantageous. The more rigorous expression using AIM properties within functional group interactions necessitates the use of a model that describes the position and orientation of the entire molecule. Even in the case where the partial charges are near neutral, they have a significant effect on interactions between molecules. The appropriate

tool to account for all the information, namely the magnitude and position of the partial charges with the interacting molecules, is the next hurdle toward the validation of such an approach.

The interaction matrix used in this work is created from TraPPE, an interaction model that contains concepts similar to that of the model developed in this work. The idea of using particular orientations to describe most of the functional group interactions exists in both methods. This methodology will be carried further in this work when a tool to determine interaction energies in various orientations is developed accordingly. Until then, the interaction matrix using TraPPE gives a sense of what is possible when first-principles properties and energies are used within a lattice-fluid model.

CHAPTER 13

PURE SPECIES THERMODYNAMIC BEHAVIOR
AND LATTICE-FLUID THEORY

The use of vacancies within lattice-fluid theory has the interesting prospect of predicting both pure and mixture system behavior. In the past, the most successful models predict either the former (through empirical equations of state) or the latter (though models of lattice fluids that do not include vacancies). This has hindered the development of a fully predictive model for the entire fluid phase, since the different modeling techniques need unrelated correlations and different modeling parameters.

The lattice-fluid models described in Chapters 4 and 5 are shown to describe either the pure state or a mixture system. Since lattice fluids have more established success at predicting mixture behavior, efforts are made to develop their ability to predict volumetric behavior. This chapter develops a pure species lattice-fluid model intended to describe both pure and mixture systems. A more generalized athermal ways expression is applied to the quasi-chemical equation to give a general model for use in any fluid system. A pure species is modeled as a mixture of molecules and vacancies, and an engineering EoS emerges from this study. Parameters are correlated to reproduce the critical point of pure fluids, and the physical significance of these quantities is evaluated. The methodology is extended to the EoS description of larger molecules through the use of the functional group properties and interaction matrix of Section 12.5. Their capabilities are evaluated by how well these models predict the critical point of fluids without any experimental data.

13.1 Use of the Generalized Guggenheim Expression

To determine the volumetric properties of a lattice fluid, one must maximize the partition function given by Equation (4.4) with respect to the number of vacancies, N_0. This is similar to the procedure taken to achieve the equations of state in Section 4.4 and in the general derivation of the quasi-chemical statistics presented in Appendix B. The general expression is given by

$$\left(\frac{\partial \ln \Omega^{ath}}{\partial N_0} \right)_{N_i} - z_0 \ln \Gamma_0 - \frac{pb_0}{k_B T} = 0 \tag{13.1}$$

where the reduced notation within Equation (5.52) is used and where it is assumed that the interaction volume and energy between two vacancies and between a vacancy and a molecule in adjacent lattice sites vanishes.

The Guggenheim expression in Equation (4.9) allows for the inclusion of vacancies on the lattice as if they were a distinct species. The application of this expression yields results that are equivalent to prior work (Smirnova and Victorov, 1987; Panayiotou, 2003b).

An athermal ways expression that is derived using the same procedure as that of Equation (4.9) is offered in Appendix A. This derivation assumes a mixture of two molecular species with vacancies on the lattice, and the result for any number of species is inferred by the new results and the Guggenheim expression presented earlier. The generalized result is given by

$$\Omega^{ath} = \prod_{i=1}^{n} \varpi_i^{N_i} \frac{N_t!}{\prod_{i=0}^{n} N_i!} \left(\frac{Q_t!}{N_t!} \right)^{z_0 \kappa / 2} \tag{13.2}$$

where

$$N_t = \sum_{i=0}^{n} r_i N_i \tag{13.3}$$

$$Q_t = \sum_{i=0}^{n} q_i N_i \tag{13.4}$$

$$q_i = \frac{z_i}{z_0} \tag{13.5}$$

and where the exponent in Equation (13.2) is given by

$$\frac{z_0 \kappa}{2} = \frac{\sum\limits_{i=0}^{n} (r_i - 1) N_i}{\sum\limits_{i=0}^{n} (r_i - q_i) N_i} \tag{13.6}$$

If the previously assumed relation between the surface area and the molecule is taken, given by Equation (4.11), then one recovers the original Guggenheim athermal ways function given by Equation (4.9).

The derivative of this expression with respect to N_0 must be taken to utilize the EoS framework in Equation (13.1). This is given by

$$\left(\frac{\partial \ln \Omega^{ath}}{\partial N_0} \right)_{N_i} = \ln\left(\frac{N_t}{N_0} \right) + \left(\frac{z_0}{2} \kappa \right) \ln\left(\frac{Q_t}{N_t} \right) \tag{13.7}$$

To make this expression explicit in the molar volume of the system, a general form of the relationship between the number of vacancies and the total volume of the system, Equation (4.5) is used. The resulting expressions for N_t / N_0 and Q_t / N_t are found and inserted into Equation (13.7) to yield

$$\left(\frac{\partial \ln \Omega^{ath}}{\partial N_0} \right)_{N_i} = \ln\left(1 + \frac{1 - \theta_0}{\theta_0} \sum_{i=1}^{n} \frac{\Theta_i}{c_i} \right) - \left(\frac{z_0}{2} \kappa \right) \ln\left(\theta_0 + (1 - \theta_0) \sum_{i=1}^{n} \frac{\Theta_i}{c_i} \right) \tag{13.8}$$

where the molar volume is found within the expression

$$\theta_0 = \frac{\left(\dfrac{V}{b_0} - \sum\limits_{i=1}^{n} x_i r_i\right) \sum\limits_{i=1}^{n} \dfrac{\theta_i'}{r_i c_i}}{1 + \left(\dfrac{V}{b_0} - \sum\limits_{i=1}^{n} x_i r_i\right) \sum\limits_{i=1}^{n} \dfrac{\theta_i'}{r_i c_i}} \tag{13.9}$$

and where θ_i', the surface area fraction of species i in a system with no vacancies, is given by Equation (5.6). This quantity is not considered a variable of the system, since the system composition and the associated structural properties of the species allows for the exact calculation of these quantities. A collection of volume and surface area characteristics is found in c_i, which is given by

$$c_i = \frac{z_i}{r_i z_0} \tag{13.10}$$

This quantity serves as an important bridge to relating this work to past works in the literature.

13.2 Vacancies within the Quasi-Chemical Equations

The next step in completing the EoS is to find the solution to the model variable Γ_0, which is based within quasi-chemical theory and is a solution to Equation (5.52). To show how vacancies fit within the lattice-fluid system, the nonlinear system is rewritten to include vacancies as a separate species. The system becomes the following

$$\frac{1}{\Gamma_i} = \sum_{j=0}^{n} \theta_j \tau_{ij} \Gamma_j, (i = 0,1,...,n) \tag{13.11}$$

where the surface area fraction of a species also includes the number of vacancies

$$\theta_j = \frac{z_j N_j}{z_0 N_0 + \sum\limits_{k=1}^{n} z_k N_k} \tag{13.12}$$

Separating the term with the vacancies from the remainder of the summation yields

$$\frac{1}{\Gamma_i} = \theta_0 \tau_{i0} \Gamma_0 + \sum_{j=1}^{n} \theta_j \tau_{ij} \Gamma_j, (i = 0,1,...,n) \tag{13.13}$$

To better see how this system reduces to the system with no vacancies, one can manipulate Equation (13.12) to result in

$$\theta_j = \frac{\theta_j'}{\dfrac{z_0 N_0}{\displaystyle\sum_{k=1}^{n} z_k N_k} + 1} = \theta_j' \frac{\displaystyle\sum_{k=1}^{n} z_k N_k}{z_0 N_0 + \displaystyle\sum_{k=1}^{n} z_k N_k} = \theta_j' (1 - \theta_0) \tag{13.14}$$

Inserting this expression into the summation in Equation (13.13) gives the more explicit relationship showing the role of vacancies on the lattice

$$\frac{1}{\Gamma_i} = \theta_0 \tau_{i0} \Gamma_0 + (1 - \theta_0) \sum_{j=1}^{n} \theta_j' \tau_{ij} \Gamma_j, (i = 0,1,...,n) \tag{13.15}$$

These expressions, combined with the equation of state given in Equations (13.1) and (13.8) through (13.10) yield volumetric properties once the variables of the nonlinear system (θ_0 and $\Gamma_i, i = 0,1,...,n$) are found.

13.3 Pure Species: A Binary Mixture of Molecules and Vacancies

The simplest lattice-fluid system that exhibits volumetric properties is where one species with multiple interaction sites of the same type inhabits a lattice with vacancies. Conceptually, this is a mixture system of a species with size r_i and number of interactions z_i with vacancies occupying one lattice site and number of interactions z_0. This is also a generalization to the Ising-like fluids (Stanley, 1971), where the molecular or magnetic species usually occupies one site also, like a vacancy.

The convenience of studying such a system is because one can achieve an analytic solution to the EoS explored in the previous section. For the case where $n = 1$, the system of equations reduces to the EoS of the form

$$\frac{pb_0}{k_B T} = \frac{b_0/b - c}{1 - c} \ln\left[\frac{1}{c} + \theta_0\left(1 - \frac{1}{c}\right)\right] + \left(\frac{z_0}{2} - 1\right)\ln\theta_0 - \frac{z_0}{2}\ln x_{00} \tag{13.16}$$

where $b = rb_0$ and

$$\theta_0 = \frac{V - b}{V - (1 - c)b} \tag{13.17}$$

The local composition of vacancies around vacancies, given by $x_{00} = \theta_0\Gamma_0^2$ is found as the solution to the polynomial

$$\left(1 - \tau_{01}^2\right)x_{00}^2 - \left[\tau_{01}^2/\theta_0 + 2\left(1 - \tau_{01}^2\right)\right]x_{00} + 1 = 0 \tag{13.18}$$

After finding the proper root from Equation (13.18) and substituting in Equation (13.17), Equation (13.16) yields an analytical relationship between the temperature, pressure and molar volume of the lattice-fluid system.

The equation in this form serves as a generalization for the lattice-fluid equations within Section 4.4. The quasi-chemical equations are used in all these methods, and various forms of the Guggenheim athermal ways contribution are applied. Equation (13.16) uses the most general athermal ways function, given by Equation (13.2). This EoS should reduce to the past expressions once one applies the relevant assumptions.

This equation has two lattice-specific parameters, b_0 and z_0, and three molecule-specific parameters, b, c and ε_{11} (within τ_{01}). Each of the molecule-specific parameters may be evaluated theoretically from the first-principles methods to determine AIM structural properties of molecules (for b and c) and energetics between molecules (for

ε_{11}). However, these structural properties are dependent on the characteristics of the lattice, since b depends on the size of a lattice site and c depends on this and the coordination number of the lattice.

The first attempt at evaluating the lattice parameters is made through correlation of the equation to reproduce the experimental critical point of the pure fluid. To eliminate one of the lattice parameters from consideration, it is assumed that the Guggenheim relation between the size of the molecule r and the surface area parameter $q = z/z_0$ is true

$$\frac{z_0}{2} = \frac{r-1}{r-z/z_0} = \frac{1-b_0/b}{1-c} \qquad (13.19)$$

This allows for the coordination number of the lattice to be describable by the fitted parameters in the equation. For the other lattice parameter, it is assumed that all the lattices on which the molecules reside is given the volume of $b_0 = 10.0 \text{ cm}^3/\text{mol}$, approximately the AIM volume of neon.

The correlation procedure to find b, c and ε_{11} consists of forcing the EoS of Equation (13.16) to reproduce the critical point (T_C, p_C and V_C) and the conditions of Equation (2.30). The correlation is performed on a number of small molecules that contain less than four heavy atoms. Table 13.1 shows the results of the correlation, and several trends within the parameters can be noted. Within the noble gases, appropriate trends are seen within the size of the molecules and the bulkiness of the molecule, given by c. Within the small molecules, those with no dipole moments have an interaction energy of about -200 K, while carbon dioxide has a larger interaction energy likely due to the quadrupole moment. For those small molecules with dipole moments, the

Table 13.1 EOS Parameters Correlated to Reproduce Critical Point

	molecule	b (cm^3/mol)	c	ε_{11}/k_B (K)
noble gases	Ne	19.0	0.812	-35.2
	Ar	33.1	0.574	-196.0
	Kr	40.2	0.501	-304.8
	Xe	51.6	0.417	-476.5
small molecules-no dipole moment	N_2	39.5	0.516	-175.2
	O_2	32.7	0.570	-207.5
	F_2	29.7	0.603	-186.5
	CO_2	42.2	0.452	-516.9
small molecules-dipole moment	CO	40.6	0.515	-180.7
	NO	27.1	0.565	-314.8
	N_2O	43.6	0.440	-537.1
	H_3N	33.8	0.472	-809.8
	H_2O	26.6	0.541	-1318.0
fluoromethanes	CF_4	61.5	0.341	-445.8
	CHF_3	60.8	0.321	-648.1
	CH_2F_2	56.2	0.312	-853.2
	CH_3F	52.6	0.328	-753.6
	CH_4	43.4	0.471	-292.0
methyl-	CH_3CH_3	64.0	0.340	-582.6
	CH_3NH_2	56.9	0.337	-918.1
	CH_3OH	55.5	0.295	-1375.6
	CH_3F	52.6	0.328	-753.6
ethyl-	$CH_3CH_3CH_3$	88.0	0.259	-814.0
	$CH_3CH_3NH_2$	80.9	0.266	-1027.1
	CH_3CH_3OH	77.2	0.237	-1389.7
	CH_3CH_3F	73.7	0.278	-851.9

interaction energies become larger as the magnitude of the dipole increases (from top to bottom). Within the remaining three sections, the volumes and the interaction energies follow appropriate trends.

Figures 13.1 and 13.2 show how the correlated structural EoS parameters relate to the critical volumes and structural parameters found in AIM calculations. Firstly, critical volumes and AIM volume show a strong linear dependence to the correlated parameter, thus qualifying its physical significance in the fluid model. With regards to the parameter c, recall its definition in Equation (13.10). The numerator contains the factor z_i, which reflects the amount of external surface area a molecule has to interact with other molecules. The denominator contains the factor r_i, which signifies the number of lattice sites a molecule occupies and is related to the volume of the molecule. Therefore, it is expected that the ratio of the AIM external surface area and the AIM volume of the molecule is related to the correlated parameter c. As Figure 13.2 suggests, there is a monotonic linear trend in the correlation.

Figure 13.3 depicts the ability of the equation of state with correlated parameters to reproduce liquid-like and vapor-like volumes of nitrogen on the critical isotherm. As required, the EoS reproduces the critical point, depicted as a red diamond near the center of the figure. It also reproduces the gas-like densities quite well except for the small error in the intermediate region between ideal gas densities and the critical density. The equation does poorly for liquid-like volumes, where the equation predicts smaller densities than experiment at a given pressure. The shortcomings with this EoS are common for those equations that are analytic through the critical region, as mentioned in
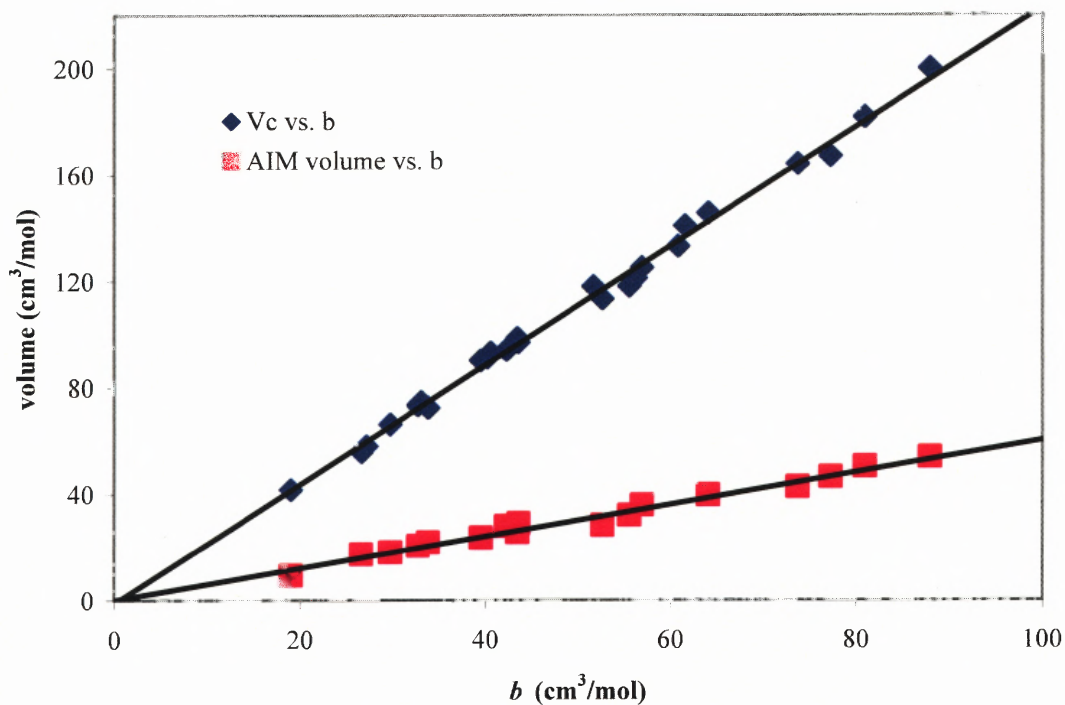
Figure 13.1 Critical volume and AIM volume related to the correlated EOS structural parameter b.
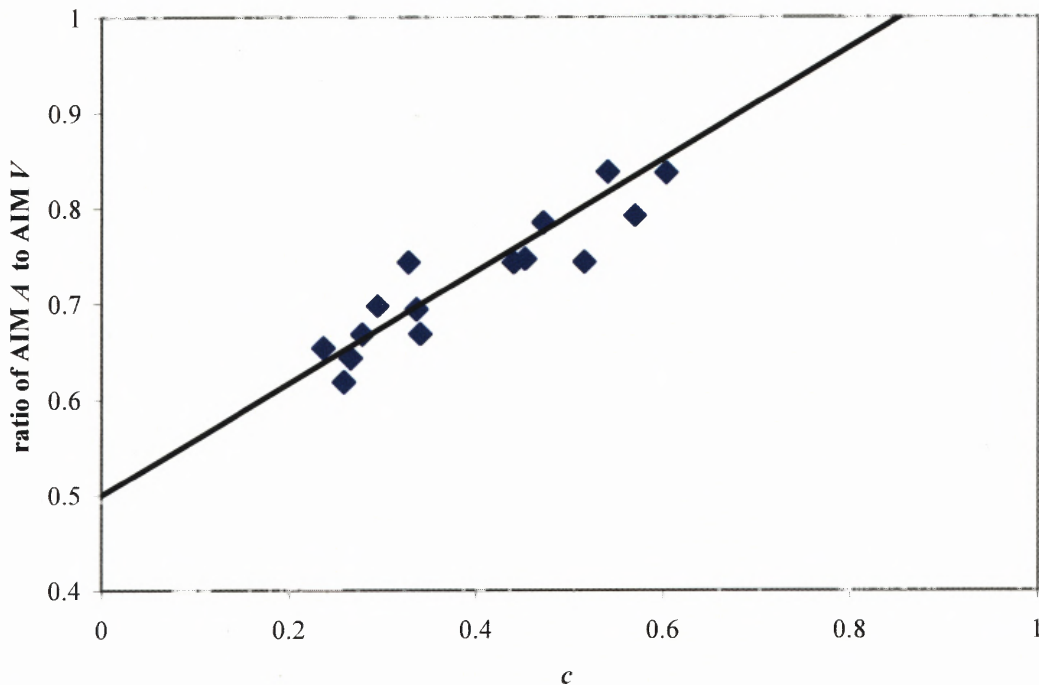
Figure 13.2 Ratio of AIM area to AIM volume related to correlated EOS structural parameter c.
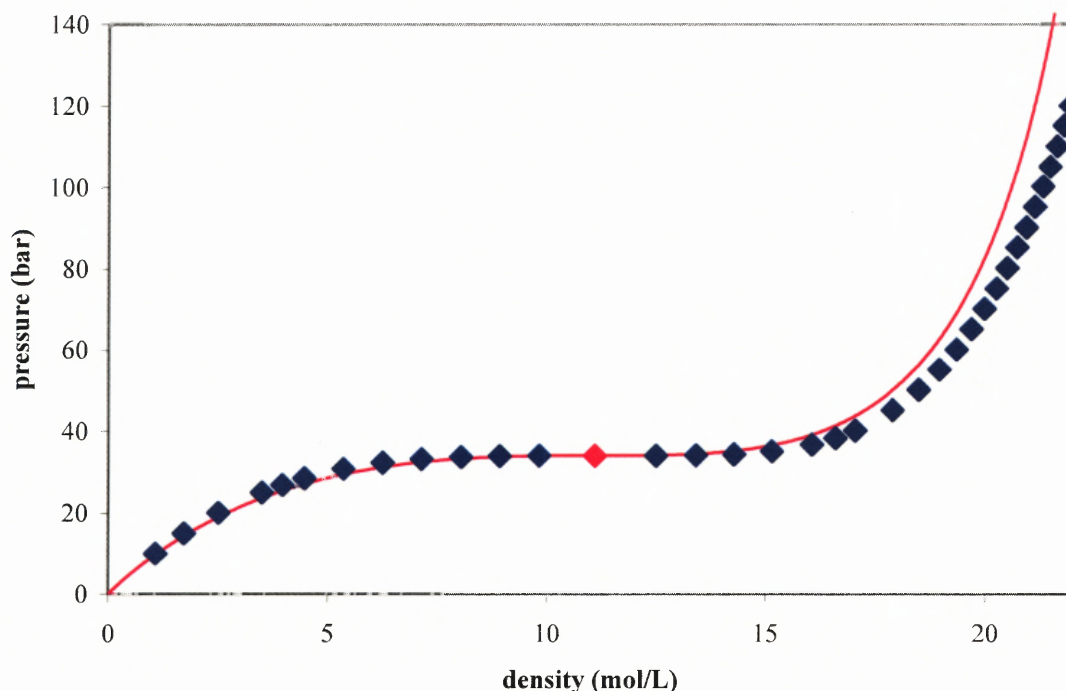
Figure 13.3 Predicted critical isotherm behavior of nitrogen compared to experiment.

Section 2.4. The equation needs to be flat for a larger range of densities in the region around the critical point, therefore accommodating all the data nearby and reaching the gas-like and liquid-like behavior. Of issue is the fact that far from the critical point, the EoS must behave as a cubic equation, while at the critical point, the function must have a power that is found to asymptotically approach 4.4 for carbon dioxide (Stanley, 1971).

13.4 Equation of State for Larger Molecules

To attain a simple expression for the volumetric behavior of small molecules, it is assumed that the interaction energies between two molecules are equivalent for all the possible combinations of interaction sites. This assumption begins to break down as molecules become larger and as the atoms that make up the molecule become more

varied. Even for a relatively small molecule like ethanol, it is difficult to accept that all the interactions between two ethanol molecules always results in an energy of $-1400\ \mathrm{K}$, (as suggested by Table 13.1) since these interactions are highly dependent on orientation effects (as depicted in the interaction matrix of Table 12.2).

The logical step in determining the volumetric properties of these larger molecules is to consider the entire matrix of interactions in the quasi-chemical approach within Equation (13.15). The structural parameters of the molecules, also necessary in the nonlinear system of equations, are available through the use of the AIM properties within Appendix G. This treatment results in a larger number of nonlinear equations to solve, and a closed-form solution such as that for small molecules is not possible. For a given isotherm, it is necessary to solve the system of equations for each fluid density to yield pressures at those conditions, and therefore a graph of continuous curves to show the volumetric behavior of the fluid.

The determination of the volumetric properties of a fluid with the use of the energetics from Chapter 12 and the AIM properties of Chapter 10 can be considered a fully theoretical prediction of macroscopic system properties from first principles. No parameters within the model are fitted to experimental data. All are found using computational techniques described in this work.

Like the small molecule EoS, two lattice parameters need to be determined to perform calculations with the lattice-fluid model. These parameters, b_0 and z_0, determine the lattice on which the fluids are modeled. For the modeling of the larger molecules, the methane molecule is chosen as the reference species. The volume of a lattice site b_0 is therefore $287.6\ \mathrm{au}$ ($25.66\ \mathrm{cm^3/mol}$), and the total surface area of a

lattice site is 213.28 au (0.5972 nm^2). This latter value combined with the interaction

area corresponding to those used in COSMO-based models, 25.0 au (0.070 nm^2), yields

the number of interactions for a vacant lattice site, $z_0 = 8.53$.

Table 13.2 lists an analysis of the predicted critical points of a range of molecules

considered later in mixture systems. The EoS does a poor job in reproducing the critical

points for all the fluids considered. The model also is unable to correctly predict the

trends with the critical temperature. Also noted with the small molecule EoS described

above, this EoS is analytical, and therefore suffers the same problems in the critical

region as those of cubic and all other analytical equations of state.

Table 13.2 Calculated Critical Parameters Compared to Experiment

	Calculation			Experiment		
	T_c	p_c	V_c	T_c	p_c	V_c
propane	205	50	120	369.83	42.48	200
butane	201	37	160	425.12	37.96	255
pentane	199	30.5	195	469.7	33.7	311
hexane	198	25	245	507.6	30.25	368
DME	375	65	142	400.1	54	170
methanol	933	215	85	512.64	80.97	118
ethanol	658	61.5	165	513.91	61.48	167
propanol	593	32.5	270	536.78	51.75	219
butanol	493	15.5	350	563.05	44.23	275

Figure 13.4 shows the isotherms of propane compared to experimental data at

various reduced temperatures. The reduced quantities within the model predictions are

related to the critical point predicted by the model, while the reduced quantities within

the experimental data are related to the experimental critical point. The structural
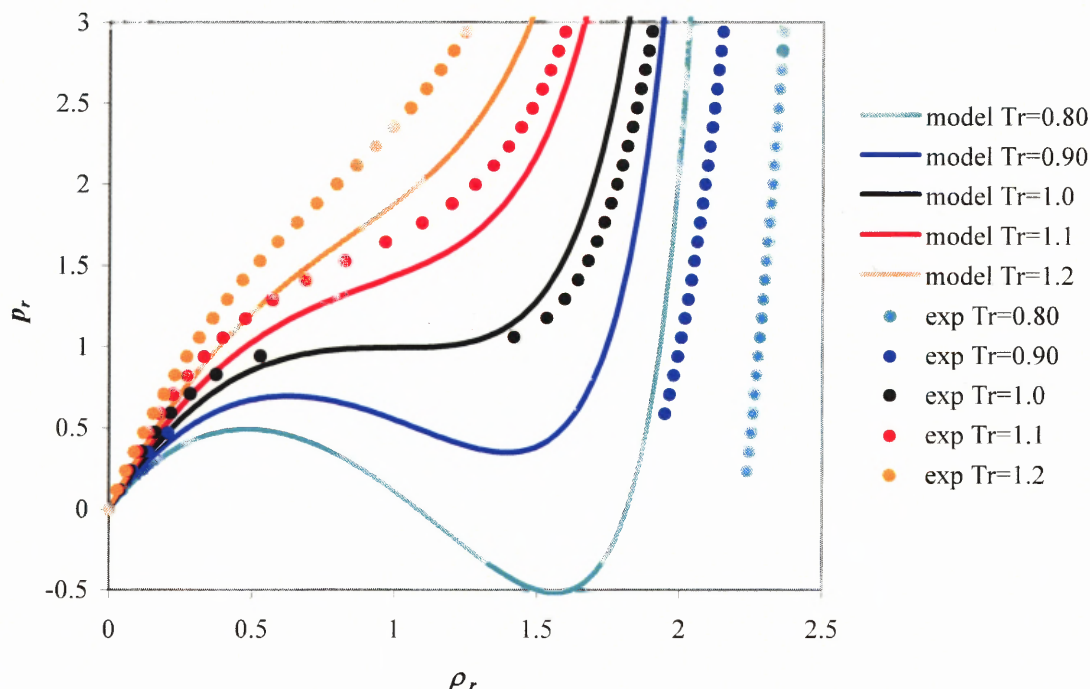
Figure 13.4 Comparisons of isotherms of propane with experimental data.

parameters for the functional groups of propane are given in Appendix G, and the energetics between functional groups are given in the interaction matrix in Table 12.2. Seen here is the characteristic volumetric behavior of a fluid at the supercritical region, the critical isotherm, and at subcritical regions where liquids and gases may coexist. The statistical model shows the ability to create such descriptions of fluids with quantities found independently of experiment. However, the model predictions of the isotherm data are poor. These can be attributed to the failure of the model around the critical region.

The issues with the first-principles predictions of the critical point of these fluids reside within the partition function. It has been an extremely difficult task to formulate a realistic EoS, mainly because of the lack of understanding of correlation effects with liquids and near-critical fluids. The low critical temperatures within the alkanes suggest a correlation effect that is not accounted for within the statistical model, an effect that

allows for greater cohesion between molecules within the liquid phase. Also, the partition function approach leaves out any description of the translational motion of the molecules, instead including molecular partition functions that describe the vibrational and rotational motion of molecules around their lattice sites. This molecular partition function is independent of the system pressure, and therefore does not contribute to the EoS (as can be seen in the relation between the partition function and the EoS in Equation (3.30)). This may be why the EoS does not reflect the mass effects (and therefore, the translational energy) within the series of alkanes. This seems likely also for the incorrect trend within the alcohols, where the larger molecules should need a higher temperature to achieve a supercritical state because of their greater mass.

13.5 Conclusions

The generalized Guggenheim statistics derived in this work offers a more general description of athermal lattice-fluid systems than previous expressions. This equation with the quasi-chemical equations is used to formulate a closed-form engineering EoS. This EoS has five parameters: two parameters characterizing the lattice and three parameters related to the molecular species. When the molecular parameters are fit to the critical point, physical significance is seen in the each of the parameters, and relations seem to exist between appropriate AIM properties of the molecules.

With the application of the generalized athermal ways expression, the structural characteristics need not obey the past assumption requiring a strict relationship between number of lattice sites occupied and number of external contacts . Such flexibility allows for the description of fluid species with parameters reflecting the molecular structure

found using *ab inito* methods. When this approach is taken for larger molecules, functional groups are used and pure species volumetric behavior is able to be predicted. However, the model fails to correctly predict the critical point of alcohols and alkanes.

The use of fundamental quantities as the parameters within the EoS allows for an assessment of the capabilities of the statistical model. With a larger set of results using the AIM properties calculated from this work, the errors within the statistical model may be diagnosed, and appropriate changes may be made.

CHAPTER 14

PREDICTIONS OF VAPOR-LIQUID EQUILIBRIUM
USING LATTICE-FLUID THEORY

The success of the COSMO-based methods is a testament to the power of the quasi-chemical approach in describing the thermodynamic properties of mixture systems. With the simple information describing the electrostatic potential around a molecule, intermolecular interactions can be analyzed and sorted using statistics. The partition function of the system emerges from these statistics, and vapor/liquid equilibrium, liquid/liquid equilibrium, partition coefficients and solubilities are made available to engineers.

This chapter considers criticisms of the COSMO-based approaches and fills in these gaps with the molecular-level properties studied in this work. The structural and electrostatic properties found from AIM Theory and the functional group interaction potentials offer first-principles information for the physical parameters used in modeling mixture systems. Equations are developed to incorporate these parameters to fully predict vapor/liquid equilibrium of a range of fluid systems. Attempts are made to predict the mixture behavior of systems containing alcohols, alkanes and ethers. Limits to the quasi-chemical approach are discussed. Modifications are made to these equations to better represent real fluid systems, and these results are compared to experiment.

This chapter outlines a methodology that attempts to fully predict the vapor/liquid system behavior without correlating any model quantity to VLE experimental data. Up to this point, no model has accomplished this; instead, past models choose to correlate structural, electrostatic or energetic information to experiment. This work uses calculated

surface areas and volumes from isolated molecule electron density profiles as the structural parameters. Energetics are found by applying a functional group interaction potential whose parameters have been found by reproducing phase equilibrium of pure species in molecular simulations. With further development, the first-principles force field of this work, where structural and electrostatic properties from isolated molecule electron densities are applied to an interaction model, will be used to offer an interaction matrix with no fitted parameters. This novel approach to predicting VLE, when validated, achieves the goal of first-principles knowledge of macroscopic fluid system behavior without the need for correlation to experiments.

14.1 Criticisms of COSMO-based Approach

Since COSMO-based methods are being used to predict a wide range of fluid-phase properties, their modeling successes and failures are being scrutinized. Several criticisms have arisen in the recent past, ranging from the fundamental assumptions of the model to the limitations in the accuracy of the model results. The criticisms focused on here are mainly on the basic premises of the model.

A criticism of the COSMO-based methods is the way the location and form of the surface charge density that exists around the molecule and takes part in interactions. The distance from the atoms at which these charge densities are sampled is a parameter that is correlated to experimental data. Although originally attributed as a surface that is 120% the van der Waals radius away from the atoms, a value used in most continuum solvation models, the radius has changed in more recently refined versions. These radii form spheres around the atoms, and the union of these spheres create crevices that have been

noted to hinder the calculations of surface charge density. These since have been fixed. Techniques that sample any electrostatic property outside the significant electron density of a molecule are subject to variation with the points that are sampled. Not only is the choice of points on the surface important, but the depth within which these points are sampled has a strong effect on properties. To achieve a more correct description of the electrostatic properties, several surfaces at different depths may be necessary.

It is assumed within the COSMO-based methods that interactions are attributable to the energetic effects that develop when surface segments overlap in the liquid phase. Figure 14.1 shows a two-dimensional representation of the layout of molecules within COSMO-based methods. As show in the figure, this assumption is based on the premise that the molecules will only interact at distances that are given by the sum of the atomic
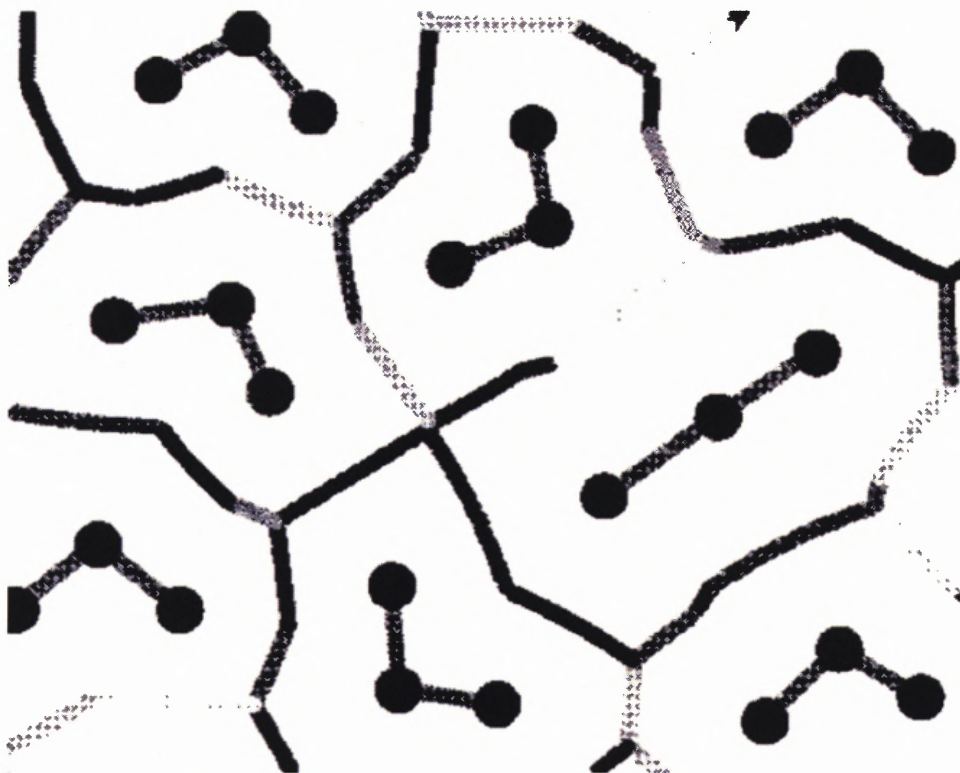
Figure 14.1 Two-dimensional depiction of the COSMO-based model lattice.

radii. Figures 14.2 and 14.3 depict the electron densities of molecules interacting under this assumption. The description of the methyl-methyl interaction in Figure 14.2 seems plausible, with the centers of the groups (the carbon atoms) at a distance of 4.00 Å apart. This corresponds roughly to the most probable interaction distance predicted by the functional group interaction model in this work (4.23 Å) and the TraPPE force field (4.21 Å), both shown in Figure 12.7. The COSMO-based interaction also agrees with the physical significance of the 0.001 au isodensity surface, that molecules interact in such a way where the 0.001 au surfaces are juxtaposed.

The water dimer interaction of Figure 14.3 is also depicted where the isodensity surfaces are juxtaposed. This interaction distance is, as suggested by the COSMO-based radii of the atoms, 3.02 Å . However, it is well known that the interaction between water
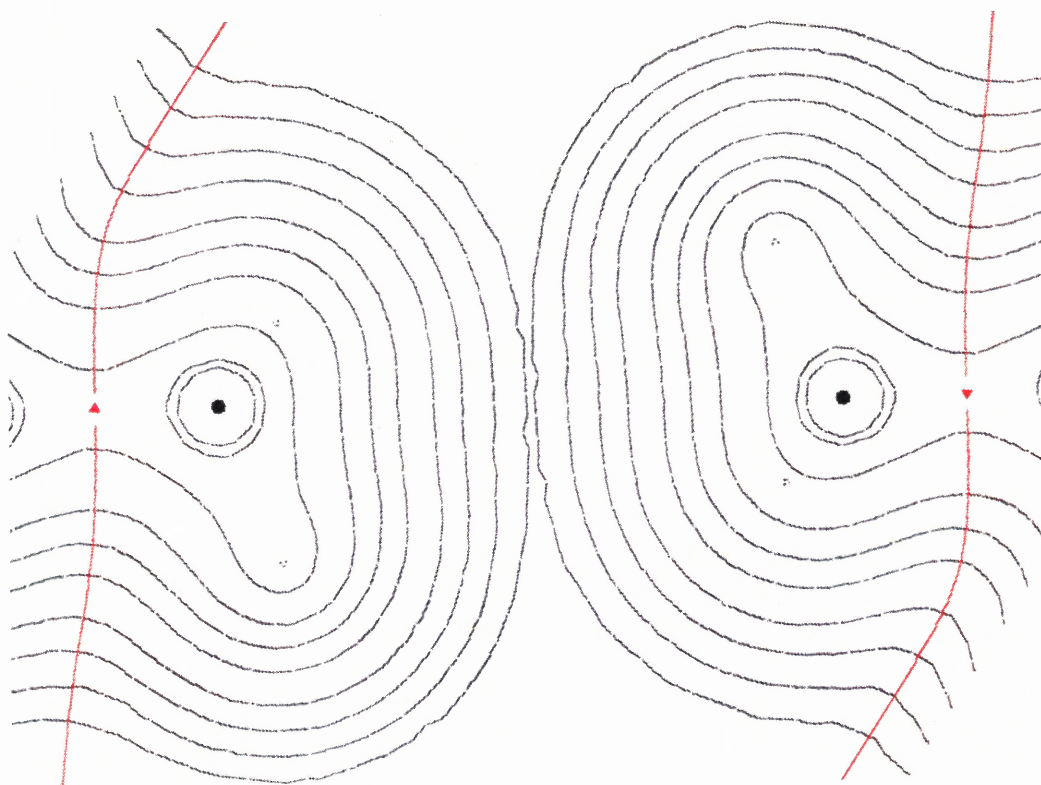
Figure 14.2 COSMO-based interaction between methyl groups.
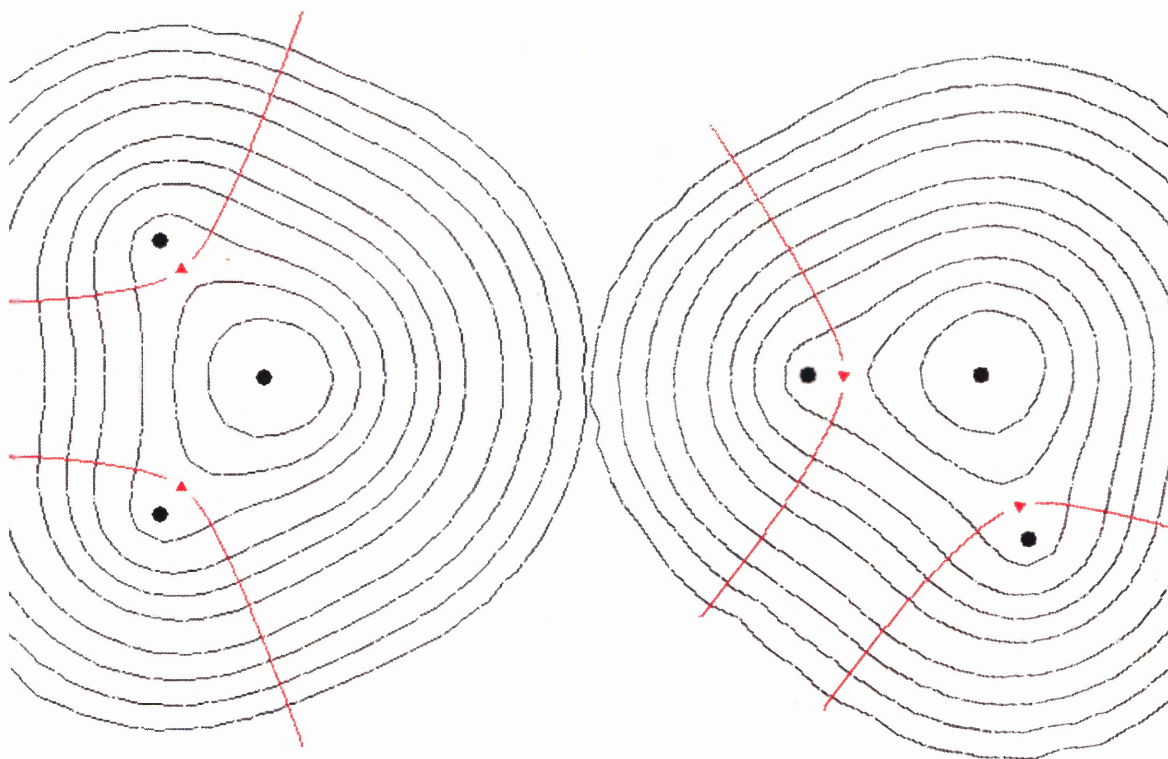
Figure 14.3 COSMO-based interaction between water molecules.

molecules is nearly 2 Å. Figure 14.4 depicts the interaction between water molecules at the more accepted interaction distance. Here the large electron density overlap expected by hydrogen bonding is apparent. The 0.001 au surface of each atom is roughly tangential to the 0.08 au surface of the other atom. Hydrogen-bonding interactions between water molecules have also been depicted in Figures 11.1 and 11.2. There too, it is shown the interaction distance assumed by the COSMO-based scheme is not a good assumption.

The sampling technique employed by the COSMO-based methods does not account for such a large discrepancy in interaction distances. The fix applied to the model is held within the correlation constants in the interaction energy terms. There are two such correlation parameters: one which scales the difference in segment charge
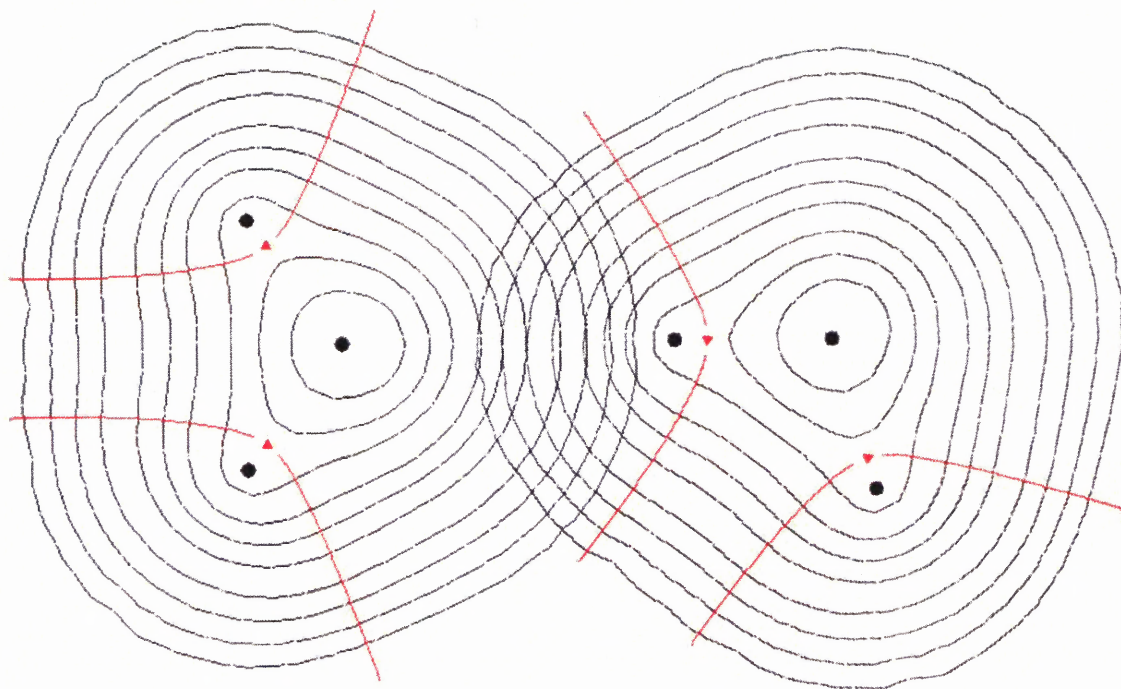
Figure 14.4 Interaction between water molecules at a distance of approximately 2 Å.

densities for non-hydrogen-bonding interaction schemes; one which scales the difference in segment charge densities for hydrogen-bonding situations. Here, an arbitrary measure must also be implemented to determine what charge density difference constitutes a hydrogen-bonding interaction.

The statistics employed within the COSMO-based methods consider all the molecules interacting with other molecules in the way described in the previous paragraphs. The statistics do not account for the possibility of vacancies within the lattice, a concept that is essential to predicting volumetric properties. If it is assumed that the COSMO-based lattices are fully occupied by molecules, and where the molecules are interacting in ways depicted in Figures 14.1 through 14.3, then the sum of the AIM volumes calculated in Appendix G should correspond to the liquid-like volume of the

species (since Figure 10.1 suggests that the 0.001 au isodensity surface roughly corresponds to the COSMO-based surfaces). Calculation of any of these AIM volumes for a molecule yields a value that is far below that of a liquid-like volume for the species. To reproduce these volumes, either larger distances must exist between the molecules (unlikely, since past studies do indicate that interactions probably occur at these distances envisioned by COSMO-base methods), or that vacancies must be included in the lattice. Including vacancies is not straightforward in COSMO-based methods, since the charge density surfaces of molecules must interact with charge densities of vacancies that are influenced by nearby molecules. The lack of vacancies is the same approach taken by UNIQUAC, UNIFAC and other past engineering models, where the liquid is on a packed lattice.

The thermodynamic inconsistency of an earlier version of COSMO, the COSMO-RS method, has been resolved by the creation of the COSMO-SAC method and the COSMOSPACE method. Both these methods are based around the quasi-chemical equations and the entropy of the athermal system described in Sections 4.2 and 5.4.

This work attempts to address the criticisms above by using the molecular-level properties within a lattice-fluid model that incorporated vacancies. The structural properties within AIM are used instead of the correlated radii to determine the extent of electron densities. The interaction model proposed in this work takes the structural and electrostatic properties and converts these into interaction energies and distances describable with an interaction curve. Since calculated curves in this work have not been achieved, similar properties, attainable using the TraPPE force field created for molecular dynamics simulations, are used. These curves avoid the predisposed interaction distances

necessary in the COSMO-based methods, and allow for a description of interaction energies at all distances. The most favorable interaction energies are then used in a statistical model similar to that of the COSMO-based methods, except where vacancies are included on the lattice. The interaction energy between a molecule and a vacancy, as well as an interaction between two vacancies, vanishes and thus agrees with acceptable physical principles.

14.2 The Development of the Lattice-Fluid Model to Predict VLE

The prediction of VLE using the lattice-fluid framework is focused around developing expressions for the activity coefficient γ of the species. These values are then used within the gamma-phi formulation reviewed in Section 2.2. The focus now is to lay out the process by which these activity coefficients are calculated and implemented.

The activity coefficient for species i in a liquid mixture is given by Equation (2.22). To evaluate this expression, equations describing the Gibbs energies of the mixture system and the pure fluid system must be determined. For lattice fluids, these can be found using the partition function in Equation (4.4). This equation is governed by the temperature, pressure and composition of the system.

The model variables for this partition function, the numbers and types of interactions M_{ij}, and the number of vacancies N_0 are found by solving the quasi-chemical equations, given by Equations (5.29) through (5.31) and developed in Appendix B. An equivalent and simpler way of determining the model variables is to consider the relation given by Equation (5.52) and solving that nonlinear system, where the interacting entities are functional groups. Both of these treatments involve solving the quasi-

chemical equations for the pure and mixture system, so one may determine the model variables for both systems for application within Equation (5.51).

The athermal contribution to the activity coefficient is found by utilizing the athermal ways function of Equations (4.9), (4.13), or (13.2), where the last of these is a generalization of the earlier two. The full expression for $\ln \gamma_i^{ath}$ is given by

$$\ln \gamma_i^{ath} = -\left(\frac{\partial \ln \Omega^{ath}}{\partial N_i}\right)_{T,p,N_{j\neq i}} + \left(\frac{\partial \ln\left(\Omega^{ath,(i)}\right)}{\partial N_i}\right)_{T,p,N_{j\neq i}} - \ln x_i \qquad (14.1)$$

where the athermal ways within the second term is for the pure system at the mixture system temperature and pressure. If one uses the classical Guggenheim expression of Equation (4.9), this relation reduces simply to

$$\ln \gamma_i^{ath} = \ln\frac{\phi_i}{x_i} + \left(1 - \frac{\phi_i}{x_i}\right) - \frac{z_i}{2}\left(1 - \frac{\phi_i}{\theta_i} + \ln\frac{\phi_i}{\theta_i}\right) \qquad (14.2)$$

However, this expression had been developed without vacancies on the lattice. Therefore, it must be assumed that the volume fraction ϕ_i and the surface area fraction θ_i are taken in a fully packed lattice, with no vacancies.

The gamma-phi formulation in its original presentation assumes that γ_i is independent of pressure. Since this work is using the Gibbs ensemble, pressure is a thermodynamic variable that influences all other variables within the equations. Therefore, for procedures that iterate the pressure, such as BUBL p and DEW p, the most general procedure involves calculating γ_i at every pressure iterate. This involves calculating the solutions of the pure and mixture species quasi-chemical equations each

time. With computers and automated algorithms, this may be accomplished while still achieving a timely result.

If one can assume that $\Phi_i = 1$ within the gamma-phi formulation, then the pressure of the mixture system is readily given by summing both sides of Equation (2.23)

$$p = \sum_i x_i \gamma_i p_i^{sat} \tag{14.3}$$

where either the experimental results or Antoine's equation is used for p_i^{sat}. The vapor compositions given by

$$y_i = x_i \gamma_i p_i^{sat} / p \tag{14.4}$$

These equations will be used in this work to predict isothermal VLE data for a variety of mixture systems. Since it is easier to know the liquid compositions x_i in experiments, the pressure and vapor compositions will be pursued, as in the BUBL p procedure.

14.3 Calculated VLE Behavior using the Full Treatment

In theory, once AIM properties of a species are found, interaction energies, pure fluid and mixture system behavior can be predicted using the interaction model in Chapter 12 and the lattice-fluid statistics in Chapters 4, 5, 13, and 14. This work focuses on the mixture systems describable with the TraPPE interaction matrix of Table 12.2, which gives the interaction energies for systems containing linear alkanes, linear alcohols and dimethylether. This information allows for the prediction of pure system properties (done in Chapter 13) and several combinations resulting in binary mixture systems.

Analyzing a mixture system using the full treatment of the theory involves allowing the EoS to determine the model variables (namely Γ_i and θ_0) of the pure

systems and the mixture system. These solutions and the generalized Guggenheim equation is used to determine the activity coefficients in Equations (5.51) and (14.1) and the pressure for the species from Equation (14.3). This is done in an iterative manner using the BUBL p procedure. The vapor composition is calculated from Equation (14.4), and the results are compared to experiment. Also available from this treatment is the excess volume of the binary mixture, assumed not to vanish in the most general version of the model.

The first system treated in this manner is an ideal binary mixture system containing 1-propanol/1-butanol at 373.15 K, depicted in Figure 14.5. The model here predicts ideal solution behavior and the associated vapor phase compositions. The experimental data shows some curvature that migrates around the predicted p-x line. The experimental data offers no vapor compositions for this mixture.

The second system treated is the binary mixture containing dimethylether and ethanol at 293.15 K, depicted in Figure 14.6. Here, the calculated p-x line is well above the experiment, while there is agreement with the p-y calculation and experiment. The disagreement in the p-x behavior is due to the inability of the EoS to accurately predict the liquid volume of the mixture system. In this case, too few vacancies exist in the ethanol-rich branch, and the dimethylether molecules tend to vaporize rather than interact with the ethanol molecules. As shown in the next section, the p-x line is modified if the experimental liquid volume is used.
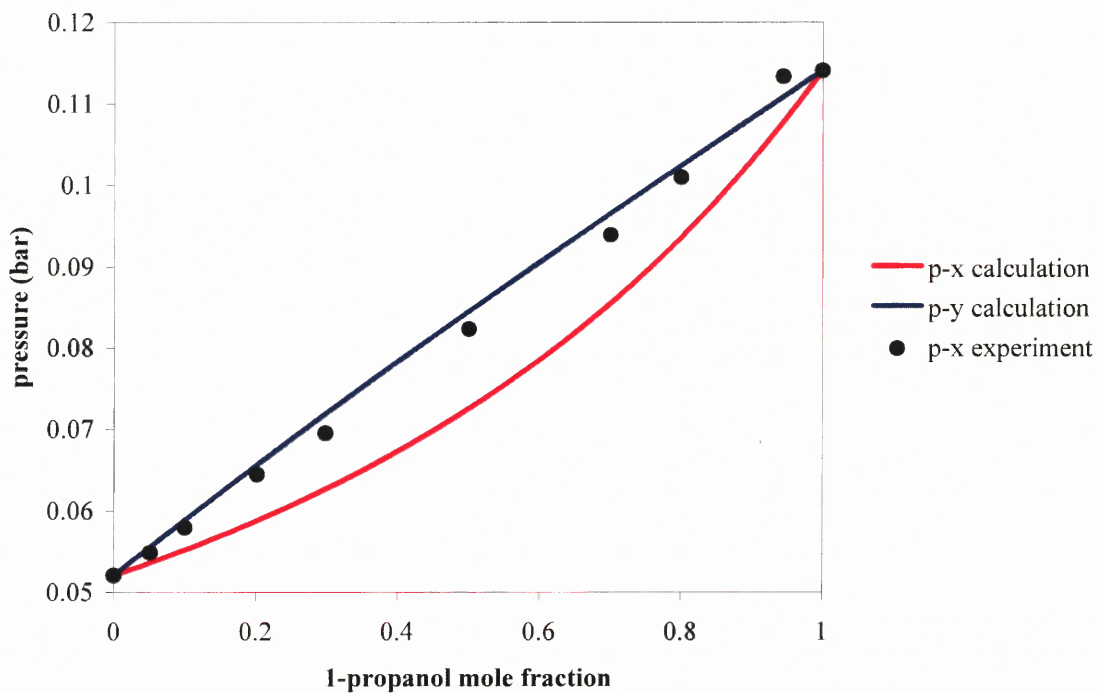
Figure 14.5 VLE for 1-propanol/1-butanol at 373.15 K.
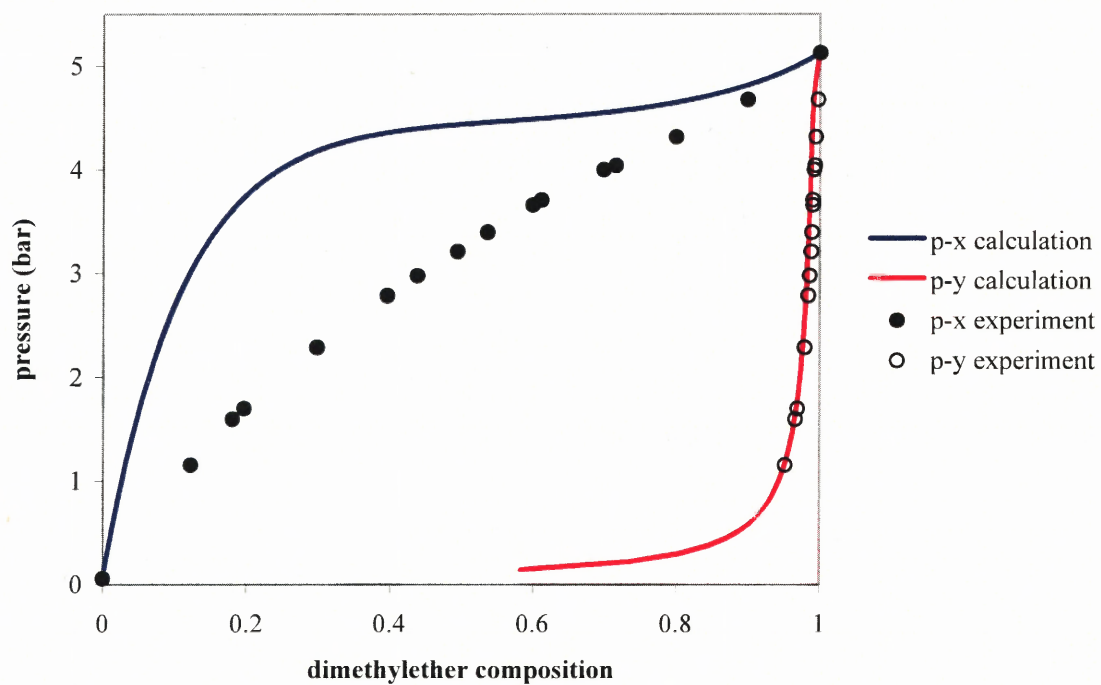
Figure 14.6 VLE for dimethylether/ethanol at 293.15 K.

14.4 The Use of Experimental Liquid Volumes

In several systems of interest to this work, the EoS fails to predict a liquid phase or a liquid-like volume for the pure species at the mixture system temperature. For instance, the EoS critical parameters predicted in Table 13.2 suggests that linear alkanes exist in a supercritical state for all temperatures above 200 K. Therefore, any mixture system prediction using the full treatment of the method fails unless the system temperature is less than 200 K. Also, although dimethylether has a predicted critical temperature of 375 K, a prediction for the mixture system with the species at 353.15 K is not possible because a liquid-like volume does not exist near the experimental or mixture system pressure at that temperature.

To alleviate this shortcoming, the experimental liquid volume at near-ambient temperatures from a standard engineering source (Poling, et al. 2001) is used to determine the number of vacancies on the lattice. In this case, the mixture system volume is considered the weighted average of the two systems, and the excess volume of the mixture is considered to vanish. While not true in general, this is a reasonable approximation because the excess volume is often quite small.

To apply experimental liquid volumes that determine the number of vacancies in the pure and mixture system, more care must be taken in calculating $\ln \gamma^{ath}$ within Equation (14.1). To start, consider the functionality of the athermal ways function for a binary system

$$\ln \Omega^{ath} = \ln \Omega^{ath} \left(T, p, N_1, N_2, N_0 \left(T, p, N_1, N_2 \right) \right) \tag{14.5}$$

The derivative of the mixture system $\ln \Omega^{ath}$ within Equation (14.1), in its most general form, is given by

$$\left(\frac{\partial \ln \Omega^{ath}}{\partial N_i}\right)_{T,p,N_j} = \left(\frac{\partial \ln \Omega^{ath}}{\partial N_i}\right)_{T,p,N_j,N_0} + \left(\frac{\partial \ln \Omega^{ath}}{\partial N_0}\right)_{T,p,N_1,N_2} \left(\frac{\partial N_0}{\partial N_i}\right)_{T,p,N_j} \qquad (14.6)$$

For the pure system $\ln \Omega^{ath,(i)}$

$$\left(\frac{\partial \ln \Omega^{ath,(i)}}{\partial N_i}\right)_{T,p} = \left(\frac{\partial \ln \Omega^{ath,(i)}}{\partial N_i}\right)_{T,p,N_0} + \left(\frac{\partial \ln \Omega^{ath,(i)}}{\partial N_0}\right)_{T,p,N_i} \left(\frac{\partial N_0}{\partial N_i}\right)_{T,p} \qquad (14.7)$$

In general, the functionality of N_0 is determined by the thermodynamic variables of the system. Within the full treatment of the method, N_0 is determined by the minimization of the Gibbs energy (maximizing the term in the Gibbs partition function) with respect to all the model variables. However, here the experimental liquid volume is used to set this quantity. The function describing N_0 here becomes

$$N_0(N_1,N_2) = N_1\left(\frac{V_1^l}{b_0} - r_1\right) + N_2\left(\frac{V_2^l}{b_0} - r_2\right) \qquad (14.8)$$

The derivatives of this expression must be included within Equations (14.6) and (14.7) for the information to be applied properly.

The first system to be calculated using the experimental liquid volumes is the pentane/hexane ideal binary mixture system at 298.7 K, depicted in Figure 14.7. Here, both the p-x and the p-y lines are in good agreement with the experimental values, where the experimental values depict slightly higher pressures in both the p-x and the p-y lines.

The second system to be calculated here is the dimethylether/methanol system at 353.15 K, depicted in Figure 14.8. As mentioned before, although dimethylether is predicted to be sub-critical at this temperature, no liquid-like volume is predicted at the experimental pressures (1.96 to 22.3 bar). Here, the model predicts a more ideal system
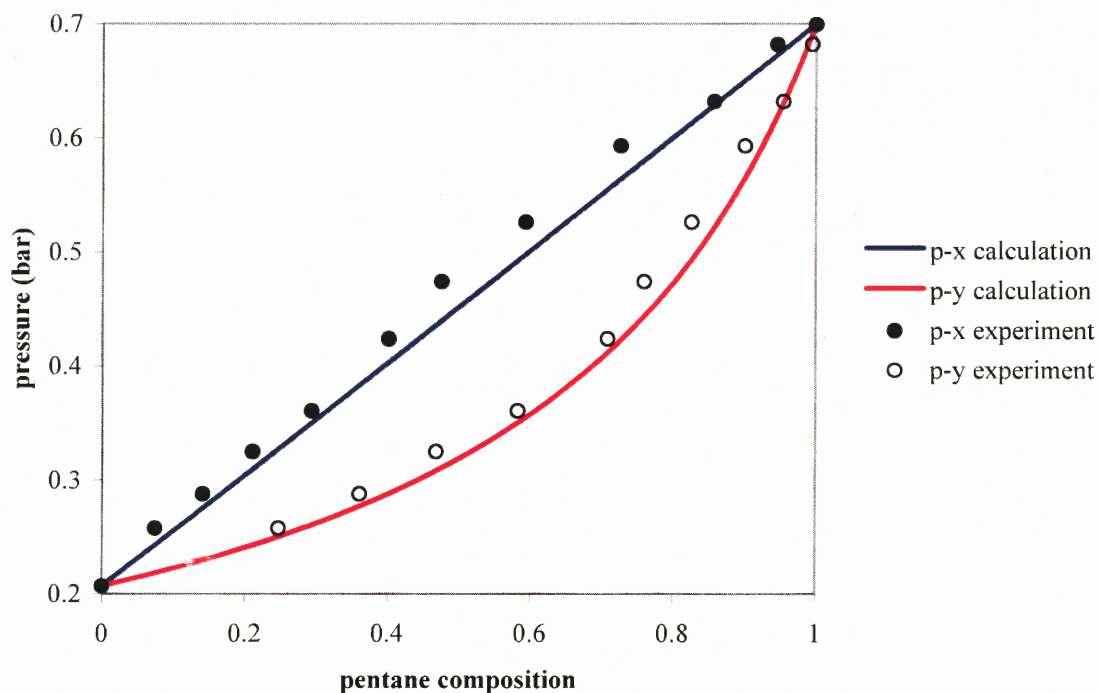
Figure 14.7 VLE for pentane/hexane at 298.7 K.
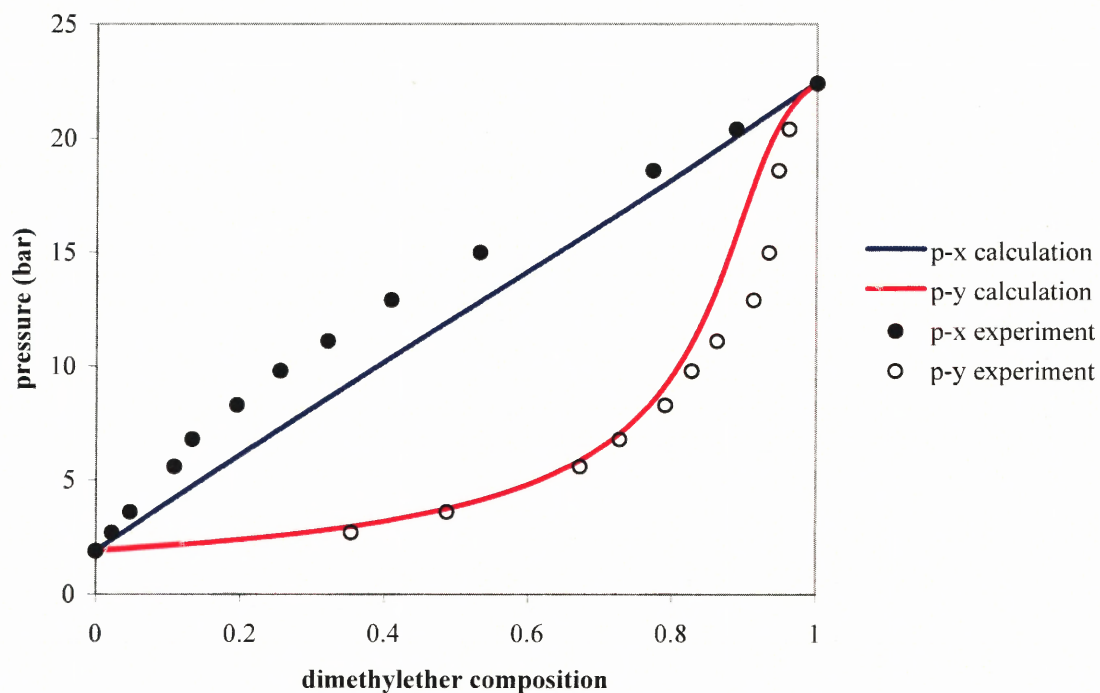
Figure 14.8 VLE for dimethylether/methanol at 353.15 K.

in the *p-x* line. The curvature in the *p-y* line at high dimethylether compositions is reproduced, although the prediction does not correspond to data well at these compositions.

The third system to be calculated with the use of experimental liquid volumes is the dimethylether/ethanol system at 293.15 K, depicted in Figure 14.9. Here, the use of the experimental liquid volume reproduces the *p-x* line very well, while the *p-y* line is still reproduced well. Successful use of the liquid volumes here is likely due to the similarity in system temperature with the temperatures at which the liquid volume is measured. The system of dimethylether/methanol is at a higher temperature (353.15 K), and thusly the experimental liquid volumes used may not correspond well to the liquid volumes for the species at 353.15 K.
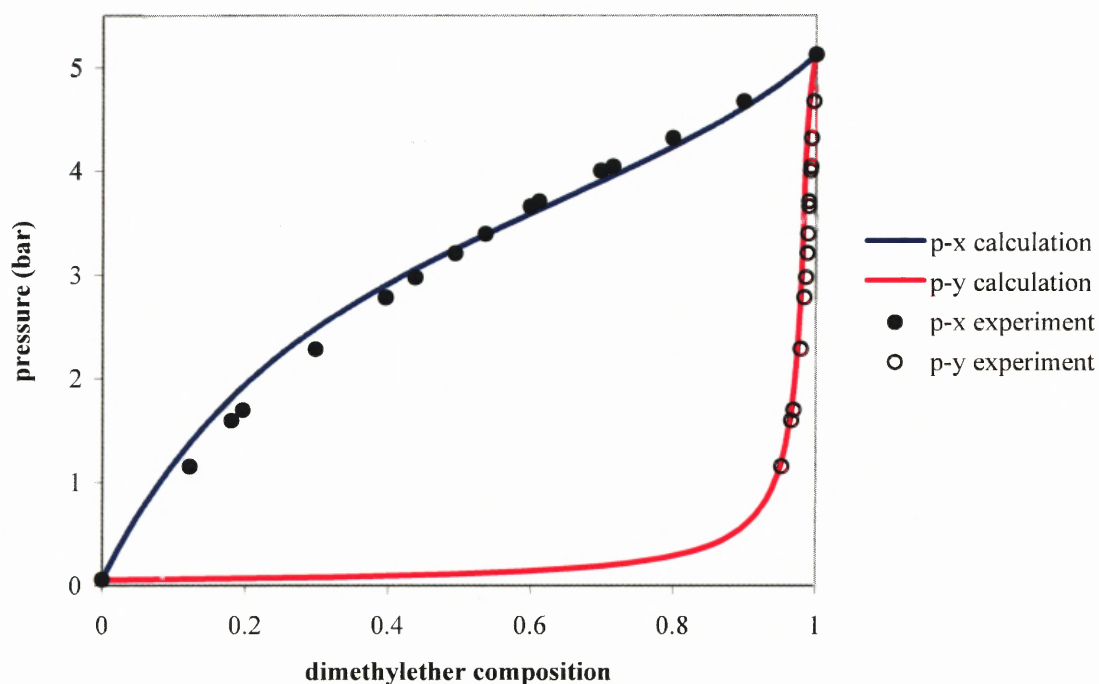
Figure 14.9 VLE for dimethylether/ethanol at 293.15 K: use of experimental liquid volume.

14.5 Alcohol-Alkane Mixture Systems

The prediction of alcohol-alkane mixture systems is a goal of many mixture models, since it offers a clear and common example of non-ideal behavior. It is also a goal to understand the nature of such mixtures, since compounds in these two groups of molecules are regularly found in the chemical industry, and their behaviors are used to model more complex biological, pharmaceutical and industrial systems.

The main problem of describing alcohol-alkane systems is the hydrogen-bonding effects between alcohol molecules. Association effects arising from hydrogen bonding are difficult to describe statistically. Models in the past have allowed aggregation to account for the long-lived complexes in the liquid phase. In the COSMO-based models, an entirely separate correlation parameter is introduced solely in the attempt to describe hydrogen bonding. By introducing such a fix, it suggests a limitation of a quasi-chemical approach in handling such systems.

An attempt is made to calculate the VLE behavior of the binary system 1-propanol/hexane at 313.15 K, depicted in Figure 14.10. As with other systems involving alkanes, the liquid volumes are utilized here. This VLE graph depicts predictions of liquid-phase splitting, and therefore analysis of VLE behavior using this technique would not be possible.

In the analysis of the solution to the quasi-chemical equations for the 1-propanol/hexane system, the model variable that represents the behavior of the hydrogen atoms, Γ_H, increases an order of magnitude in the region of small concentrations of 1-propanol ($\Gamma_H = 0.0156$ at $x_1 = 0.04$ while $\Gamma_H = 0.1522$ at $x_1 = 0.002$). This combined with the value in pure 1-propanol ($\Gamma_H = 0.0007$ at $x_1 = 1$) yields a large enough
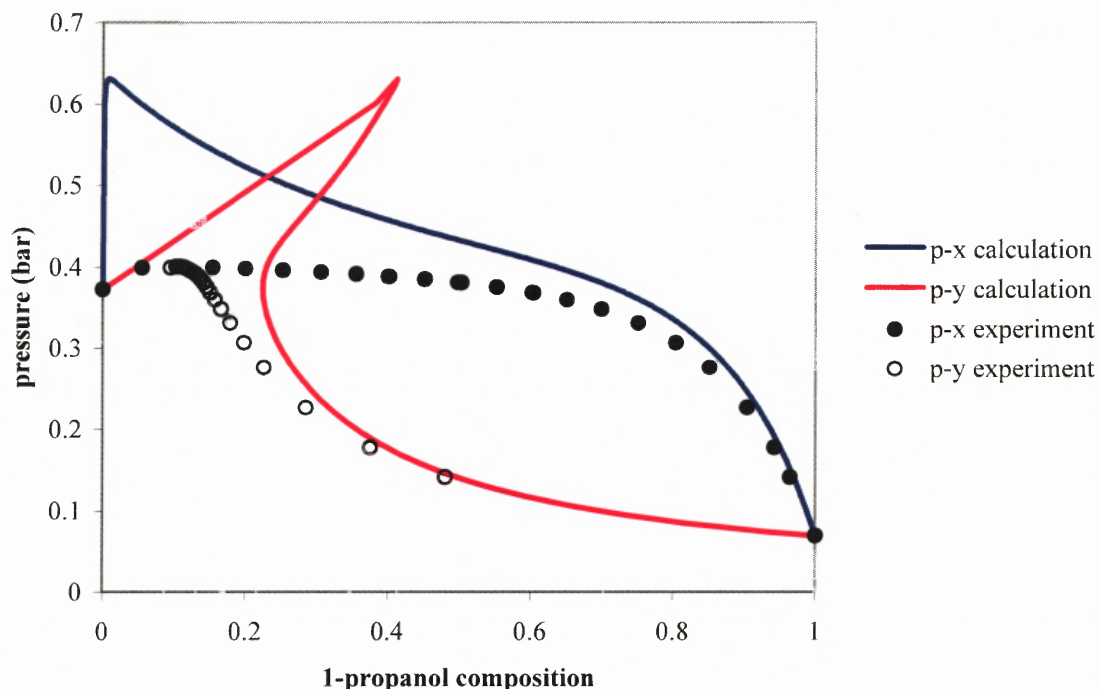
Figure 14.10 VLE for 1-propanol/hexane at 313.15 K.

contribution to the residual activity within Equation (5.51) to give a large jump in the pressure at low 1-propanol concentrations, depicted in Figure 14.10.

This phenomenon can be analyzed further by considering the definition of the model variable Γ_H. Equation (5.50) states the following relationship

$$\Gamma_H = \sqrt{\frac{y_{HH}}{\Theta_H}} \qquad (14.9)$$

where the local composition y_{HH} reflects the number of hydrogen-hydrogen interactions in the system. Considering the interaction energy between two hydrogen atoms (690.5 K), the interaction energy between a hydrogen atom and an oxygen atom (-2896 K) and the number of interactions a hydrogen atom can take part in ($z_H = 0.909$), it is conjectured that the quasi-chemical equations are predicting too many interactions

between hydrogen atoms. A suggested solution at this point is to set the model variable $\Gamma_H = 0$, thus forcing the model to predict zero hydrogen-hydrogen interactions for the pure or mixture systems at any composition.

The results of the fix stated above for the 1-propanol/hexane system at 313.15 K is depicted in Figure 14.11. The VLE behavior depicted here is much more in line with the experimental data. An azeotrope is depicted at low concentrations of 1-propanol near the composition determined through experiment. The general shape of the p-x, p-y envelope also reflects that of the data. The model approaches the experimental data at high 1-propanol concentrations for both the p-x and p-y lines. The fix stated above has taken the model from the completely incorrect prediction from Figure 14.10 to the more sensible result in Figure 14.11.
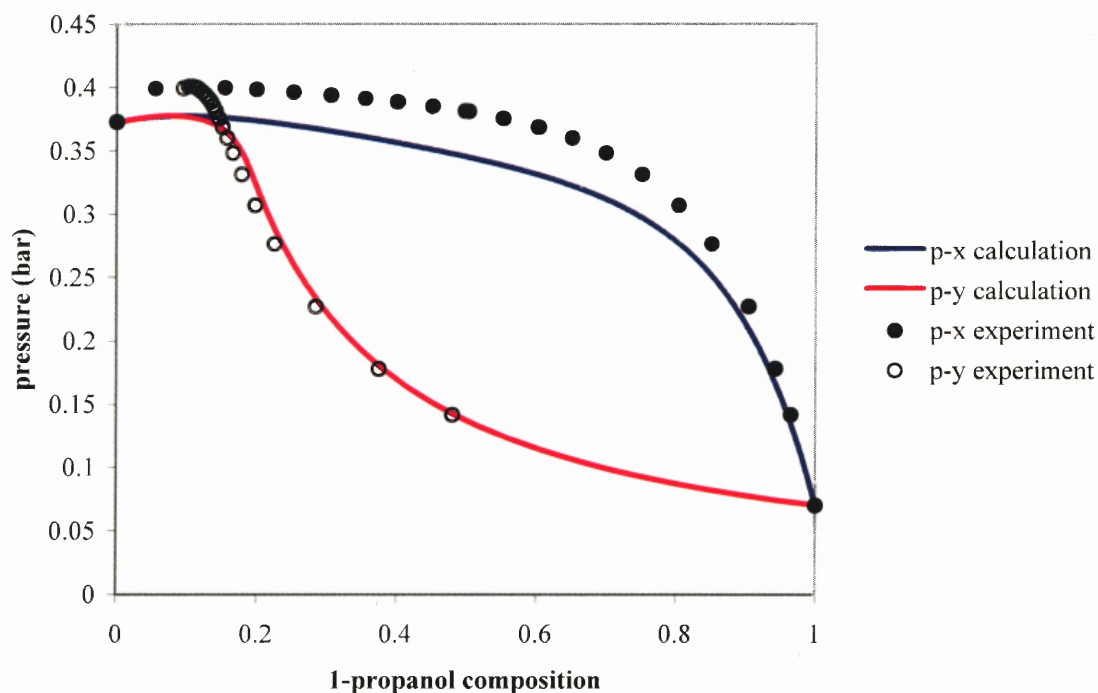
Figure 14.11 VLE for 1-propanol/hexane at 313.15 K: $\Gamma_H = 0$.

The application of setting $\Gamma_H = 0$ is also applied to the calculation of the VLE behavior of ethanol/pentane at a near-critical temperature of 422.6 K, depicted in Figure 14.12. The model again predicts an azeotrope near the composition found in experiment. The model also approaches the behavior of experimental data at high concentrations of the alcohol, while underpredicting the pressures at low concentrations.

By forcing $\Gamma_H = 0$, no contributions from the hydrogen atoms are included in the activity of the alcohol. This in turn reduces the pressure contribution from the alcohol in Equation (14.3). Since the activity of the hydrogen atom is most significant at low alcohol concentrations (suggested by the abnormally high pressures predictions in Figure 14.10), a significant pressure contribution from Γ_H would mostly be seen at these low
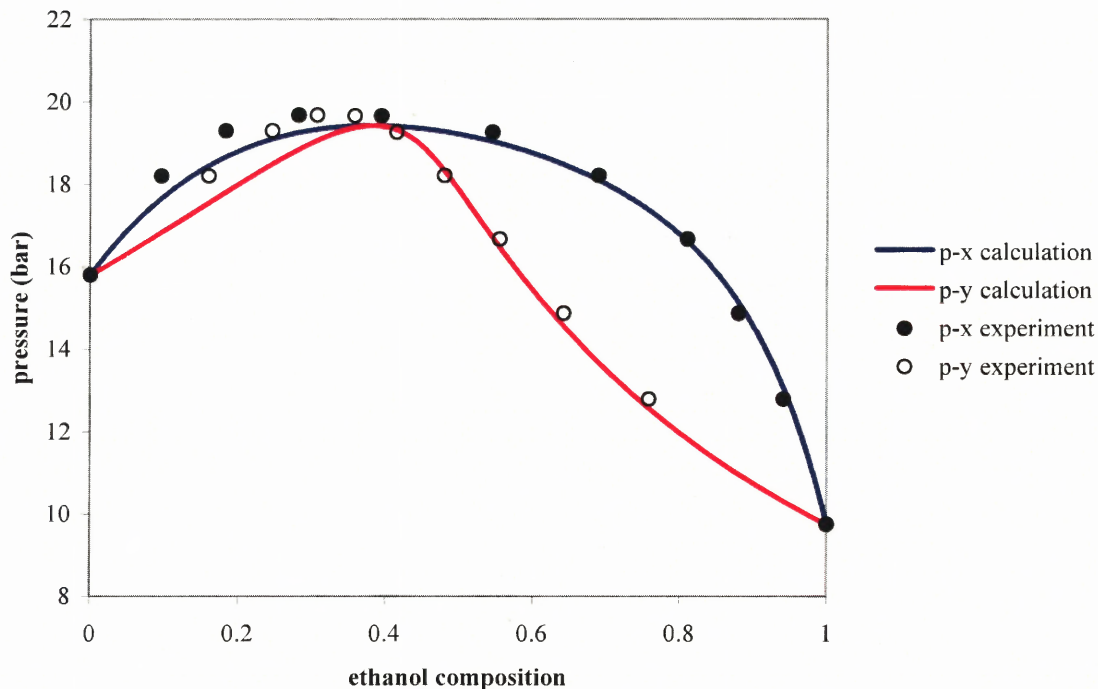
Figure 14.12 VLE for ethanol/pentane at 422.6 K: $\Gamma_H = 0$.

concentrations. This may be why the system pressures are underpredicted on the p-x line at low alcohol concentrations in Figures 14.11 and 14.12. The arbitrary value of $\Gamma_H = 0$ is seen as an overcorrection to the model result. Since statistical models always predict a non-zero solution even though an effect is highly unlikely, a more correct model description through an extension of the quasi-chemical method would likely yield a small, non-zero result for Γ_H at these low concentrations.

14.6 Conclusions

The use of molecular characteristics and energetics from first principles offers a new perspective on the statistical models used to predict fluid-phase behavior. In the past, errors within the modeling technique could not be analyzed, since they would be masked by the use of empirical parameters. Now that these characteristics are describable using first principles, focus can be applied to analyzing the faults of the statistical methodologies.

The use of the quasi-chemical method for the full prediction of VLE behavior has its shortcomings. The method depends on an EoS that is required to predict liquid-like volumes at the mixture system temperature and pressure. It also fails to predict the proper number of interactions within alcohol/alkane systems, thus predicting phase splitting where an azeotrope should exists. Fixes to these problems have been suggested in this work, such as the use of experimental liquid volumes and the arbitrary observation that $\Gamma_H = 0$ in alcohol/alkane mixtures. Now that these more specific problems with the statistical methods have been diagnosed, a more focused effort can be made to fix the quasi-chemical method used in fluid-phase property prediction.

CHAPTER 15

CONCLUSION

15.1 Summary of Contributions

The goal of this work is to create a general model that uses molecular-level properties to predict macroscopic fluid system properties. This motivates research in three areas: molecular and functional group property calculation; applied intermolecular interaction theory; and statistical thermodynamics of lattice-fluid systems. Research has been conducted in these fields, and contributions have been made to achieve an unbroken algorithm toward the goal of the work.

Within the field of molecular and functional group property calculation, the contributions within this work focus on how the methods in computational chemistry can be used in engineering applications. This work demonstrates that molecular-level properties are attainable using modest computational equipment. This has resulted in the calculation of properties for a wide range of molecules and functional groups within the molecules. The transferability assumption is analyzed using AIM computational results, thus giving quantitative evidence of its impossibility. Novel AIM properties, such as the exposed surface area and the polarizability of non-alkyl groups, are calculated for use in this work.

Within the field of intermolecular interactions, the goal of finding a functional group interaction model motivates the combination of rigorous theory with the AIM properties. The extent of the electron density of molecules and functional groups is used to describe short-range repulsive effects. The AIM electrostatic properties are used in

classical expressions to describe long-range effects. These two contributions are combined to form first-principles interaction potentials for small molecules and functional groups. The intermolecular potentials allow for the prediction of second virial coefficients for small molecules. The functional group interaction model yields interaction energies between groups over all interaction ranges. These potentials are comparable to those used in molecular dynamics simulations. The model also predicts a hydrogen-bonding interaction energy and distance in the water dimer system that is comparable to experimental results.

Within the field of lattice-fluid theory, an all-encompassing fluid model is developed that uses the available structural and energetic information from the fields described above. A generalization of the Guggenheim statistical model for athermal systems is derived, so as to utilize the rigorously calculated volumes and surface areas from AIM theory. Local-composition concepts are resolved with vacancies on a lattice to create a lattice-fluid model that may be used in both pure fluid and mixture systems. A closed-form engineering EoS is derived, where the fitting parameters in the equation show physical significance. An excess Gibbs energy is calculated for several ideal and non-ideal mixture systems, and VLE behavior is predicted without the use of any parameters correlated from VLE data.

15.2 Future Work

As with all novel contributions to a research field, questions about assumptions arise as the capabilities of the methods grow more rigorous. Future work within the three areas of research will go a long way toward reaching the goal of a fully predictive fluid model.

Future work in the field of molecular and functional group property calculation involves creating a database of AIM properties and making such information available to engineers. The use of higher level *ab initio* and DFT methods will give molecular properties that are more comparable to experiment. With these results, AIM properties of functional groups will reflect the improvement of the molecular calculations. These properties can then be used in new group-contribution methods which are based on the rigorous properties, likely allowing for a more detailed description of macroscopic properties with a more user-friendly method. To make the information more readily accessible to engineers, a suite of programs can be created to calculate functional group properties merely from the initial guess of the geometry of the molecule.

Future work in the field of intermolecular interactions involves approaching the problem from a more theoretical standpoint, as well as validation of the current methodology. Several expressions that describe the short-range contribution have been presented, and application of the more theoretical expressions can likely improve the model. The use of Slater-type orbitals and *p*-type orbitals within the orbital overlap models will make the model reflect system behavior more rigorously. The inclusion of quadrupole moment information will improve the description of the long-range effects in the intermolecular and functional group interaction models. A suite of programs can be created that considers two molecules and calculates interaction energy curves and an interaction matrix for all the possible group interactions. These curves can also be applied within molecular dynamics simulations in order to validate their effectiveness at macroscopic property prediction.

Future work in the field of lattice-fluid theory involves assessment of the limitations of the quasi-chemical approach and application of new statistics to alleviate these problems. This work has already shown the inability of the statistics to accurately predict critical points and liquid volumes of pure fluids, and the distribution of interaction numbers within nonideal mixture systems. More molecular effects can be included in the model, such as non-nearest-neighbor interactions, proximity effects, and the inclusion of a translational partition function within lattice-fluids. A more realistic EoS can be used within the generalized Guggenheim statistics, such as the Carnahan-Starling equation of state for hard-spheres (Carnahan and Starling, 1969), as opposed to the modified ideal gas law. Studies can be conducted into the approach taken for Ising models and renormalization theory to account for critical behavior. The results from the quasi-chemical equations can be compared to excess Gibbs and excess enthalpy data to assess further the capabilities of the predictive model.

Of great importance to the engineering community is the ease with which information from fundamental studies are made available for use in engineering applications. Great effort can be made to make disseminate the methods in this work using software. This entire work is computation-oriented, and a program that takes an engineer from a molecular structure to macroscopic system properties is envisioned. This would allow for the dissemination and, hopefully, successful use of the new methodologies.

APPENDIX A

GUGGENHEIM STATISTICS
FOR A BINARY SYSTEM WITH VACANCIES

The statistics proposed by Guggenheim (1944a) attempt to determine the number of ways an athermal, lattice-fluid system may be arranged. The method relates two approaches that describe the Helmholtz energy of the system: through the partition function and through the fundamental property relation. The original derivation considers a binary mixture of polymer with monomers, and has been incorporated in numerous engineering models of a lattice-fluid. Here, a generalization to the approach is taken, so as to include vacancies in the lattice initially and to eliminate any assumption about the structure of the molecules on the lattice. The result is a generalized function that reduces to past work.

A.1 The Gibbs Ensemble Approach

Assume the Gibbs partition function can be approximated by the maximum summand. Therefore, for a binary system with vacancies

$$\Delta = f_1^{N_1} f_2^{N_2} \Omega e^{-\frac{W'}{k_B T}} e^{-\frac{pV'}{k_B T}} \tag{A.1}$$

where f_i is the molecular partition function for species i. Considering an athermal system, $W' = 0$ and

$$\Delta = f_1^{N_1} f_2^{N_2} \Omega^{ath} e^{-\frac{pV'}{k_B T}} \tag{A.2}$$

271

Vacancies occur on the lattice, and it is assumed that the partition function of a vacant lattice site, f_0, is unity. Therefore, this term is included within the partition function without effect on the expression

$$\Delta = f_0^{N_0} f_1^{N_1} f_2^{N_2} \Omega^{ath} e^{-\frac{pV^t}{k_B T}} \tag{A.3}$$

The total Gibbs energy of the system is therefore

$$\frac{G^t \left(N_0, N_1, N_2 \right)}{k_B T} = -\ln \Delta = -\sum_{i=0}^{2} N_i \ln f_i - \ln \Omega^{ath} \left(N_0, N_1, N_2 \right) + \frac{pV^t}{k_B T} \tag{A.4}$$

Here, an EOS for an athermal system must be used to relate the pressure to the occupancy of the lattice. Consider a modified ideal gas EoS, where the hard-sphere molecular volumes affect the pressure

$$p \left[V^t - b_0 \left(r_1 N_1 + r_2 N_2 \right) \right] = \left(N_1 + N_2 \right) k_B T \tag{A.5}$$

where the total volume of the system is given by

$$V^t = b_0 \sum_{i=0}^{n} r_i N_i \tag{A.6}$$

Inserting these into Equation (A.4) gives

$$\frac{G^t \left(N_0, N_1, N_2 \right)}{k_B T} = -\sum_{i=0}^{2} N_i \ln f_i - \ln \Omega^{ath} \left(N_0, N_1, N_2 \right) + \frac{N_0 + r_1 N_1 + r_2 N_2}{N_0} \left(N_1 + N_2 \right) \tag{A.7}$$

It is helpful to express this relationship in terms of the total number of lattice sites N_t, the occupation fraction of the molecules on the lattice ω, and the volume fraction ϕ of molecule 1 in the system where no vacancies exist. The definitions of these variables are given by

$$N_t = N_0 + r_1 N_1 + r_2 N_2 \tag{A.8}$$

$$\omega = \frac{r_1 N_1 + r_2 N_2}{N_t} \tag{A.9}$$

$$\phi = \frac{r_1 N_1}{r_1 N_1 + r_2 N_2} \tag{A.10}$$

Rearranging yields

$$N_0 = N_t (1 - \omega) \tag{A.11}$$

$$N_1 = \frac{1}{r_1} N_t \omega \phi \tag{A.12}$$

$$N_2 = \frac{1}{r_2} N_t \omega (1 - \phi) \tag{A.13}$$

Inserting the new variables into Equation (A.7) gives

$$\frac{G^t (N_t, \omega, \phi)}{k_B T} = -N_t (1 - \omega) \ln f_0 - \frac{1}{r_1} N_t \omega \phi \ln f_1 - \frac{1}{r_2} N_t \omega (1 - \phi) \ln f_2$$
$$- \ln \Omega^{ath} (N_t, \omega, \phi) + N_t \frac{\omega}{1 - \omega} \left[\frac{1}{r_1} \phi + \frac{1}{r_2} (1 - \phi) \right] \tag{A.14}$$

This expression will be revisited later, after another expression for the total Gibbs energy of the system is derived.

A.2 The Fundamental Property Relation Approach

Consider the fundamental property relation for the total Gibbs energy of a binary system

$$dG^t = -S^t dT + V^t dp + \mu_1 dN_1 + \mu_2 dN_2 \tag{A.15}$$

Any changes in the thermodynamic variables of the system will be made isothermally. Therefore

$$dG^t = V^t dp + \mu_1 dN_1 + \mu_2 dN_2 \tag{A.16}$$

Vacancies occur on the lattice, and the chemical potential, μ_0, within the athermal system is assumed to be zero. Therefore, the term is included in the expression without changing the differential. Also, rewriting the expression in dimensionless form yields

$$\frac{dG^t}{k_B T} = V^t \frac{dp}{k_B T} + \frac{\mu_0}{k_B T} dN_0 + \frac{\mu_1}{k_B T} dN_1 + \frac{\mu_2}{k_B T} dN_2 \qquad (A.17)$$

An EoS is necessary here to express the differential pressure and the total volume of the system in terms of the occupancy of the lattice. Here too, the modified version of the ideal gas EoS is used, Equation (A.5). The differential pressure given in terms of the numbers of molecules is

$$\frac{dp}{k_B T} = \frac{1}{b_0}\left(\frac{1}{N_0} dN_1 + \frac{1}{N_0} dN_2 - \frac{(N_1 + N_2)}{N_0^2} dN_0\right) \qquad (A.18)$$

Inserting this expression into Equation (A.17) and organizing by the differentials yields

$$\begin{aligned}\frac{dG^t (N_0, N_1, N_2)}{k_B T} &= \left(\ln\gamma_0 - \frac{N_t}{N_0^2}(N_1 + N_2)\right) dN_0 \\ &\quad + \left(\ln\gamma_1 + \frac{N_t}{N_0}\right) dN_1 \qquad (A.19) \\ &\quad + \left(\ln\gamma_2 + \frac{N_t}{N_0}\right) dN_2\end{aligned}$$

where the chemical potential is expressed now through the absolute activity γ

$$\ln\gamma_i = \frac{\mu_i}{k_B T} \qquad (A.20)$$

Here, the variables within Equation (A.19) are to be changed to the variables of Equations (A.8) through (A.10). The differentials of the numbers of vacancies and molecules on the lattice, found using Equations (A.11) through (A.13), are given by

$$dN_0 = (1 - \omega) dN_t - N_t d\omega \qquad (A.21)$$

$$dN_1 = \frac{1}{r_1}\left(\omega\phi dN_t + N_t\phi d\omega + N_t\omega d\phi\right) \qquad (A.22)$$

$$dN_2 = \frac{1}{r_2}\left[\omega\left(1-\phi\right)dN_t + N_t\left(1-\phi\right)d\omega - N_t\omega d\phi\right] \qquad (A.23)$$

Lengthy algebra yields the following expression for the differential total Gibbs energy in terms of the new variables

$$\frac{dG'\left(N_t,\omega,\phi\right)}{k_B T} = \left[\ln\gamma_0 + \frac{\phi\omega}{r_1}\ln\left(\frac{\gamma_1}{\gamma_0^{r_1}}\right) + \frac{(1-\phi)\omega}{r_2}\ln\left(\frac{\gamma_2}{\gamma_0^{r_2}}\right)\right]dN_t$$
$$+N_t\left[\begin{array}{c}\dfrac{\phi}{r_1\left(1-\omega\right)^2}+\dfrac{1-\phi}{r_2\left(1-\omega\right)^2}\\[3mm]+\dfrac{\phi}{r_1}\ln\left(\dfrac{\gamma_1}{\gamma_0^{r_1}}\right)+\dfrac{(1-\phi)}{r_2}\ln\left(\dfrac{\gamma_2}{\gamma_0^{r_2}}\right)\end{array}\right]d\omega \qquad (A.24)$$
$$+N_t\omega\left[\begin{array}{c}\dfrac{1}{r_1 r_2}\ln\left(\dfrac{\gamma_1^{r_2}}{\gamma_2^{r_1}}\right)\\[3mm]+\dfrac{1}{r_1\left(1-\omega\right)}-\dfrac{1}{r_2\left(1-\omega\right)}\end{array}\right]d\phi$$

The remainder of the derivation is only concerned with changes in the overall occupancy of the lattice ω. Therefore, the other differential expressions are neglected, simplifying the expression to

$$\frac{dG'\left(N_t,\omega,\phi\right)}{k_B T} = N_t\left[\begin{array}{c}\dfrac{\phi}{r_1\left(1-\omega\right)^2}+\dfrac{1-\phi}{r_2\left(1-\omega\right)^2}\\[3mm]+\dfrac{\phi}{r_1}\ln\left(\dfrac{\gamma_1}{\gamma_0^{r_1}}\right)+\dfrac{(1-\phi)}{r_2}\ln\left(\dfrac{\gamma_2}{\gamma_0^{r_2}}\right)\end{array}\right]d\omega \qquad (A.25)$$

Here, let α_i express the ratio of absolute activities to the molecular partition functions

$$\alpha_i = \frac{\gamma_i f_i}{\left(\gamma_0 f_0\right)^{r_i}} \qquad (A.26)$$

The reason for this will be apparent later, when this Gibbs energy is compared with that from the partition function approach. Inserting this into Equation (A.25) yields

$$\frac{dG^t\left(N_t,\omega,\phi\right)}{k_BT} = N_t \begin{bmatrix} \dfrac{\phi}{r_1\left(1-\omega\right)^2} + \dfrac{1-\phi}{r_2\left(1-\omega\right)^2} \\[2mm] +\ln f_0 - \dfrac{\phi}{r_1}\ln f_1 - \dfrac{\left(1-\phi\right)}{r_2}\ln f_2 \\[2mm] +\dfrac{\phi}{r_1}\ln\alpha_1 + \dfrac{\left(1-\phi\right)}{r_2}\ln\alpha_2 \end{bmatrix} d\omega \qquad (A.27)$$

The indefinite integral of this expression is given by

$$\frac{G^t\left(N_t,\omega,\phi\right)}{k_BT} = N_t \begin{bmatrix} \dfrac{\phi}{r_1\left(1-\omega\right)} + \dfrac{1-\phi}{r_2\left(1-\omega\right)} \\[2mm] +\omega\ln f_0 - \dfrac{\phi\omega}{r_1}\ln f_1 - \dfrac{\left(1-\phi\right)\omega}{r_2}\ln f_2 \\[2mm] +\dfrac{\phi}{r_1}\int\ln\alpha_1 d\omega + \dfrac{1-\phi}{r_2}\int\ln\alpha_2 d\omega \end{bmatrix} + C\left(N_t,\phi\right) \qquad (A.28)$$

A.3 Equating the Approaches

Next step is to consider the difference of the Gibbs energies, to apply limits to the indefinite integrals in Equation (A.28). Consider the process going from a system of N_t lattice sites with no occupancy to the same size system with the desired occupancy ω and volume fraction ϕ

$$\frac{\Delta G^t}{k_BT} = \frac{G^t\left(N_t,\omega,\phi\right)}{k_BT} - \frac{G^t\left(N_t,0,\phi\right)}{k_BT} \qquad (A.29)$$

For the partition function approach, plugging Equation (A.14) into Equation (A.29) yields

$$\frac{\Delta G^t}{k_B T} = N_t \left[\frac{\phi \omega}{r_1 (1-\omega)} + \frac{(1-\phi)\omega}{r_2 (1-\omega)} + \omega \ln f_0 - \frac{\phi \omega}{r_1} \ln f_1 - \frac{(1-\phi)\omega}{r_2} \ln f_2 \right]$$
$$- \ln \left(\frac{\Omega^{ath} (N_t, \omega, \phi)}{\Omega^{ath} (N_t, 0, \phi)} \right)$$

(A.30)

For the FPR approach, plugging Equation (A.28) into Equation (A.29) yields

$$\frac{\Delta G^t}{k_B T} = N_t \left[\frac{\phi \omega}{r_1 (1-\omega)} + \frac{(1-\phi)\omega}{r_2 (1-\omega)} + \omega \ln f_0 - \frac{\phi \omega}{r_1} \ln f_1 - \frac{(1-\phi)\omega}{r_2} \ln f_2 \right]$$
$$+ \frac{\phi N_t}{r_1} \int_0^\omega \ln \alpha_1 d\omega + \frac{(1-\phi) N_t}{r_2} \int_0^\omega \ln \alpha_2 d\omega$$

(A.31)

Equating these yields

$$\ln \left(\frac{\Omega^{ath} (N_t, \omega, \phi)}{\Omega^{ath} (N_t, 0, \phi)} \right) = -N_t \left[\frac{\phi}{r_1} \int_0^\omega \ln \alpha_1 d\omega + \frac{(1-\phi)}{r_2} \int_0^\omega \ln \alpha_2 d\omega \right]$$

(A.32)

Since the number of ways a system of N_t lattice sites with no molecules occupying the lattice is one, then $\Omega^{ath} (N_t, 0, \phi) = 1$ and

$$\ln \Omega^{ath} (N_t, \omega, \phi) = -N_t \left[\frac{\phi}{r_1} \int_0^\omega \ln \alpha_1 d\omega + \frac{(1-\phi)}{r_2} \int_0^\omega \ln \alpha_2 d\omega \right]$$

(A.33)

After evaluating α_i, the athermal ways function is solved for.

A.4 The α Integrals

The concept of α_i in the derivation of the Guggenheim statistics is that it represents the ratio of probabilities of placing molecules on the lattice. The explanation of the quantity is available within the original derivation of the statistics (Guggenheim, 1952).

In the numerator is the probability that a group of r_1 sites is entirely occupied by a molecule of species 1. This is given by

$$P_1 = \varpi_1 \frac{N_1}{N_0 + r_1 N_1 + r_2 N_2} \tag{A.34}$$

where ϖ_1 holds symmetry factors that describe the total number of ways a single molecule of species 1 can occupy a given r_1 group of sites. The denominator of the ratio is the probability that those same r_1 sites are completely unoccupied (occupied by vacancies). This is given by

$$P_0 = \frac{N_0}{N_0 + r_1 N_1 + r_2 N_2} \left(\frac{z_0 N_0}{z_0 N_0 + z_1 N_1 + z_2 N_2} \right)^{r_1 - 1} \tag{A.35}$$

This probability is the product of finding one vacancy at the first of the r_1 lattice sites and the probability of finding vacancies along each of the remainder of the $r_1 - 1$ sites. The definition of α_1 from Equation (A.26) is a ratio of these probabilities

$$\alpha_1 = \frac{\gamma_1 f_1}{\left(\gamma_0 f_0 \right)^{r_1}} = \frac{P_1}{P_0} = \varpi_1 \frac{N_1}{N_0} \left(\frac{z_0 N_0 + z_1 N_1 + z_2 N_2}{z_0 N_0} \right)^{r_1 - 1} \tag{A.36}$$

Similarly for α_2

$$\alpha_2 = \frac{\gamma_2 f_2}{\left(\gamma_0 f_0 \right)^{r_2}} = \frac{P_2}{P_0} = \varpi_2 \frac{N_2}{N_0} \left(\frac{z_0 N_0 + z_1 N_1 + z_2 N_2}{z_0 N_0} \right)^{r_2 - 1} \tag{A.37}$$

In the reduced coordinates, these expressions are found to be

$$\alpha_1 \left(N_t, \omega, \phi \right) = \varpi_1 \frac{\omega \phi}{r_1 \left(1 - \omega \right)} \left[1 + \frac{q_1}{r_1} \frac{\omega \phi}{1 - \omega} + \frac{q_2}{r_2} \frac{\omega \left(1 - \phi \right)}{1 - \omega} \right]^{r_1 - 1} \tag{A.38}$$

$$\alpha_2 \left(N_t, \omega, \phi \right) = \varpi_2 \frac{\omega \phi}{r_2 \left(1 - \omega \right)} \left[1 + \frac{q_1}{r_1} \frac{\omega \phi}{1 - \omega} + \frac{q_2}{r_2} \frac{\omega \left(1 - \phi \right)}{1 - \omega} \right]^{r_2 - 1} \tag{A.39}$$

Here, it is convenient to define q_i as the ratio of contacts between molecule i and a vacant site. This is given by

$$q_i = \frac{z_i}{z_0} \qquad (A.40)$$

This differs from the original expressions from Guggenheim (1944, 1952). The original expressions require a relationship between r_i and q_i, namely

$$\frac{1}{2}z_0\left(r_i - q_i\right) = r_i - 1 \qquad (A.41)$$

This relationship will not be used in the present derivation, since here it is assumed that the volume and surface area of a molecule can be found rigorously using other techniques.

A.5 Athermal Ways for the System

Returning back to Equation (A.33), the athermal ways of the system is now able to be evaluated by using Equations (A.36) and (A.37). Algebraic manipulation is required to achieve the final result. An intermediate result for Ω^{ath} is given by an expression found just prior to converting natural logarithm functions back into factorials using Stirling's approximation. This is given by

$$-\ln\Omega^{ath}\left(N_t, \omega, \phi\right) = N_t \frac{\phi\omega}{r_1}\left(\ln\varpi_1 - \ln r_1\right) + N_t \frac{(1-\phi)\omega}{r_2}\left(\ln\varpi_2 - \ln r_2\right)$$

$$+ N_t \frac{\phi\omega}{r_1}\ln\phi + N_t \frac{(1-\phi)\omega}{r_2}\ln(1-\phi)$$

$$+ N_t\left(\frac{\phi\omega}{r_1} + \frac{(1-\phi)\omega}{r_2}\right)\left(\ln\omega - 1\right) \qquad (A.42)$$

$$+ N_t\left(1-\omega\right)\left[\ln\left(1-\omega\right) - 1\right] + N_t$$

$$+ N_t\left(\begin{array}{c} \dfrac{\phi(r_1-1)}{r_1} \\ + \dfrac{(1-\phi)(r_2-1)}{r_2} \end{array}\right)\int_0^\omega \ln\left\{1 - \omega\left[1 - \frac{q_1}{r_1}\phi - \frac{q_2}{r_2}(1-\phi)\right]\right\}d\omega$$

Once the integral in the final line is evaluated and the expression is reverted back into factorial form, the expression for the number of ways an athermal binary mixture with vacancies can be arranged is found to be

$$\Omega^{ath} = \varpi_1^{N_1} \varpi_2^{N_2} \frac{N_t!}{N_0! N_1! N_2!} \left[\frac{(N_0 + q_1 N_1 + q_2 N_2)!}{(N_0 + r_1 N_1 + r_2 N_2)!} \right]^{z_0 \kappa / 2} \tag{A.43}$$

where the exponent is given by

$$\frac{z_0 \kappa}{2} = \frac{(r_1 - 1) N_1 + (r_2 - 1) N_2}{(r_1 - q_1) N_1 + (r_2 - q_2) N_2} \tag{A.44}$$

A general expression is inferred from this result, given by

$$\Omega^{ath} = \prod_{i=1}^{n} \varpi_i^{N_i} \frac{N_t!}{\prod_{i=0}^{n} N_i!} \left(\frac{Q_t!}{N_t!} \right)^{z_0 \kappa / 2} \tag{A.45}$$

where q_i is given by Equation (A.40),

$$N_t = \sum_{i=0}^{n} r_i N_i \tag{A.46}$$

$$Q_t = \sum_{i=0}^{n} q_i N_i \tag{A.47}$$

and where the exponent is given by

$$\frac{z_0 \kappa}{2} = \frac{\sum_{i=0}^{n} (r_i - 1) N_i}{\sum_{i=0}^{n} (r_i - q_i) N_i} \tag{A.48}$$

It can be seen here that if Equation (A.41) were satisfied, then κ within these expressions must be unity. This applied to a system without vacancies results in the athermal ways function in the previous Guggenheim derivations.

APPENDIX B

GENERALIZED QUASI-CHEMICAL APPROACH

The quasi-chemical method is used within many modern engineering models of fluid-phase behavior. Its derivation exists in many places, from the originating methods of Guggenheim (1944b) through to the parallel derivation offered by the COSMO-based methods (Lin and Sandler, 2002; Klamt, et al., 2002). This method is rederived here, utilizing the concepts of local composition and the notation of Knox and coworkers (1984, 1987). Vacancies are included on the lattice initially, and a variable substitution that leads to the reduction of these equations to that of the COSMO-based methods is offered.

B.1 Model Equations

The equations for a lattice-fluid are written to determine the following model variables: numbers and types of nearest-neighbor interactions M_{kl}; the interaction distances between nearest neighbors r_{kl}; and the number of vacancies on the lattice N_0. Also, for convenience of notation, let $N_0 = M_0$. Assume there are n different species in the system, and there are m distinct types of functional groups within the molecules. All the summations without bounds in the following derivation include vacancies with all the molecule and functional groups.

The total number of interactions a functional group takes part in is related to the interaction numbers by the relation

$$\frac{z_k M_k}{2} = \sum_l M_{kl} \quad \text{for } k = 0, 1, \ldots n \tag{B.1}$$

For the entire system, let the total number of interactions of the system be represented by I, given by

$$I = \sum_k \frac{z_k M_k}{2} = \sum_k \sum_l M_{kl} \tag{B.2}$$

Also necessary in the quasi-chemical approach is the equality between interaction numbers and interaction distances with those quantities with opposite indices. Therefore, the following relationships must hold

$$M_{kl} = M_{lk} \tag{B.3}$$

$$r_{kl} = r_{lk} \tag{B.4}$$

With these model equations and the partition function for the lattice-fluid, the macroscopic system properties are definable through the model variables. The solution method involves finding the most probable set of interaction numbers and interaction distances.

B.2 Partition Function Approach

The general partition function for a lattice-fluid is given by

$$\Delta(T,p,N_i) = \Delta_{internal}(T,p,N_i) \sum_{W^t} \sum_{V^t} \Omega(T,p,N_i;W^t,E^t) e^{-\frac{pV^t}{k_B T}} e^{-\frac{W^t}{k_B T}} \tag{B.5}$$

Let τ_Δ stand for the summand

$$\tau_\Delta(T,p,N_i;V^t,W^t) = \Omega(T,p,N_i;V^t,W^t) e^{-\frac{pV^t}{k_B T}} e^{-\frac{W^t}{k_B T}} \tag{B.6}$$

An acceptable assumption for systems far away from the critical region is that the maximum of τ_Δ with respect to W^t and V^t represents the entire summation.

$$\Delta\left(T,p,N_i\right) = \Delta_{internal}\left(T,p,N_i\right)\tau_\Delta\left(T,p,N_i;\left(V^t\right)^*,\left(W^t\right)^*\right) \qquad (B.7)$$

The total volume and the total energy of the system are explicit in the model variables. These relations are given by

$$V^t = \sum_i N_i b_i + \sum_k \sum_l M_{kl} v_{kl}\left(r_{kl}\right) \qquad (B.8)$$

$$E^t = \sum_k \sum_l M_{kl} u_{kl}\left(r_{kl}\right) \qquad (B.9)$$

where v_{kl} is the volume of interaction and u_{kl} is the energy of interaction evolved between an interaction of groups k and l at a distance r_{kl}. The partition function now is expressed through the modeling variables

$$\Delta\left(T,p,N_i\right) = \Delta_{internal}\left(T,p,N_i\right)\tau_\Delta\left(T,p,N_i;N_0^*,M_{kl}^*,r_{kl}^*\right) \qquad (B.10)$$

The task now is to find the maximum of τ_Δ with respect to the model variables.

Taking the natural logarithm of Equation (B.6) and expanding

$$\ln\tau_\Delta = \ln\Omega - \frac{p}{k_B T}\left(\sum_i N_i b_i + \sum_k \sum_l M_{kl} v_{kl}\left(r_{kl}\right)\right) - \frac{1}{k_B T}\sum_k \sum_l M_{kl} u_{kl}\left(r_{kl}\right) \qquad (B.11)$$

Here, define η_{kl} as the enthalpy of interaction, a combination of the energy effects and the volumetric effects

$$\eta_{kl}\left(r_{kl}\right) = u_{kl}\left(r_{kl}\right) + p v_{kl}\left(r_{kl}\right) \qquad (B.12)$$

This simplifies Equation (B.11) to

$$\ln\tau_\Delta = \ln\Omega - \frac{p}{k_B T}\sum_i N_i b_i - \frac{1}{k_B T}\sum_k \sum_l M_{kl}\eta_{kl}\left(r_{kl}\right) \qquad (B.13)$$

Not all the M_{kl} and r_{kl} within this expression are independent variables, since $M_{kl} = M_{lk}$ and $r_{kl} = r_{lk}$. This derivation will proceed by using M_{kl} and r_{kl} where

$0 < k \leq l$ as the independent variables. Lengthy algebra on Equation (B.13) yields the partition function with only independent variables

$$\ln \tau_\Delta = \ln \Omega - N_0 \left[\frac{pb_0}{k_B T} + \frac{z_0}{2} \frac{\eta_{00}}{k_B T} \right] - \sum_{i=1}^{n} N_i \left[\frac{pb_i}{k_B T} + \frac{z_i}{2} \left(2\frac{\eta_{0i}}{k_B T} - \frac{\eta_{00}}{k_B T} \right) \right]$$
$$- 2 \sum_{k=1}^{m-1} \sum_{l=k+1}^{m} M_{kl} \left(\frac{\eta_{kl}}{k_B T} + \frac{\eta_{00}}{k_B T} - \frac{\eta_{0k}}{k_B T} - \frac{\eta_{0l}}{k_B T} \right) - \sum_{s=1}^{m} M_{ss} \left(\frac{\eta_{ss}}{k_B T} + \frac{\eta_{00}}{k_B T} - 2\frac{\eta_{0s}}{k_B T} \right) \quad (B.14)$$

B.3 Maximizing Conditions

The values of the model variables are found by finding the extrema within Equation (B.14). The derivatives that must be evaluated are

$$\left(\frac{\partial \ln \tau_\Delta}{\partial N_0} \right)_{T,p,M_{kl},r_{kl}} = 0 \quad (B.15)$$

$$\left(\frac{\partial \ln \tau_\Delta}{\partial M_{kl}} \right)_{T,p,N_0,r_{kl}} = 0 \text{ for } 0 < k \leq l \quad (B.16)$$

$$\left(\frac{\partial \ln \tau_\Delta}{\partial r_{kl}} \right)_{T,p,N_0,M_{kl}} = 0 \text{ for } 0 < k \leq l \quad (B.17)$$

Equation (B.15) results in the following expression

$$\left(\frac{\partial \ln \Omega}{\partial N_0} \right)_{T,p,M_{kl},r_{kl}} - \frac{pb_0}{k_B T} - \frac{z_0}{2} \frac{\eta_{00}}{k_B T} = 0 \quad (B.18)$$

This expression is usually referred to as the EoS, since the partial derivative of $\ln \Omega$ with respect to N_0 usually results in an expression explicit in system volume, thereby yielding an equation that related p, T, and V'.

Equation (B.16) is dealt with in two parts. First, for the off-diagonal interactions, where $k < l$

$$0 = \left(\frac{\partial \ln \Omega}{\partial M_{kl}}\right)_{T,p,M_{ss},N_0,r_{kl}} - 2\left(\frac{\eta_{kl}}{k_B T} + \frac{\eta_{00}}{k_B T} - \frac{\eta_{0k}}{k_B T} - \frac{\eta_{0l}}{k_B T}\right) \qquad (B.19)$$

Next, for the diagonal terms, where $k = l = s$

$$0 = \left(\frac{\partial \ln \Omega}{\partial M_{ss}}\right)_{T,p,M_{kl},N_0,r_{kl}} - \left(\frac{\eta_{ss}}{k_B T} + \frac{\eta_{00}}{k_B T} - 2\frac{\eta_{0s}}{k_B T}\right) \qquad (B.20)$$

Equation (B.17) is also considered for off-diagonal and diagonal terms, similar to the last two derivatives. For the off-diagonal terms, only those where $k > 0$ is considered, since it will be assumed that $\eta_{0l} = 0$ for all l (there are no energetic or volumetric consequences in an interaction between two vacancies or a group and a vacancy). Therefore, for $0 < k < l$

$$0 = \left(\frac{\partial \ln \Omega}{\partial r_{kl}}\right)_{T,p,r_{ss},N_0,M_{kl}} - 2M_{kl}\left(\frac{\partial(\eta_{kl}/k_B T)}{\partial r_{kl}}\right)_{T,p,r_{ss},N_0,M_{kl}} \qquad (B.21)$$

Similarly for diagonal terms

$$0 = \left(\frac{\partial \ln \Omega}{\partial r_{ss}}\right)_{T,p,r_{kl},N_0,M_{kl}} - M_{ss}\left(\frac{\partial(\eta_{ss}/k_B T)}{\partial r_{ss}}\right)_{T,p,r_{kl},N_0,M_{kl}} \qquad (B.22)$$

Now the derivatives of the ways function Ω must be evaluated.

B.4 The Ways Function

It is assumed that Ω is the product of a naïve combinatorial formula and a proportionality factor

$$\Omega = K \frac{I!}{\prod_k \prod_l M_{kl}!} \qquad (B.23)$$

This relation is normalized to the number of ways an athermal system can be arranged. The proportionality factor is found to be (Knox, et al., 1984; Knox, 1987)

$$K = K\left(M_k, N_i\right) = \Omega^{ath}\left(M_0, N_i\right)\left[\prod_k \left(\frac{z_k M_k}{2}\right)! \Big/ I!\right]^2 \qquad \text{(B.24)}$$

Combining the last two expressions, one finds that

$$\Omega\left(M_{kl}, N_i\right) = \Omega^{ath}\left(M_0, N_i\right)\frac{\left[\prod_k \left(\frac{z_k M_k}{2}\right)!\right]^2}{I!\prod_k\prod_l M_{kl}!} \qquad \text{(B.25)}$$

After applying Stirling's approximation, $\ln\Omega$ is given by

$$\ln\Omega = \ln\Omega^{ath} + 2\sum_k\left(\frac{z_k M_k}{2}\ln\frac{z_k M_k}{2}\right) - I\left(M_0\right)\ln I\left(M_0\right)$$
$$-\sum_k\sum_l\left(M_{kl}\ln M_{kl}\right) \qquad \text{(B.26)}$$

Before evaluating the necessary derivatives with Equation (B.26), the expression should be represented only in terms of the independent model variables. The derivation henceforth assumes that the independent variables have the subscripts $0 < k \le l$. The evaluations of the derivatives of $\ln\Omega$ with respect to r_{kl} is straightforward, since it is seen within Equation (B.25) that the ways function is independent of all r_{kl}. Therefore, Equations (B.21) and (B.22) are easily computed, and will be revisited later.

Writing the ways function in terms of the remaining independent variables is somewhat difficult, and the necessary derivatives in Equation (B.19) and (B.20) are found implicitly. Start by rewriting the double summation within Equation (B.26) to give

$$\ln \Omega = \ln \Omega^{ath} + 2\sum_{k}\left(\frac{z_k M_k}{2}\ln\frac{z_k M_k}{2}\right) - I\ln I$$

$$-\left(\begin{array}{c} M_{00}\ln M_{00} + 2\sum_{k=1}^{m}\left(M_{0k}\ln M_{0k}\right) \\ +2\sum_{k=1}^{m-1}\sum_{l=k+1}^{m}\left(M_{kl}\ln M_{kl}\right) + \sum_{s=1}^{m}\left(M_{ss}\ln M_{ss}\right) \end{array}\right) \qquad (B.27)$$

The task now is to find the dependence of all the M_{k0} and M_{00} on the independent model variables. For the former, consider the expression

$$M_{0q} = \frac{z_q M_q}{2} - \sum_{p=1}^{m} M_{pq} \qquad (B.28)$$

Attempting to isolate the independent variables within the expression, one finds that

$$M_{0q} = \frac{z_q M_q}{2} - \sum_{p=1}^{q-1} M_{pq} - M_{qq} - \sum_{p=q+1}^{m} M_{qp} \qquad (B.29)$$

All the numbers of interactions within this expression are independent variables. Taking the derivative with respect to a general M_{kl}, where $k < l$

$$\left(\frac{\partial M_{0q}}{\partial M_{kl}}\right)_{M_{ss}} = -\delta_{kq} - \delta_{lq} \qquad (B.30)$$

where the δ here is the Kronecker delta. For the diagonal element

$$\left(\frac{\partial M_{0q}}{\partial M_{ss}}\right)_{M_{kl}} = -\delta_{sq} \qquad (B.31)$$

Equation (B.29) is independent of M_0, therefore

$$\left(\frac{\partial M_{0q}}{\partial M_0}\right)_{M_{kl},M_{ss}} = 0 \qquad (B.32)$$

At this point, it is time to consider M_{00}. Following the example starting at Equation (B.28), consider the expression

$$M_{00} = \frac{z_0 M_0}{2} - \sum_{p=1}^{m} M_{p0} \tag{B.33}$$

Taking the derivative with respect the off-diagonal element

$$\left(\frac{\partial M_{00}}{\partial M_{kl}} \right)_{M_{ss}, M_0} = -\sum_{p=1}^{m} \left(\frac{\partial M_{p0}}{\partial M_{kl}} \right)_{M_{ss}, M_0} = 1 \tag{B.34}$$

For the diagonal element

$$\left(\frac{\partial M_{00}}{\partial M_{ss}} \right)_{M_{kl}, M_0} = -\sum_{p=1}^{m} \left(\frac{\partial M_{p0}}{\partial M_{ss}} \right)_{M_{kl}} = 2 \tag{B.35}$$

Finally, with respect to the number of vacancies

$$\left(\frac{\partial M_{00}}{\partial M_0} \right)_{M_{kl}, M_{ss}} = \frac{z_0}{2} \tag{B.36}$$

At this point, one may revisit the derivatives of $\ln \Omega$. When taking the derivative with respect to the off-diagonal elements M_{kl}, one finds that

$$\left(\frac{\partial \ln \Omega}{\partial M_{kl}} \right)_{T, p, M_0, M_{ss}, r_{kl}} = 2 \ln \left(\frac{M_{k0} M_{l0}}{M_{00} M_{kl}} \right) \tag{B.37}$$

Taking the derivative with respect to the diagonal element, one finds that

$$\left(\frac{\partial \ln \Omega}{\partial M_{ss}} \right)_{T, p, M_0, M_{kl}, r_{kl}} = \ln \left(\frac{M_{0s} M_{s0}}{M_{00} M_{ss}} \right) = \ln \left(\frac{M_{0s}^2}{M_{00} M_{ss}} \right) \tag{B.38}$$

Finally, taking the derivative with respect to the number of vacancies yields

$$\left(\frac{\partial \ln \Omega}{\partial M_0} \right)_{T, p, M_{kl}, r_{kl}} = \left(\frac{\partial \ln \Omega^{ath}}{\partial M_0} \right)_{N_i} + \frac{z_0}{2} \ln \left(\frac{(z_0 M_0 / 2)^2}{M_{00} I} \right) \tag{B.39}$$

This expression now only depends on the choice of Ω^{ath}.

B.5 The Quasi-Chemical Equations in Reduced Variables

Now that the derivatives of $\ln\Omega$ have been found, the conditions where the maximum of τ_Δ is achieved can be summarized. Firstly, the trivial case of the r_{kl} , Equations (B.21) and (B.22), is considered. Since Ω is independent of interaction distance, these expressions lead the conclusion that

$$\left(\frac{\partial\left(\eta_{kl}/k_BT\right)}{\partial r_{kl}}\right)_{T,p,M_0,M_{kl}} = 0 \tag{B.40}$$

for $0 < k \leq l$. This implies that functional groups, on average, are interacting at the minimum of η_{kl} with the interaction energy of u_{kl} and with the volumetric effect of pv_{kl}. With regards to the number of holes in the system, the EoS is given by

$$\left(\frac{\partial\ln\Omega^{ath}}{\partial M_0}\right)_{T,p,N_i} + \frac{z_0}{2}\left[\ln\left(\frac{\left(z_0M_0/2\right)^2}{M_{00}I}\right) - \frac{\eta_{00}}{k_BT}\right] - \frac{pb_0}{k_BT} = 0 \tag{B.41}$$

Again, the explicit EoS can be found after choosing the appropriate athermal ways function.

The numbers of interactions for the off-diagonal elements are governed by the relation

$$\frac{M_{k0}M_{l0}}{M_{00}M_{kl}} = \exp\left[-\left(\frac{\eta_{0k}}{k_BT} + \frac{\eta_{0l}}{k_BT} - \frac{\eta_{kl}}{k_BT} - \frac{\eta_{00}}{k_BT}\right)\right] \tag{B.42}$$

The diagonal terms are given by

$$\frac{M_{0s}^2}{M_{00}M_{ss}} = \exp\left[-\left(2\frac{\eta_{0s}}{k_BT} - \frac{\eta_{ss}}{k_BT} - \frac{\eta_{00}}{k_BT}\right)\right] \tag{B.43}$$

These terms can be combined to give the familiar interchange energy of the quasi-chemical equations

$$\frac{M_{kl}M_{lk}}{M_{kk}M_{ll}} = \exp\left[\frac{-\left(2\eta_{kl} - \eta_{kk} - \eta_{ll}\right)}{k_B T}\right] \tag{B.44}$$

This relation is true for combinations of $k \geq 0$ and $l \geq 0$. It is also the case that equations like Equation (B.44) are used to calculate the actual maximizing variables, and thus, the solution to the quasi-chemical equations.

It is at this point useful to concisely express the results of the derivations as a set of non-linear expressions. The equations include the EoS, Equation (B.41), the relations for interaction distances, Equation (B.40), the quasi-chemical relations, Equation (B.44), and the fundamental model relations, Equations (B.1) through (B.4).

Within these expressions, define a set of dimensionless relations that simplify the derivation results even further. Let the local composition of l groups around k groups be defined as

$$y_{kl} = \frac{M_{kl}}{z_k M_k / 2} \tag{B.45}$$

Let the surface area fraction of group k be defined as

$$\Theta_k = \frac{z_k M_k / 2}{I} \tag{B.46}$$

With these relations and the equations from the derivation, the non-linear system of equations determining the model variables is give by

$$\left(\frac{\partial \ln \Omega^{ath}}{\partial N_0}\right)_{T,p,N_i} + \frac{z_0}{2}\left[\ln\left(\frac{\Theta_0}{y_{00}}\right) - \frac{\eta_{00}}{k_B T}\right] - \frac{pb_0}{k_B T} = 0 \tag{B.47}$$

$$\Theta_k y_{kl} = \Theta_l y_{lk}, \left(k,l = 1,2,...,m\right) \tag{B.48}$$

$$\sum_l y_{kl} = 1, \left(k = 0,1,...,m\right) \tag{B.49}$$

$$\frac{y_{kl}y_{lk}}{y_{kk}y_{ll}} = \tau_{kl}^2, \left(k,l = 0,1,...,m\right) \tag{B.50}$$

with τ_{kl} defined as

$$\tau_{kl} = \exp\left(-\frac{2\eta_{kl} - \eta_{kk} - \eta_{ll}}{2k_B T}\right) \tag{B.51}$$

This system of equations is the generalization of the quasi-chemical equations, where in the past vacancies were not included on the lattice, and interaction volumes always vanished, leaving $\eta_{kl} = \varepsilon_{kl}$.

A significant simplification of the model equations can be made to relate this model to that employed by the COSMO-based methods. Define Γ_k as a quantity related to the activity of functional group k through the relation

$$\Gamma_k = \sqrt{y_{kk}/\Theta_k} \tag{B.52}$$

Using this relation, Equations (B.47) through (B.50) reduce significantly, to the system described by

$$\left(\frac{\partial \ln \Omega^{ath}}{\partial N_0}\right)_{T,p,N_i} - \frac{z_0}{2}\left[\ln \Gamma_0^2 - \frac{\eta_{00}}{k_B T}\right] - \frac{pb_0}{k_B T} = 0 \tag{B.53}$$

$$\frac{1}{\Gamma_k} = \sum_l \Theta_l \tau_{kl} \Gamma_l, \left(k = 0,1,...,m\right) \tag{B.54}$$

This latter result is given for COSMO-based models, while the EoS is necessary to describe vacancies on the lattice.

APPENDIX C

INPUT FILES FOR GAUSSIAN 98W

C.1 Full Input File

The following input file contains all five sections outlined in Section 6.1: header section, route section, title, Z-matrix and the wavefunction file name. A geometry optimization is performed on ethanamide (CH_3CONH_2), and a .wfn file is given as an output. If a .wfn file is not desired, the keyword "output=wfn" and the final two lines of the input file may be omitted.

```
%chk=ethanamide.chk
%mem=400MB
%nproc=1
#opt rmp2(full)/aug-cc-pvdz density output=wfn

OPT calculation of ethanamide at mp2 level

0  1
 C1
 O2       1    B1
 C3       1    B2    2    A1
 H4       3    B3    1    A2    2    D1
 H5       3    B4    1    A3    2    D2
 H6       3    B5    1    A4    2    D3
 N7       1    B6    2    A5    3    D4
 H8       7    B7    1    A6    2    D5
 H9       7    B8    1    A7    2    D6

     B1              1.231687
     B2              1.511214
     B3              1.089296
     B4              1.087833
     B5              1.084815
     B6              1.371706
     B7              1.007534
     B8              1.005152
     A1            122.740081
     A2            109.047689
     A3            112.200378
     A4            108.574315
     A5            121.926423
     A6            116.973255
     A7            120.925002
     D1             90.076546
```

```
D2              -149.414353
D3               -27.879112
D4              -179.722359
D5               -11.048318
D6              -167.113695
```

ethanamide.wfn

C.2 Input File using .chk Checkpoint File

The following input file utilizes an already existing checkpoint file to read the title, charge, multiplicity and Z-matrix for the molecule, outlined in Section 6.4. Note the use of the "geom=allcheck" and "guess=read" keywords listed in the route section. A single point energy calculation is performed on ethanamide (CH_3CONH_2), and a .wfn file is given as an output. If a .wfn file is not desired, the keyword "output=wfn" and the final two lines of the input file may be omitted.

```
%chk=ethanamide.chk
%mem=400MB
%nproc=1
#hf/6-31++G(d,p) geom=allcheck guess=read
#scf=tight output=wfn

ethanamide.wfn
```

C.3 Output File from Gaussian 98W

The output file from Gaussian 98W listed below includes the following: common beginning lines; coordinates of the nuclei; statement of method used in analysis of the electron density; the electrostatic properties of the molecule; and the AIM properties. The listing of AIM properties includes: location of attractors; location of critical points; electrostatic properties; and energetic properties. Each of the sections described here is separated by a '+ + + +' line below.

```
Entering Gaussian System, Link 0=g98
 Input=hfluoride3j3o.speaim.txt
 Output=hfluoride3j3o.speaim.out
 Initial command:
 l1.exe "E:\G98W\gxx.inp" "hfluoride3j3o.speaim.out" /scrdir="E:\G98W\"
 Entering Link 1 = l1.exe PID=      664.

 Copyright (c) 1988,1990,1992,1993,1995,1998 Gaussian, Inc.
                  All Rights Reserved.

 This is part of the Gaussian(R) 98 program.  It is based on
 the Gaussian 94(TM) system (copyright 1995 Gaussian, Inc.),
 the Gaussian 92(TM) system (copyright 1992 Gaussian, Inc.),
 the Gaussian 90(TM) system (copyright 1990 Gaussian, Inc.),
 the Gaussian 88(TM) system (copyright 1988 Gaussian, Inc.),
 the Gaussian 86(TM) system (copyright 1986 Carnegie Mellon
 University), and the Gaussian 82(TM) system (copyright 1983
 Carnegie Mellon University). Gaussian is a federally registered
 trademark of Gaussian, Inc.

 ++++

                        Z-Matrix orientation:
 ---------------------------------------------------------------------
 Center     Atomic     Atomic              Coordinates (Angstroms)
 Number     Number      Type          X           Y           Z
 ---------------------------------------------------------------------
    1          9          0        0.000000    0.000000    0.092619
    2          1          0        0.000000    0.000000   -0.833570
 ---------------------------------------------------------------------

 ++++
 *********************************************************************


           Population analysis using the MP2 density.

 *********************************************************************


 ++++


 Total atomic charges:
                 1
     1  F   -0.280735
     2  H    0.280735
 Sum of Mulliken charges=   0.00000
 Atomic charges with hydrogens summed into heavy atoms:
                 1
     1  F    0.000000
     2  H    0.000000
 Sum of Mulliken charges=   0.00000
 Electronic spatial extent (au):  <R**2>=     14.4297
 Charge=      0.0000 electrons
 Dipole moment (Debye):
     X=     0.0000    Y=     0.0000    Z=    -1.8175  Tot=      1.8175
 Quadrupole moment (Debye-Ang):
    XX=     -5.9620   YY=    -5.9620   ZZ=    -3.7762
    XY=      0.0000   XZ=     0.0000   YZ=     0.0000
 Octapole moment (Debye-Ang**2):
    XXX=     0.0000  YYY=     0.0000  ZZZ=    -1.8087  XYY=      0.0000
    XXY=     0.0000  XXZ=    -0.1093  XZZ=     0.0000  YZZ=      0.0000
    YYZ=    -0.1093  XYZ=     0.0000
 Hexadecapole moment (Debye-Ang**3):
 XXXX=     -4.4001 YYYY=    -4.4001 ZZZZ=    -3.9085 XXXY=      0.0000
 XXXZ=      0.0000 YYYX=     0.0000 YYYZ=     0.0000 ZZZX=      0.0000
 ZZZY=      0.0000 XXYY=    -1.4667 XXZZ=    -1.6428 YYZZ=     -1.6428
 XXYZ=      0.0000 YYXZ=     0.0000 ZZXY=     0.0000
 N-N= 5.142141874930D+00 E-N=-2.498608080782D+02  KE =9.992363305921D+01
```

++++

```
*********************************************************************

          Properties of atoms in molecules using the MP2 density.

*********************************************************************
```

I. ATTRACTORS

| Attr. | Cartesian Coordinates | | | Nucleus | Density | |
	X	Y	Z	(Distance)	Total	Spin
1	0.000000	0.000000	0.175020	F (0.000005)	0.42851E+03	0.00000E+00
2	0.000000	0.000000	-1.465056	H (0.110163)	0.36197E+00	0.00000E+00

II. CRITICAL POINTS ON ATTRACTOR INTERACTION LINES

| Line | Attractors | | Cartesian Coordinates | | | Density | |
	A	B	X	Y	Z	Total	Spin
1	2	1	0.000000	0.000000	-1.282010	0.35102E+00	0.00000E+00

++++

III. PROPERTIES OF ATTRACTORS

| Attr. | Number of electrons | | Charge |
	total	spin	
1	9.720001	0.000000	-0.720001
2	0.280083	0.000000	0.719917
Total	10.000084	0.000000	-0.000084

| Attr. | Kinetic energy | | | Dipole moment | | |
	G	K	L	X	Y	Z
1	0.996488E+02	0.996499E+02	0.112E-02	0.000000	0.000000	0.420704
2	0.274697E+00	0.274281E+00	-0.416E-03	0.000000	0.000000	0.045586
Total	0.999235E+02	0.999242E+02	0.703E-03	0.000000	0.000000	0.466290

| Attr. | Traceless quadrupole moment | | | | | |
	XX	YY	ZZ	XY	XZ	YZ
1	-0.149910	-0.149910	0.299820	0.000000	0.000000	0.000000
2	0.062823	0.062823	-0.125646	0.000000	0.000000	0.000000
Total	-0.087087	-0.087087	0.174174	0.000000	0.000000	0.000000

++++

Normal termination of Gaussian 98.

APPENDIX D

WAVEFUNCTION OUTPUT

A wavefunction file for water at the MP2/3-21G level and basis set is presented. The .wfn file yields the characteristics of the wavefunction approximation, and when calculated properly, will give the computed electron density at any point relative to the position of the nuclei. The electron density at a point is found through the summation of the electron density contributed by each orbital ψ^{MO} multiplied by the occupation of that orbital ν

$$\rho(\mathbf{r}) = \sum_{l=1}^{n_{MO}} \nu_l \left[\psi_l^{MO}(\mathbf{r}) \right]^2 \tag{D.1}$$

where n_{MO} is the total number of orbitals used in the calculation, and is not necessarily equal to half the number of electrons (to yield two electrons within each orbital). Here, the orbitals used with the calculation are molecular orbitals (MOs), which have contributions centered on all nuclei within the system. Also, the occupation number ν of an orbital is two for HF calculations and not necessarily two for methods based on higher theory, since more states are available for electrons to inhabit in methods that include electron correlation.

The MOs are further described by GTFs and with number given by the basis set chosen for calculation. A given MO consists of n_{PRIM} primitives of the form of a GTF which, for AIM calculations, can take the form of s-type orbitals ($i + j + k = 0$), p-type orbitals ($i + j + k = 1$), d-type orbitals ($i + j + k = 2$), or f-type orbitals ($i + j + k = 3$). In general, a MO is represented as

$$\psi_l^{MO}(\mathbf{r}) = \sum_{m=1}^{n_{PRIM}} c_{l,m} \chi_m(\mathbf{r}) \tag{D.2}$$

where $c_{l,m}$ gives the coefficient that multiplies the mth primitive within the lth orbital.

Each GTF within the wavefunction file has a different exponent and orbital center, and all the information is given within the .wfn file.

```
Title Card Required
GAUSSIAN              13 MOL ORBITALS     21 PRIMITIVES       3 NUCLEI
   O    1   (CENTRE  1)  0.00000000  0.00000000  0.20940408  CHARGE =  8.0
   H    2   (CENTRE  2)  0.00000000  1.48149962 -0.83761633  CHARGE =  1.0
   H    3   (CENTRE  3)  0.00000000 -1.48149962 -0.83761633  CHARGE =  1.0
CENTRE ASSIGNMENTS     1  1  1  1  1  1  1  1  1  1  1  1  1  1  2  2  2  3  3
CENTRE ASSIGNMENTS     3
TYPE ASSIGNMENTS       1  1  1  1  1  2  2  3  3  4  4  1  2  3  4  1  1  1  1  1
TYPE ASSIGNMENTS       1
EXPONENTS  0.3220370D+03 0.4843080D+02 0.1042060D+02 0.7402940D+01 0.1576200D+01
EXPONENTS  0.7402940D+01 0.1576200D+01 0.7402940D+01 0.1576200D+01 0.7402940D+01
EXPONENTS  0.1576200D+01 0.3736840D+00 0.3736840D+00 0.3736840D+00 0.3736840D+00
EXPONENTS  0.5447178D+01 0.8245472D+00 0.1831916D+00 0.5447178D+01 0.8245472D+00
EXPONENTS  0.1831916D+00
MO   1     MO 0.0         OCC NO =     2.0000059  ORB. ENERGY =     0.0000000
   0.31343219D+01  0.44912805D+01  0.28565813D+01 -0.13214659D+00  0.12510020D+00
  -0.19801657D-13 -0.99991662D-14 -0.39659636D-15 -0.20026773D-15 -0.14882765D-01
  -0.75152926D-02 -0.62013784D-02 -0.23982472D-14  0.13217562D-16  0.26597630D-02
   0.19085423D-02  0.26811292D-02  0.14149938D-02  0.19085423D-02  0.26811292D-02
   0.14149938D-02
MO   2     MO 0.0         OCC NO =     1.9923686  ORB. ENERGY =     0.0000000
  -0.86268546D+00 -0.12361724D+01 -0.78624059D+00 -0.29709090D+00  0.28124928D+00
  -0.90766252D-13 -0.45833884D-13 -0.36232990D-15 -0.18296433D-15  0.50984632D+00
   0.25745513D+00  0.27525592D+00 -0.10659108D-13 -0.33074145D-16  0.60634827D-01
   0.19865283D-01  0.27906844D-01 -0.55570495D-02  0.19865283D-01  0.27906844D-01
  -0.55570495D-02
MO   3     MO 0.0         OCC NO =     1.9873068  ORB. ENERGY =     0.0000000
  -0.10796407D-13 -0.15470553D-13 -0.98397087D-14 -0.14258108D-13  0.13497830D-13
   0.22142760D+01  0.11181344D+01 -0.61019419D-15 -0.30812740D-15 -0.42764437D-13
  -0.21594593D-13  0.10398403D-13  0.26370166D+00 -0.56189151D-16 -0.44858988D-14
   0.30082167D-14  0.42259569D-14  0.58554825D-15  0.30299521D-14  0.42564910D-14
   0.55746521D-15
MO   4     MO 0.0         OCC NO =     1.9817530  ORB. ENERGY =     0.0000000
   0.92828368D-01  0.13301704D+00  0.84602598D-01  0.35023256D-01 -0.33155729D-01
   0.73838435D-13  0.37285910D-13  0.17917952D-15  0.90479592D-16  0.18757167D+01
   0.94717342D+00  0.21486354D-01  0.88486590D-14  0.66894373D-17  0.20607671D+00
  -0.66170870D-01 -0.92957149D-01 -0.22498740D-01 -0.66170870D-01 -0.92957149D-01
  -0.22498740D-01
MO   5     MO 0.0         OCC NO =     1.9788228  ORB. ENERGY =     0.0000000
  -0.57947634D-17 -0.83035210D-17 -0.52812739D-17 -0.93353628D-16  0.88375781D-16
   0.44031146D-15  0.22234238D-15  0.17232217D+01  0.87016861D+00 -0.11145607D-15
  -0.56281542D-16  0.49124612D-16  0.55483603D-16  0.14562428D+00 -0.94271143D-17
   0.93478726D-01  0.13131935D+00  0.37757062D-01 -0.93478726D-01 -0.13131935D+00
  -0.37757062D-01
MO   6     MO 0.0         OCC NO =     0.0187608  ORB. ENERGY =     0.0000000
   0.12138664D-16  0.17393920D-16  0.11063025D-16 -0.10886717D-14  0.10306210D-14
   0.21347783D-15  0.10779907D-15  0.35120093D+01  0.17734458D+01  0.16072792D-14
   0.81162159D-15 -0.24043637D-15  0.20839420D-16  0.49441379D-01 -0.81774318D-16
  -0.13200753D+00 -0.18544479D+00 -0.82447623D-01  0.13200753D+00  0.18544479D+00
   0.82447623D-01
```

```
MO    7    MO 0.0         OCC NO =     0.0172343  ORB. ENERGY =   0.0000000
    0.30670183D+00   0.43948388D+00   0.27952416D+00   0.25162149D+00  -0.23820441D+00
   -0.20111260D-13  -0.10155505D-13  -0.22056340D-14  -0.11137705D-14   0.35081256D+01
    0.17714846D+01  -0.15706575D+00   0.18840350D-14   0.29293024D-15  -0.89695821D-01
    0.12852689D+00   0.18055519D+00   0.57220123D-01   0.12852689D+00   0.18055519D+00
    0.57220123D-01
MO    8    MO 0.0         OCC NO =     0.0120382  ORB. ENERGY =   0.0000000
    0.22775760D-15   0.32636191D-15   0.20757539D-15  -0.76110701D-14   0.72052289D-14
   -0.43843706D+01  -0.22139587D+01  -0.15794759D-16  -0.79758185D-17  -0.40075608D-13
   -0.20236825D-13  -0.45948960D-14   0.40166082D+00   0.34871437D-17   0.45819365D-14
   -0.16232121D-14  -0.22802960D-14   0.20361049D-14  -0.16214247D-14  -0.22777851D-14
    0.20391541D-14
MO    9    MO 0.0         OCC NO =     0.0073579  ORB. ENERGY =   0.0000000
    0.41612575D+00   0.59628127D+00   0.37925173D+00   0.14239074D+01  -0.13479811D+01
    0.14404074D-13   0.72735694D-14  -0.54352495D-15  -0.27446170D-15  -0.21748004D+01
   -0.10982006D+01   0.24796304D+00  -0.12685824D-14   0.57955343D-15   0.30279138D+00
    0.93644475D-01   0.13155220D+00   0.15866831D-03   0.93644475D-01   0.13155220D+00
    0.15866831D-03
MO   10    MO 0.0         OCC NO =     0.0020107  ORB. ENERGY =   0.0000000
    0.34823045D-15   0.49899171D-15   0.31737282D-15   0.98094875D-14  -0.92864213D-14
   -0.11953610D-15  -0.60361681D-16   0.18487725D+01   0.93356750D+00   0.84486974D-14
    0.42663061D-14   0.31849857D-14  -0.12250648D-16  -0.45271402D+00  -0.24616399D-14
    0.35793192D+00   0.50282445D+00  -0.79784526D-01  -0.35793192D+00  -0.50282445D+00
    0.79784526D-01
MO   11    MO 0.0         OCC NO =     0.0016872  ORB. ENERGY =   0.0000000
   -0.95515979D-02  -0.13686822D-01  -0.87052052D-02  -0.10502437D+01   0.99424211D+00
   -0.75935607D-14  -0.38344910D-14   0.15189197D-13   0.76700302D-14  -0.10042892D+01
   -0.50713204D+00  -0.36207041D+00   0.68935776D-15  -0.41602276D-14   0.30856594D+00
    0.37508012D+00   0.52691432D+00  -0.56118237D-01   0.37508012D+00   0.52691432D+00
   -0.56118237D-01
MO   12    MO 0.0         OCC NO =     0.0004158  ORB. ENERGY =   0.0000000
    0.45421399D-15   0.65085926D-15   0.41396487D-15   0.35230268D-14  -0.33351702D-14
   -0.13433725D-15  -0.67835765D-16   0.26747545D+01   0.13506604D+01   0.31757604D-14
    0.16036515D-14   0.13424691D-14  -0.10658274D-16  -0.54118385D+00  -0.64618900D-15
   -0.92885616D-01  -0.13048615D+00   0.26040258D+00   0.92885616D-01   0.13048615D+00
   -0.26040258D+00
MO   13    MO 0.0         OCC NO =     0.0002379  ORB. ENERGY =   0.0000000
    0.24140963D-01   0.34592438D-01   0.22001768D-01  -0.11612109D+01   0.10992923D+01
    0.50236340D-13   0.25367651D-13   0.54033323D-14   0.27284999D-14  -0.17025948D+01
   -0.85975272D+00  -0.60877104D+00  -0.45845399D-14  -0.95570172D-15   0.33398091D+00
   -0.60380842D-01  -0.84823291D-01   0.19696365D+00  -0.60380842D-01  -0.84823291D-01
    0.19696365D+00
END DATA
 THE   HF ENERGY =     -75.585809975842 THE VIRIAL(-V/T)=    2.00183768
```

APPENDIX E

TABLES OF MOLECULAR PROPERTIES

The table of molecular properties contains data for dipole moments ("Dipole Moments", 2004), molecular polarizabilities (Miller, 2004), and ionization potentials (Lias, 2004).

Table E.1 Molecular Properties: Alkanes

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
methane	74-82-8				2.43	2.593		12.61
ethane	74-84-0				4.18	4.47		11.56
propane	74-98-6	0.082	0.091	0.084	5.96	6.29		10.95
butane	106-97-8				7.77	8.2		10.53
g-butane	106-97-8-conf	0.086	0.096		7.69			
isobutane	75-28-5	0.129	0.144	0.132	7.76		8.14	(10.57)
pentane	109-66-0	0.081	0.090		9.57	9.99		10.28
g-pentane	109-66-0-conf	0.062	0.067		9.51			
isopentane	78-78-4	0.093	0.104	0.13	9.50			10.32
neopentane	463-82-1				9.56		10.2	(10.2)
hexane	110-54-3				11.44	11.9		10.13
g-hexane	110-54-3-conf	0.085	0.092		11.31			
2-methylpentane	107-83-5	0.132	0.147	0.1	11.33			(10.12)
3-methylpentane	96-14-0	0.083	0.092		11.26			(10.08)
2,2-dimethylbutane	75-83-2	0.054	0.063		11.25			(10.06)
2,3-dimethylbutane	79-29-8	0.000	0.000	0.2	11.23			(10.02)

Table E.2 Molecular Properties: Alkenes

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
ethene	74-85-1				4.02	4.252		10.5138
propene	115-07-1	0.410	0.370	0.366	5.88	6.26		9.86
1-butene	106-98-9	0.452	0.373	0.438	7.73	8.27		9.55
trans-2-butene	624-64-6	0.000	0.000		7.79	8.49		9.10.
cis-2-butene	590-18-1	0.246	0.260	0.253	7.67			9.11
2-methylpropene	115-11-7	0.592	0.502	0.503	7.70	8.29		9.239
1-pentene	109-67-1	0.433	0.382	0.4	9.58		9.65	9.51
trans-2-pentene	646-04-8	0.042	0.050		9.63		9.84	9.04
cis-2-pentene	627-20-3		0.293		9.52		9.84	9.01
2-methyl-2-butene	513-35-9		0.197		9.54			8.69
3-methyl-1-butene	563-45-1		0.318	0.32	9.55			9.52
1-hexene	592-41-6	0.501	0.420	0.4	11.43		11.65	9.44

Table E.3 Molecular Properties: Amines

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
methanamine	74-89-5	1.541	1.325	1.31	3.79	4.01		(8.80)
ethanamine	75-04-7	1.533	1.323	1.22	5.60	7.10.		8.86
t-1-propanamine	107-10-8	1.600	1.406	1.17	7.39	9.20.	7.7	(8.78)
g-1-propananmine	107-10-8-conf	1.454	1.283		7.31			
2-propanamine	75-31-0	1.438	1.293	1.19	7.41		7.77	(8.72)
t-1-butanamine	109-73-9	1.572	1.378	1.3	9.20	13.5		8.7
g-1-butanamine	109-73-9-conf	1.417	1.243		9.12			
2-butanamine	13952-84-6	1.417	1.238	[1.28]	9.13			8.46
2-methyl-1-propanamine	78-81-9	1.429	1.266	1.2	9.11			8.50.
2-methyl-2-propanamine	75-64-9	1.370	1.222	[1.29]	9.17			8.46
1-pentanamine	100-58-7	1.626	1.439		11.06			

Table E.4 Molecular Properties: Diamines

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
dimethylamine	124-40-3	1.228	1.067	1.01	5.60	6.37		8.24
methylethylamine	624-78-2	1.162	1.018		7.45			
diethylamine	109-89-7	1.100	0.974	0.92	9.27	10.2	9.61	7.85
methyl-n-propanamine	627-35-0	1.118	0.981		9.22			
methylisopropanamine	4747-21-1	1.101	0.977		9.13			

Table E.5 Molecular Properties: Triamines

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
trimethylamine	75-50-3	0.849	0.655	0.612	7.49	8.15		7.82
n,n-dimethyl-ethylamine	598-56-1	0.798	0.616		9.35			
methyldiethylamine	616-39-7	0.753	0.582		11.10			

Table E.6 Molecular Properties: Nitriles

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
ethanenitrile	75-05-8	4.323	3.936	3.92519	4.31	4.4	4.48	12.2
propanenitrile	107-12-0	4.438	4.029	4.05	6.15	6.7	6.24	11.84
butanenitrile	109-74-0	4.559	4.153	3.91	7.94	8.4		(11.2)
2-methylpropanenitrile	78-82-0	4.481	4.042	4.29	7.88		8.05	(11.3)

Table E.7 Molecular Properties: Alcohols

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
methanol	67-56-1	2.062	1.705	1.7	3.10	3.23	3.32	10.85
ethanol	64-17-5	1.928	1.619	1.69	4.89	5.41	5.11	10.43
g-1-propanol	71-23-8	1.821	1.534	1.58	6.64	6.74		10.18
t-1-propanol	71-23-8-conf	1.820	1.528	1.55	6.68			
2-propanol	67-63-0	1.938	1.669	1.58	6.67	7.61	6.97	10.17
g-1-butanol	71-36-3	1.806	1.521	1.66	8.45	8.88		9.99
t-1-butanol	71-36-3-conf	1.874	1.595		8.45			
2-butanol	78-92-2	1.970	1.710	1.7	8.41			9.88
2-methyl-1-propanol	78-83-1	1.720	1.446	1.64	8.43	8.92		10.02
2-methyl-2-propanol	75-65-0	1.880	1.634	1.7	8.44			9.90.
g-1-pentanol	71-41-0	1.797	1.515	1.7	10.26			(10.00)
t-1-pentanol	71-41-0-conf	1.789	1.505		10.34			
2-methyl-2-butanol	75-85-4	1.782	1.541	1.9	10.14			(9.8)
3-methyl-1-butanol	123-51-3	1.732	1.448		10.19			
1,2 propanediol	57-55-6	3.083	2.687	[2.25]	7.32			
1,3 propanediol	504-63-2	2.647	2.255	[2.55]	7.35			
1-methoxy-2-propanol	107-98-2	2.823	2.397	2.36	9.20			

Table E.8 Molecular Properties: Ethers

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
dimethyl ether	115-10-6	1.693	1.355	1.30.	4.91	5.29		10.025
methyl ethyl ether	540-67-0	1.571	1.261	1.17	6.80	7.93		9.72
diethyl ether	60-29-7	1.463	1.175	1.15	8.59	8.73		9.51
methyl propyl ether	557-17-5	1.484	1.183	1.107	8.54		8.86	9.41
methyl isopropyl ether	598-53-8	1.582	1.311	1.247	8.48			9.45
methyl tert-butyl ether	1634-04-4	1.554	1.311		10.15			(9.24)
ethyl tert-butyl ether	637-92-3	1.467	1.242		12.00			

Table E.9 Molecular Properties: Aldehydes

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
methanal	50-00-0	3.061	2.430	2.332	2.57	2.8	2.45	10.88
ethanal	75-07-0	3.450	2.834	2.75	4.43		4.59	10.229
propanal	123-38-6	3.301	2.711	2.72	6.13	6.5		9.96
butanal	123-72-8	3.218	2.635	2.72	7.94	8.2		9.84
2-methylpropanal	78-84-2	3.383	2.787	2.69	7.90			9.71

Table E.10 Molecular Properties: Ketones

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
propanone	67-64-1	3.621	3.009	2.88	6.18	6.33	6.39	9.703
butanone	78-93-3	3.445	2.852	2.779	7.91	8.13		9.52
2-pentanone	107-87-9	3.336	2.745	2.7	9.71	9.93		9.38
3-pentanone	96-22-0	3.275	2.702	2.8	9.61	9.93		9.31
3-methyl-2-butanone	563-80-4	3.473	2.874	2.5	9.62			9.30.
methyl vinyl ketone	78-94-3	3.324	2.769		8.00			

Table E.11 Molecular Properties: Carboxylic Acids

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
methanoic	64-18-6	1.733	1.465	1.425	3.32	3.4		11.33
ethanoic	64-19-7	2.038	1.690	1.7	5.05	5.1		10.65
propanoic	79-09-4	1.894	1.551	1.75	6.78	6.9		10.525
butanoic	107-92-6	1.824	1.477	1.5	8.59		8.58	10.17
2-methylpropanoic	79-31-2	1.986	1.634	1.3	8.54			10.33

Table E.12 Molecular Properties: Esters

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
methyl methanoate	107-31-3	2.061	1.853	1.77	5.06		5.05	10.835
methyl ethanoate	79-20-9	2.097	1.808	1.72	6.83	6.94	6.81	10.25
ethyl methanoate	109-94-4	2.289	2.105	1.93	6.90	8.01	6.88	10.61
methyl propanoate	554-12-1	1.931	1.653	1.7	8.59		8.97	10.15
ethyl ethanoate	141-78-6	2.193	1.943	1.78	8.69	9.7	8.62	10.01
propyl methanoate	110-74-7	2.360	2.179	1.9	8.65			10.52
propyl ethanoate	109-60-4	2.321	2.048	[1.78]	10.42			(9.92)
methyl acrylate	96-33-3	1.695	1.517	[1.77]	8.73			(9.9)
vinyl ethanoate	108-05-4	2.068	1.764	[1.79]	8.95			9.19
methyl butanoate	623-42-7	1.842	1.579		10.43		10.41	(10.07)
ethyl propanoate	105-37-3	2.078	1.817	[1.74]	10.41		10.41	(10.00)
butyl methanoate	592-84-7	2.361	2.181	[2.03]	10.49			10.52

Table E.13 Molecular Properties: Fluorides

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
fluoromethane	593-53-3	2.351	1.951	1.858	2.45	2.97		12.47
fluoroethane	353-36-6	2.403	2.051	1.937	4.23	4.96		(11.78)
g-1-fluoropropane	460-13-9	2.307	1.974	1.9	5.97			(11.3)
t-1-fluoropropane	460-13-9-conf	2.450	2.112	2.05	5.98			
2-fluoropropane	420-26-8	2.397	2.069	1.958	5.98			(11.08)
g-1-fluorobutane	2366-52-1	2.246	1.913		7.76			
t-1-fluorobutane	2366-52-1-conf	2.523	2.188		7.75			
2-fluorobutane		2.294	1.979		7.74			
2-methyl-1-fluoropropane		2.301	1.971		7.72			
2-methyl-2-fluoropropane	353-61-7	2.368	2.052		7.74			

Table E.14 Molecular Properties: Amides

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
methanamide	75-12-7	4.451	3.895	3.73	4.15	4.2	4.08	10.16
ethanamide	60-35-5	4.402	3.820	3.68	5.85		5.67	9.65
propanamide	79-05-0	4.262	3.685		7.59			
butanamide	541-35-5	4.085	3.542		9.32			
2-methylpropanamide	563-83-7	4.183	3.656		9.35			

Table E.15 Molecular Properties: Nitros

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
nitromethane	75-52-5	4.388	3.505	3.46	4.83	7.37		11.08
nitroethane	79-24-3	4.569	3.689	3.23	6.56	9.63		10.88
1-nitropropane	108-03-2	4.516	3.670	3.66	8.25	8.5		(10.81)
2-nitropropane	79-46-9	4.653	3.768	3.73	8.27			(10.71)

Table E.16 Molecular Properties: Inorganics

molecule name	CAS number	μ (D)			α (nm^3)			I (eV)
		OPT	SPE	exp	SPE	exp	calc	
carbon monoxide	630-08-0	0.438	0.243	0.1098	1.96	1.95		14.014
carbon dioxide	124-38-9				2.64	2.911		13.773
hydrogen fluoride	7664-39-3	2.088	1.818	1.826178	0.74		0.8	16.044
fluorine	7782-41-4				1.08		1.38	15.697
molecular hydrogen	1333-74-0				0.75	0.8059		15.42593
water	7732-18-5	2.276	1.857	1.8546	1.37	1.45		12.6206
ammonia	7664-41-7	1.813	1.468	1.4718	2.05	2.81		10.070.
nitrogen	7727-37-9				1.73	1.7403		15.5808
nitrous oxide	10024-97-2	0.594	0.125	0.16083	2.97	3.03		12.886
neon	7440-01-9				0.30			21.56454
molecular oxygen	7782-44-7				1.31	1.5812		12.0692
hydrogen cyanide	74-90-8	3.342	3.017	2.985188	2.47	2.59		13.60.

APPENDIX F

INPUT AND OUTPUT FILES FOR PROAIMV

F.1 Input for Basic Integration

```
hexane3j3o
   C     2
PROAIM
   2 0 0
-2.667587   3.991982   0.000000
-1.345536   1.998661   0.000000
180 90 200
OPTIONS
INTEGER 2
1 1
3 500
REAL 1
1 12.0
```

F.2 Input for Integration of groups in Electric Field

```
hexane3j3o_z+50
   C     2
PROAIM
   2 0 0
-2.667588   3.991992   0.002022
-1.345539   1.998663   0.001876
90 45 100
OPTIONS
INTEGER 1
1 1
REAL 1
1 12.0
```

F.3 Input for External Surface Area Calculation

```
hexane3j3o
   C     2
PROAIM
   2 0 0
-2.667587   3.991982   0.000000
-1.345536   1.998661   0.000000
90 45 100
OPTIONS
INTEGER 2
2 100
3 100
REAL 1
```

```
4 1.0D-9
xsa isodensity surface
0.001d0
montecarlo volume routine
100    !ntrials
100000      !number of random points per trial
```

F.4 Output for Integration

```
PROAIMV - Version 94 - Revision B

 hexane geometry 2t [opt=tight rmp2=full/6-31++g(d,p)]
 -V/T FOR THIS WAVEFUNCTION =      2.00070274000
 MOLECULAR SCF ENERGY (AU)  =     -235.39269906658


hexane3j3o
PROAIM SURFACE ALGORITHM USED
Critical Points in Atomic Surface:
  1 Bond  -2.66758700E+00   3.99198200E+00   0.00000000E+00
  2 Bond  -1.34553600E+00   1.99866100E+00   0.00000000E+00
Optional Parameters Read From Input
 INTEGRATION IS OVER ATOM   C    2
 DYNAMIC CUTOFFS USED THROUGHOUT
 CUTOFF VALUE USED IS 1.00E-09
 PRE-INTEGRATION PRIMITIVE CUTOFF ALGORITHM USED
 TOTAL NUMBER OF PRIMITIVES =    412
 NUMBER OF PRIMITIVES USED OUTSIDE BETA SPHERE=    287
 NUMBER OF PRIMITIVES USED INSIDE BETA SPHERE=    291
 23200 OF THE  75768 PRIM COEFFS ZEROED OUTSIDE BETA SPHERE
 20001 OF THE  76824 PRIM COEFFS ZEROED INSIDE BETA SPHERE
 RADIUS OF BETA SPHERE:  1.4363 WITH    200 POINTS PER RAY
Default number of theta and phi planes used for Beta Sphere
     36 PHI AND    24 THETA PLANES IN BETA SPHERE
    180 PHI AND    90 THETA PLANES OUTSIDE BETA SPHERE
Integrate Out to 1.20E+01 For Rays Intersecting Surface at Infinity
 VOL1 RHO CONTOUR THRESHOLD=   0.0010
 VOL2 RHO CONTOUR THRESHOLD=   0.0020
Doing Beta Sphere Integration ...
Beta Sphere Integration is done ...
Doing Proaim Surface Routine ...
    500 Initial GradRho Trajectories Per Interatomic Surface
    140 Points per GradRho Surface Trajectory
Max. Dist. Between Ends of Adjacent GradRho Surface Trajectories = 6.00E-01
Maximum Length of GradRho Surface Trajectories = 8.00E+00
 INSERTION LIMIT USED        =    6
 INSERTION LIMIT REACHED      0 TIMES FOR SURFACE   1
 INSERTION LIMIT REACHED      0 TIMES FOR SURFACE   2
 FOR SURFACE # 1 NUMBER OF INSERTED PATHS =     0
 FOR SURFACE # 2 NUMBER OF INSERTED PATHS =     0
 TOTAL NUMBER OF INSERTED PATHS=     0
Surface is done ...
Doing Integration Outside of Beta Sphere ...
Default number of radial points used
Integration Outside of Beta Sphere is Done ...


 RESULTS OF THE INTEGRATION
          N  7.98486527866057E+00    NET CHARGE  -1.98486527866057E+00
          G  3.91029327904546E+01
          K  3.91035319975581E+01     E(ATOM)  -3.91310116136340E+01
```

```
          L    5.99207103551932E-04
          I    8.29160733615813E-01
       R(-1)   1.58342580029904E+01
         R1    1.09658789575428E+01
         R2    2.22311657662070E+01
         R4    1.59897138020210E+02
      GR(-1)  -2.98563050921308E+01
        GR0   -2.05314521124194E+01
        GR1   -3.69516503316555E+01
        GR2   -9.58751598161868E+01
       VNEO   -9.50055480179420E+01      VNEO(COR)  -9.50389183920008E+01
       VNET   -1.71506574163116E+02      VNET(COR)  -1.71566815261165E+02
       VEET    6.62015210827790E+01      VEET(COR)   6.62247741408148E+01
        EHF   -6.62015210827790E+01

                                         VREP(COR)   9.32952836099598E+01
                                         V(ATOM)    -7.82715316512053E+01

       EL DX    2.31133904626963E+00
       EL DY    1.56496468222067E+00
       EL DZ   -2.40889541646776E-06
 EL DIPOLE MAG  2.79130841065878E+00

     ATOMIC QUADRUPOLE MOMENT TENSOR
        QXX    2.39288228225037E+00
        QXY   -5.08179667916091E+00
        QXZ    1.90970456971477E-06
        QYY    6.62666798331359E+00
        QYZ    2.93343542977484E-05
        QZZ   -9.01955026556396E+00

     EIGENVALUES OF QUADRUPOLE MOMENT TENSOR:
       -9.95303688499775E-01    1.00148539541321E+01    -9.01955026563236E+00

     EIGENVECTORS OF QUADRUPOLE MOMENT TENSOR:
        8.32025997649714E-01    5.54736639526925E-01    -1.17162945615666E-06
        5.54736639523696E-01   -8.32025997649634E-01    -2.25539081813172E-06
        2.22597409040472E-06   -1.22659800827073E-06     9.99999999996770E-01

       FAXA   -1.11458660206909E+00
       FAYA   -7.89340638668150E-01
       FAZA    1.55945921953161E-07
       FBXA   -6.55837568210484E+00
       FBYA    2.13922593424540E-01
       FBZA    8.40670729638210E-07
      RHO*L    1.03155324375102E+03
       VOL1    1.59604371906269E+02
       VOL2    1.31199459024643E+02
     N(VOL1)   7.93688954123273E+00
     N(VOL2)   7.89609848053934E+00
```

MAXIMUM DISTANCE REACHED FROM NUCLEUS = 1.1996329351E+01

Atomic Overlap Matrix Not Calculated

NORMAL TERMINATION OF PROAIMV

F.5 Output for Surface Area Calculation

```
PROAIMV - Version 94 - Revision B
 hexane geometry 2t [opt=tight rmp2=full/6-31++g(d,p)]
 -V/T FOR THIS WAVEFUNCTION =        2.00070274000
 MOLECULAR SCF ENERGY (AU)   =     -235.39269906658
hexane3j3o
PROAIM SURFACE ALGORITHM USED
Critical Points in Atomic Surface:
   1 Bond  -2.66758700E+00    3.99198200E+00    0.00000000E+00
   2 Bond  -1.34553600E+00    1.99866100E+00    0.00000000E+00
Optional Parameters Read From Input
 INTEGRATION IS OVER ATOM   C     2
 DYNAMIC CUTOFFS USED THROUGHOUT
 CUTOFF VALUE USED IS 1.00E-09
 PRE-INTEGRATION PRIMITIVE CUTOFF ALGORITHM USED
 TOTAL NUMBER OF PRIMITIVES =     412
 NUMBER OF PRIMITIVES USED OUTSIDE BETA SPHERE=    287
 NUMBER OF PRIMITIVES USED INSIDE BETA SPHERE=    291
 23200 OF THE   75768 PRIM COEFFS ZEROED OUTSIDE BETA SPHERE
 20001 OF THE   76824 PRIM COEFFS ZEROED INSIDE BETA SPHERE
 RADIUS OF BETA SPHERE:   1.4363 WITH     100 POINTS PER RAY
Default number of theta and phi planes used for Beta Sphere
    36 PHI AND     24 THETA PLANES IN BETA SPHERE
    90 PHI AND     44 THETA PLANES OUTSIDE BETA SPHERE
Integrate Out to 9.00E+00 For Rays Intersecting Surface at Infinity
 VOL1 RHO CONTOUR THRESHOLD=    0.0010
 VOL2 RHO CONTOUR THRESHOLD=    0.0020
Doing Beta Sphere Integration ...
Beta Sphere Integration is done ...
Doing Proaim Surface Routine ...
    100 Initial GradRho Trajectories Per Interatomic Surface
    100 Points per GradRho Surface Trajectory
Max. Dist. Between Ends of Adjacent GradRho Surface Trajectories = 6.00E-01
Maximum Length of GradRho Surface Trajectories = 8.00E+00
 INSERTION LIMIT USED        =    6
 INSERTION LIMIT REACHED       0 TIMES FOR SURFACE   1
 INSERTION LIMIT REACHED       0 TIMES FOR SURFACE   2
 FOR SURFACE # 1 NUMBER OF INSERTED PATHS =     52
 FOR SURFACE # 2 NUMBER OF INSERTED PATHS =     85
 TOTAL NUMBER OF INSERTED PATHS=    137
Surface is done ...
 proaimxsa process has taken   3501.881 seconds

 atomxsa routine has begun
 the external surface area will be calculated for
      the 0.001000 isodensity surface
   0 pts fixed using the band-aid routine
 writing surface integration (.surfint) file
 .surfint file is called hexane3j3o_C02.surfint
 xsa vaules within 0.001000 by 0.0000000100
 initial shaving uses    1991 wfns of a possible    4140
 writing of surface file (.surface) complete
 .surface file is called hexane3j3o_C02.surface
   23267 wfns used for atomxsa routine
 max distance of surface enclosing domain is  4.664816 au
 min distance of surface enclosing domain is  1.438520 au
 mean distance (r**2*sin(th)) of surface enclosing domain is  4.249719 au
 mean distance    (r*sin(th))  of surface enclosing domain is  4.236977 au
 atomxsa process has taken   1233.962 seconds
 ending atomxsa routine
```

```
external surface area and mc volume routine has begun
100 mc trials will be conducted
 each with    100000 randomly drawn pts between rmin and rmax
 the calculated xsa is   82.4754 au
             for closer inspection, refer to xsasum.txt
 the mc volume is  159.3228 au
             with a std dev of    0.6836 au
             for closer inspection, refer to atmcvolpts.txt
aimstruc process has taken     97.446 seconds
 ending aimstruc routine
```

F.6 Executable file

```
proaimv hexane3j3o_met01 hexane3j3o
proaimxsa hexane3j3o_met01surf hexane3j3o
atomxsa
aimstruc
proaimv hexane3j3o_mee02 hexane3j3o
proaimxsa hexane3j3o_mee02surf hexane3j3o
atomxsa
aimstruc
proaimv hexane3j3o_mee06 hexane3j3o
proaimxsa hexane3j3o_mee06surf hexane3j3o
atomxsa
aimstruc
proaimv hexane3j3o_mee09 hexane3j3o
proaimxsa hexane3j3o_mee09surf hexane3j3o
atomxsa
aimstruc
proaimv hexane3j3o_mee14 hexane3j3o
proaimxsa hexane3j3o_mee14surf hexane3j3o
atomxsa
aimstruc
proaimv hexane3j3o_met15 hexane3j3o
proaimxsa hexane3j3o_met15surf hexane3j3o
atomxsa
aimstruc
proaimv hexane3j3o_x+50_met01 hexane3j3o_x+50
proaimv hexane3j3o_x+50_mee02 hexane3j3o_x+50
proaimv hexane3j3o_x+50_mee06 hexane3j3o_x+50
proaimv hexane3j3o_x+50_mee09 hexane3j3o_x+50
proaimv hexane3j3o_x+50_mee14 hexane3j3o_x+50
proaimv hexane3j3o_x+50_met15 hexane3j3o_x+50
proaimv hexane3j3o_y+50_met01 hexane3j3o_y+50
proaimv hexane3j3o_y+50_mee02 hexane3j3o_y+50
proaimv hexane3j3o_y+50_mee06 hexane3j3o_y+50
proaimv hexane3j3o_y+50_mee09 hexane3j3o_y+50
proaimv hexane3j3o_y+50_mee14 hexane3j3o_y+50
proaimv hexane3j3o_y+50_met15 hexane3j3o_y+50
proaimv hexane3j3o_z+50_met01 hexane3j3o_z+50
proaimv hexane3j3o_z+50_mee02 hexane3j3o_z+50
proaimv hexane3j3o_z+50_mee06 hexane3j3o_z+50
proaimv hexane3j3o_z+50_mee09 hexane3j3o_z+50
proaimv hexane3j3o_z+50_mee14 hexane3j3o_z+50
proaimv hexane3j3o_z+50_met15 hexane3j3o_z+50
```

APPENDIX G

ATOMS IN MOLECULES PROPERTIES FOR FUNCTIONAL GROUPS

All quantities in the tables are given in atomic units. the conversion factors are as follows:

$$E : 1 \text{ au} = 627.51 \frac{\text{kcal}}{\text{mol}} = 2625.5 \frac{\text{kJ}}{\text{mol}} = 27.212 \text{ eV} = 315780.1 \text{ K}$$

$$\mu : 1 \text{ au} = 2.542 \text{ Debye}$$

$$\alpha : 1 \text{ au} = 0.14819 \text{ nm}^3$$

$$V : 1 \text{ au} = 0.089237 \frac{\text{cm}^3}{\text{mol}}$$

$$A : 1 \text{ au} = 0.0028 \text{ nm}^2$$

$$r_{avg} : 1 \text{ au} = 0.529177 \text{ Å}$$

The superscripts on values or functional group definitions stand for the following:

tr denotes value found using assumption of transferability

+ denotes a value determined by the summation of atoms within the group

y denotes value partly determined through symmetry

Table G.1 AIM Properties for Atoms and Functional Groups: Alkanes

	name	Z	q	E	L	μ	α	V	A	r_{avg}
methane 74-82-8	(1) C	6	0.136	-37.8100	6.0E-06	0.000	5.7	78.20	34.48	4.413^{tr}
	(2) H	1	-0.034	-0.6259	-1.7E-06	-0.138	2.8	52.35	44.70	2.749
	(3) H	1	-0.034	-0.6259	-1.9E-06	-0.138	2.6	52.35	44.70	2.749
	(4) H	1	-0.034	-0.6259	-1.8E-06	-0.138	2.6	52.35	44.70	2.749
	(5) H	1	-0.034	-0.6259	-1.7E-06	-0.138	2.6	52.35	44.71	2.749
ethane 74-84-0	(1) CH_3	9	0.000	-39.7259	-8.3E-07	-0.274	14.1	222.29	148.61	4.116
	(2) $CH_3{}^y$	9	0.000	-39.7259	-8.3E-07	-0.274	14.1	222.29	148.61	4.116
propane 74-98-6	(1) CH_3	9	-0.015	-39.7344	2.7E-06	-0.272	14.2	222.11	141.38	4.163
	(2) CH_2	8	0.030	-39.1224	2.6E-04	-0.322	12.0	160.47	91.20	4.187
	(3) $CH_3{}^y$	9	-0.015	-39.7344	2.7E-06	-0.272	14.2	222.11	141.38	4.163
butane (trans) 106-97-8	(1) CH_3	9	-0.013	-39.7285	2.1E-06	-0.274	14.3	222.10	141.48	4.170
	(2) CH_2	8	0.014	-39.1370	5.9E-04	-0.322	11.3	159.59	84.89	4.249
	(3) $CH_2{}^y$	8	0.014	-39.1370	5.9E-04	-0.322	11.3	159.59	84.89	4.249
	(4) $CH_3{}^y$	9	-0.013	-39.7285	2.1E-06	-0.274	14.3	222.10	141.48	4.170

Table G.1 (Continued)

		name	Z	q	E	L	μ	α	V	A	r_{avg}
butane (gauche) 106-97-8-conf		(1) CH_3	9	-0.015	-39.7326	-7.6E-05	-0.276	12.6	220.01	139.28	4.190
		(2) CH_2	8	0.015	-39.1312	1.1E-06	-0.313	9.2	159.85	94.68	4.232
		(3) CH_2	8	0.015	-39.1312	-4.9E-07	-0.313	12.1	159.86	94.64	4.232
		(4) CH_3	9	-0.015	-39.7326	-7.6E-05	-0.276	13.5	220.01	139.28	4.190
isobutane 75-28-5		(1) CH_3	9	-0.024	-39.7444	8.2E-06	-0.276	14.5	220.92	137.34	4.216
		(2) CH	7	0.072	-38.4974	4.0E-04	-0.271	8.4	99.41	41.18	4.323
		(3) CH_3 [y]	9	-0.024	-39.7444	8.2E-06	-0.276	14.5	220.92	137.34	4.216
		(4) CH_3 [y]	9	-0.024	-39.7444	8.2E-06	-0.276	14.5	220.92	137.34	4.216
pentane (trans) 109-66-0		(1) CH_3	9	-0.012	-39.7258	6.3E-07	-0.275	14.4	222.13	141.07	4.164
		(2) CH_2	8	0.014	-39.1331	4.1E-04	-0.320	11.9	159.62	83.05	4.236
		(3) CH_2	8	-0.002	-39.1534	4.3E-04	-0.319	11.9	158.82	75.21	4.294
		(4) CH_2 [y]	8	0.014	-39.1331	4.1E-04	-0.320	11.9	159.62	83.05	4.236
		(5) CH_3 [y]	9	-0.012	-39.7258	6.3E-07	-0.275	14.4	222.13	141.07	4.164
g-pentane (gauche) 109-66-0-conf		(1) CH_3	9	-0.016	-39.7301	-1.0E-03	-0.280	13.0	219.93	153.46	4.235
		(2) CH_2	8	0.011	-39.1261	-5.6E-03	-0.296	11.6	160.55	97.93	4.261
		(3) CH_2	8	-0.006	-39.1462	-5.6E-03	-0.292	12.1	159.71	82.89	4.313
		(4) CH_2	8	0.007	-39.1357	-4.8E-03	-0.335	13.0	158.33	103.03	4.398
		(5) CH_3	9	-0.012	-39.7253	2.1E-06	-0.276	14.8	221.99	144.54	4.183
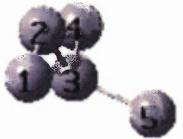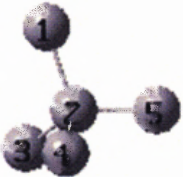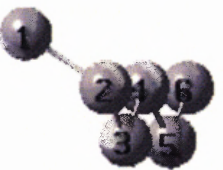
	name	Z	q	E	L	μ	α	V	A	r_{avg}
isopentane 78-78-4	(1) CH_3	9	-0.022	-39.7264	-8.6E-03	-0.290	16.5	221.87	168.35	4.278
	(2) CH_2	8	-0.001	-39.1428	-4.6E-03	-0.290	12.8	159.44	82.29	4.309
	(3) CH	7	0.053	-38.5062	-4.7E-03	-0.271	8.8	99.56	50.62	4.369
	(4) CH_3	9	-0.024	-39.7419	7.2E-06	-0.278	14.7	220.32	132.41	4.215
	(5) CH_3	9	-0.028	-39.7438	-3.7E-03	-0.286	12.9	218.74	150.53	4.275
neopentane 463-82-1	(1) CH_3	9	-0.028	-39.7528	1.3E-05	-0.283	14.1	218.78	123.60	4.225
	(2) C	6	0.115	-37.8582	2.2E-03	0.000	8.1	40.84	0.00	5.014
	(3) $CH_3{}^y$	9	-0.028	-39.7528	1.3E-05	-0.283	14.1	218.78	123.60	4.225
	(4) $CH_3{}^y$	9	-0.028	-39.7528	1.3E-05	-0.283	14.1	218.78	123.60	4.225
	(5) CH_3 y	9	-0.028	-39.7528	1.3E-05	-0.283	14.1	218.78	123.60	4.225
hexane (trans) 110-54-3	(1) CH_3	9	-0.013	-39.7246	1.8E-06	0.275	14.5	222.11	141.57	4.170
	(2) CH_2	8	0.015	-39.1310	6.0E-04	0.321	11.9	159.60	82.48	4.250
	(3) CH_2	8	-0.002	-39.1504	6.5E-04	0.319	11.7	158.63	76.00	4.302
	(4) $CH_2{}^y$	8	-0.002	-39.1504	6.5E-04	0.319	11.7	158.63	76.00	4.302
	(5) $CH_2{}^y$	8	0.015	-39.1310	6.0E-04	0.321	11.9	159.60	82.48	4.250
	(6) $CH_3{}^y$	9	-0.013	-39.7246	1.8E-06	0.275	14.5	222.11	141.57	4.170
hexane (gauche) 110-54-3-conf	(1) CH_3	9	-0.017	-39.7287	-1.8E-03	-0.280	17.2	219.93	303.03	4.479
	(2) CH_2	8	0.014	-39.1246	-3.3E-03	-0.304	12.2	160.13	87.58	4.260
	(3) CH_2	8	-0.005	-39.1433	-5.2E-03	-0.294	11.4	159.52	79.19	4.314
	(4) CH_2	8	-0.006	-39.1537	-2.2E-03	-0.321	13.6	156.76	98.43	4.478
	(5) CH_2	8	0.015	-39.1304	-2.1E-04	-0.319	11.9	159.49	85.55	4.275
	(6) CH_3	9	-0.013	-39.7240	2.9E-06	-0.275	13.2	222.15	143.43	4.180
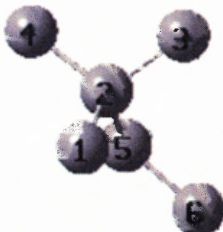
	name	Z	q	E	L	μ	α	V	A	r_{avg}
2-methylpentane 107-83-5	(1) CH$_3$	9	-0.011	-39.7231	5.5E-07	-0.277	16.3	221.85	144.31	4.182
	(2) CH$_2$	8	0.008	-39.1331	-4.4E-03	-0.332	17.7	157.90	77.54	4.285
	(3) CH$_2$	8	-0.019	-39.1601	-6.4E-03	-0.283	12.7	158.89	75.28	4.377
	(4) CH	7	0.051	-38.5029	-5.9E-03	-0.273	13.5	100.25	45.05	4.374
	(5) CH$_3$	9	-0.023	-39.7402	4.2E-06	-0.279	16.9	220.38	137.36	4.218
	(6) CH$_3$	9	-0.023	-39.7440	1.3E-03	-0.281	15.2	217.78	126.46	4.223
2,2-dimethylbutane 75-83-2	(1) CH$_3$	9	-0.037	-39.7521	-6.9E-03	-0.298	13.9	217.50	140.00^{tr}	4.356^{tr}
	(2) C	6	0.099	-37.8618	-2.4E-03	0.041	4.7	41.94	0.00	4.781
	(3) CH$_3$	9	-0.037	-39.7521	-6.9E-03	-0.298	14.5^{tr}	217.49	140.49	4.356
	(4) CH$_3$	9	-0.030	-39.7532	1.6E-05	-0.288	14.5^{tr}	217.77	125.35	4.255
	(5) CH$_2$	8	-0.004	-39.1492	-1.2E-03	-0.295	9.5	156.51	72.58	4.337
	(6) CH$_3$	9	-0.013	-39.7285	-9.7E-04	-0.284	14.5^{tr}	217.22	140.00^{tr}	4.356^{tr}

Table G.2 AIM Properties for Atoms and Functional Groups: Alkenes

		name	Z	q	E	L	μ	α	V	A	r_{avg}
ethene 74-85-1		(1) CH_2	8	0.000	-39.1462	-3.1E-05	-0.198	13.6	199.8	133.5	4.025
		(2) $CH_2{}^y$	8	0.000	-39.1462	-3.1E-05	-0.198	13.6	199.8	133.5	4.025
propene 115-07-1		(1) CH_2	8	-0.031	-39.1473	7.2E-06	-0.225	14.1	201.96	132.23	4.049
		(2) CH	7	-0.011	-38.5577	-3.0E-04	-0.194	11.3	136.31	75.09	4.046
		(3) CH_3	9	0.042	-39.7272	2.3E-06	-0.266	14.2	220.70	144.06	4.131
1-butene 106-98-9		(1) CH_2	8	-0.033	-39.1434	9.2E-06	-0.219	14.3	202.14	132.24	4.047
		(2) CH	7	-0.023	-38.5642	1.1E-04	-0.207	11.3	135.64	67.84	4.084
		(3) CH_2	8	0.052	-39.1269	-9.3E-04	-0.287	11.9	159.74	90.42	4.207
		(4) CH_3	9	0.002	-39.7346	-2.0E-06	-0.271	14.2	220.80	142.81	4.161
trans-2-butene 624-64-6		(1) CH_3	9	0.034	-39.7193	-2.9E-08	-0.276	14.5	221.18	144.35	4.138
		(2) CH	7	-0.034	-38.5669	2.4E-04	-0.215	11.6	137.15	71.54	4.077
		(3) $CH{}^y$	7	-0.034	-38.5669	2.4E-04	-0.215	11.6	137.15	71.54	4.077
		(4) $CH_3{}^y$	9	0.034	-39.7193	-2.9E-08	-0.276	14.5	221.18	144.35	4.138

Table G.2 (Continued)

	name	Z	q	E	L	μ	α	V	A	r_{avg}
2-methylpropene 115-11-7	(1) CH_2	8	-0.051	-39.1524	1.8E-05	-0.253	14.8	202.52	125.99	4.127
	(2) C	6	0.002	-37.9560	1.2E-03	-0.045	9.4	72.66	18.37	4.011
	(3) CH_3	9	0.025	-39.7326	7.7E-06	-0.268	14.0	220.32	137.14	4.154
	(4) $CH_3{}^+$	9	0.025	-39.7326	7.7E-06	-0.268	14.0	220.32	137.14	4.154
1-pentene 109-67-1	(1) CH_2	8	-0.032	-39.1349	3.7E-06	-0.223	14.5	202.09	132.06	4.051
	(2) CH	7	-0.023	-38.5623	9.7E-05	-0.197	11.7	135.15	69.23	4.100
	(3) CH_2	8	0.032	-39.1471	-1.4E-04	-0.296	12.0	158.49	80.41	4.278
	(4) CH_2	8	0.028	-39.1341	-7.0E-05	-0.309	12.0	159.43	96.31	4.258
	(5) CH_3	9	-0.007	-39.7327	1.9E-06	-0.272	14.3	221.63	141.51	4.174
trans-2-pentene 646-04-8	(1) CH_2	9	0.035	-39.7142	2.4E-07	-0.276	14.5	221.19	144.17	4.138
	(2) CH	7	-0.035	-38.5621	-2.1E-05	-0.211	11.9	137.22	70.52	4.104
	(3) CH	7	-0.046	-38.5800	-9.6E-06	-0.215	11.7	135.89	65.82	4.151
	(4) CH_2	8	0.046	-39.1268	-6.3E-06	-0.306	12.4	159.40	86.42	4.217
	(5) CH_3	9	0.000	-39.7281	-2.7E-06	-0.274	14.6	222.06	141.77	4.166
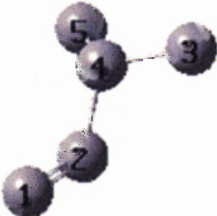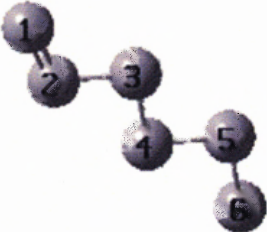
	name	Z	q	E	L	μ	α	V	A	r_{avg}
3-methyl-1-butene 563-45-1	(1) CH_2	8	-0.032	-39.1331	3.2E-06	-0.223	14.2	202.28	131.52	4.102
	(2) CH	7	-0.032	-38.5797	-1.4E-05	-0.196	11.4	133.20	59.85	4.231
	(3) CH_3	9	-0.005	-39.7402	-2.2E-06	-0.273	14.5	220.52	133.70	4.202
	(4) CH	7	0.075	-38.5162	4.0E-04	-0.254	9.8	98.79	38.19	4.349
	(5) CH_3 [y]	9	-0.005	-39.7402	-2.2E-06	-0.273	14.5	220.51	134.42	4.202
1-hexene 592-41-6	(1) CH_2	8	-0.032	-39.1358	8.3E-06	-0.220	14.2	202.33	133.40	4.048
	(2) CH	7	-0.022	-38.5542	-9.0E-05	-0.205	11.3	135.61	68.01	4.084
	(3) CH_2	8	0.036	-39.1371	-1.7E-03	-0.282	11.9	159.01	79.64	4.253
	(4) CH_2	8	0.008	-39.1533	-3.8E-04	-0.314	12.0	158.00	76.12	4.293
	(5) CH_2	8	0.016	-39.1374	-7.9E-04	-0.316	12.3	159.49	83.08	4.247
	(6) CH_3	9	-0.009	-39.7304	2.4E-06	-0.273	14.8	221.78	141.82	4.168

Table G.3 AIM Properties for Atoms and Functional Groups: Amines
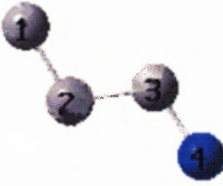
	name	Z	q	E	L	μ	α	V	A	r_{avg}
methanamine 74-89-5	(1) CH_3	9	0.372	-39.5514	-3.8E-06	-0.697	13.2	212.55	148.26	4.113
	(2) NH_2	9	-0.372	-55.9255	6.7E-06	0.595	10.6^t	190.13	131.35	3.896
	N	7	-1.160	-55.0302	6.5E-05	-0.234	7.6	126.08	70.70	3.778
	H	1	0.380	-0.4646	2.5E-04	-0.188	1.5^{tr}	31.83	29.78	2.553
	H ^{y}	1	0.380	-0.4646	2.5E-04	-0.188	1.5^{tr}	31.83	29.78	2.553
ethanamine 75-04-7	(1) CH_3	9	-0.026	-39.7669	2.3E-07	-0.248	14.4	222.95	144.40	4.168
	(2) CH_2	8	0.409	-38.9298	-4.5E-04	-0.601	10.8	151.26	103.46	4.173
	(3) NH_2	9	-0.383	-55.9210	9.4E-06	0.589	12.6^+	189.64	123.51	3.914
	N	7	-1.164	-55.0261	7.1E-05	-0.235	9.6	124.34	67.51	3.778
	H	1	0.391	-0.4475	-1.7E-05	-0.220	1.5	32.27	28.84	2.565
	H ^{y}	1	0.391	-0.4475	-1.7E-05	-0.220	1.5	32.27	28.84	2.565
1-propanamine (trans) 107-10-8	(1) CH_3	9	-0.008	-39.7376	3.1E-06	-0.266	14.8	221.65	141.81	4.166
	(2) CH_2	8	-0.003	-39.1686	-5.0E-04	-0.312	13.6	160.50	84.66	4.247
	(3) CH_2	8	0.393	-38.9411	-3.6E-04	-0.604	10.9	150.45	94.83	4.218
	(4) NH_2	9	-0.383	-55.9107	9.2E-06	0.590	11.9^+	189.60	124.18	3.915
	N	7	-1.164	-55.0156	6.5E-05	-0.235	8.9	124.35	69.12	3.783
	H	1	0.384	-0.4556	-2.2E-05	-0.205	1.5^{tr}	32.34	28.68	2.564
	H ^{y}	1	0.384	-0.4556	8.4E-05	-0.205	1.5^{tr}	32.34	28.68	2.564
1-propanamine (gauche) 107-10-8-conf	(1) CH_3	9	0.008	-39.7361	-5.5E-05	-0.301	13.7	219.31	137.91	4.179
	(2) CH_2	8	0.015	-39.1534	-1.5E-04	-0.315	12.0	159.64	90.97	4.255
	(3) CH_2	8	0.357	-38.9518	-6.4E-05	-0.666	11.5	153.44	87.34	4.239
	(4) NH_2	9	-0.379	-55.9132	-2.0E-05	0.569	12.1^+	186.20	118.04	3.949
	N	7	-1.161	-55.0170	3.0E-05	-0.226	8.9	121.63	60.10	3.807
	H	1	0.389	-0.4487	-2.2E-05	-0.220	1.5	32.48	28.64	2.566
	H	1	0.393	-0.4475	-1.5E-05	-0.218	1.7	31.83	29.90	2.569

Table G.3 (Continued)
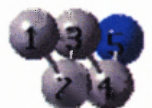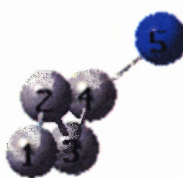
name	Z	q	E	L	μ	α	V	A	r_{avg}
2-propanamine 75-31-0									
(1) CH_3	9	0.000	-39.7570	9.1E-06	-0.279	14.0	219.07	133.32	4.182
(2) CH	7	0.385	-38.3173	-2.5E-03	-0.596	9.6	96.17	53.30^{tr}	4.292^{tr}
(3) CH_3	9	0.000	-39.7570	9.1E-06	-0.279	13.9	219.07	133.31	4.182
(4) NH_2	9	-0.387	-55.9257	2.5E-05	0.587	12.4$^+$	187.83	113.30	3.974
N	7	-1.165	-55.0286	-2.2E-04	-0.235	9.2	122.42	57.42	3.841
H	1	0.389	-0.4486	-1.5E-05	-0.221	1.6	32.52	28.04	2.580
H ^{y}	1	0.389	-0.4486	-1.5E-05	-0.221	1.6	32.52	28.04	2.580
1-butanamine (trans) 109-73-9									
(1) CH_3	9	-0.007	-39.7346	5.2E-07	-0.275	14.3	221.77	146.33	4.167
(2) CH_2	8	0.014	-39.1419	4.6E-04	-0.314	12.6	159.34	82.39	4.246
(3) CH_2	8	-0.018	-39.1828	6.3E-04	-0.312	11.3	159.39	83.80	4.301
(4) CH_2	8	0.393	-38.9363	7.8E-05	-0.602	11.0	150.38	82.23	4.218
(5) NH_2	9	-0.382	-55.9032	6.0E-06	0.592	12.3$^+$	189.64	123.73	3.915
N	7	-1.162	-55.0077	7.9E-05	-0.237	9.3	124.25	64.87	3.780
H	1	0.390	-0.4477	-1.7E-05	-0.220	1.5	32.25	28.67	2.564
H ^{y}	1	0.390	-0.4477	-1.7E-05	-0.220	1.5	32.25	28.67	2.564
1-butanamine (gauche) 109-73-9-conf									
(1) CH_3	9	-0.019	-39.7371	2.1E-06	-0.278	14.3	222.49	143.63	4.187
(2) CH_2	8	0.041	-39.1373	-1.0E-04	-0.298	11.6	156.28	78.04	4.267
(3) CH_2	8	0.000	-39.1671	-4.1E-05	-0.314	12.0	158.77	83.03	4.310
(4) CH_2	8	0.357	-38.9463	1.5E-04	-0.665	11.4	153.51	90.82	4.240
(5) NH_2	9	-0.379	-55.9064	-1.8E-05	0.569	12.3$^+$	186.24	119.26	3.950
N	7	-1.160	-55.0099	-1.3E-04	-0.227	9.1	121.62	60.93	3.806
H	1	0.388	-0.4487	-2.3E-05	-0.220	1.5	32.50	28.54	2.570
H	1	0.393	-0.4478	-1.7E-05	-0.218	1.7	31.84	29.87	2.570

Table G.3 (Continued)

name	Z	q	E	L	μ	α	V	A	r_{avg}
2-butanamine 13952-84-6									
(1) CH_3	9	-0.033	-39.7653	5.0E-06	-0.259	14.4	221.24	134.21	4.212
(2) CH	7	0.402	-38.3189	-1.9E-04	-0.532	9.0	92.94	37.04	4.336
(3) CH_2	8	0.005	-39.1627	-1.9E-04	-0.314	12.0	158.05	84.59	4.301
(4) CH_3	9	0.011	-39.7306	-5.9E-05	-0.304	14.1	219.13	137.23	4.179
(5) NH_2	9	-0.386	-55.9189	-4.5E-05	0.563	12.2^+	184.61	111.98	3.973
N	7	-1.160	-55.0218	6.4E-05	-0.234	9.1	118.69	52.59	3.804
H	1	0.389	-0.4486	-1.4E-05	-0.220	1.6	32.37	28.56	2.580
H	1	0.385	-0.4486	-2.0E-05	-0.224	1.5	33.12	26.87	2.582
2-methyl-1-propanamine 78-81-9									
(1) CH_2	8	0.347	-38.9620	9.6E-05	-0.663	11.2	151.88	79.93	4.287
(2) CH	7	0.054	-38.5300	6.9E-05	-0.271	10.0	98.89	41.27	4.363
(3) CH_3	9	-0.021	-39.7487	2.5E-06	-0.276	14.2	220.22	132.46	4.212
(4) CH_3	9	-0.002	-39.7463	-5.6E-05	-0.307	14.0	217.81	131.31	4.212
(5) NH_2	9	-0.378	-55.9063	-1.9E-05	0.566	12.2^+	185.67	117.88	3.948
N	7	-1.161	-55.0103	-3.1E-05	-0.225	9.1	121.25	56.78	3.807
H	1	0.389	-0.4487	-2.2E-05	-0.220	1.5	32.41	28.68	2.566
H	1	0.394	-0.4473	-1.5E-05	-0.218	1.6	31.82	29.80	2.568
2-methyl-2-propanamine 75-64-9									
(1) CH_3	9	-0.003	-39.7631	1.4E-05	-0.289	14.0	216.70	124.32	4.208
(2) C	6	0.428	-37.6708	-2.4E-03	0.407	7.7	37.54	0^{tr}	4.413^{tr}
(3) CH_3	9	-0.003	-39.7631	1.4E-05	-0.289	14.0	216.70	124.32	4.208
(4) CH_3	9	-0.036	-39.7748	1.3E-05	-0.268	14.3	219.32	126.12	4.245
(5) NH_2	9	-0.389	-55.9276	3.3E-05	0.584	11.9^+	185.61	104.68	3.992
N	7	-1.158	-55.0284	6.9E-05	-0.250	8.9	119.01	50.67	3.837
H	1	0.385	-0.4496	-1.3E-05	-0.223	1.5	33.04	26.62	2.588
H^y	1	0.385	-0.4496	-1.3E-05	-0.223	1.5	33.04	26.62	2.588

	name	Z	q	E	L	μ	α	V	A	r_{avg}
	(1) CH_3	9	-0.010	-39.7333	2.3E-06	-0.273	14.7	221.88	142.66	4.169
	(2) CH_2	8	0.017	-39.1383	-5.2E-04	-0.318	11.8	159.57	82.55	4.248
	(3) CH_2	8	-0.003	-39.1573	-4.1E-04	-0.309	12.3	158.78	74.83	4.298
1-pentanamine (trans)	(4) CH_2	8	-0.019	-39.1775	-1.4E-03	-0.306	11.7	159.76	76.59	4.299
100-58-7	(5) CH_2	8	0.393	-38.9328	-4.4E-04	-0.600	11.1	150.61	80.29	4.217
	(6) NH_2	9	-0.382	-55.8988	8.8E-06	0.592	12.7^+	189.67	123.25	3.914
	N	7	-1.162	-55.0034	7.2E-05	-0.237	9.7	124.30	67.61	3.779
	H	1	0.390	-0.4478	-1.7E-05	-0.220	1.5	32.25	28.65	2.564
	H ^{y}	1	0.390	-0.4478	-1.7E-05	-0.220	1.5	32.25	28.65	2.564

Table G.4 AIM Properties for Atoms and Functional Groups: Diamines
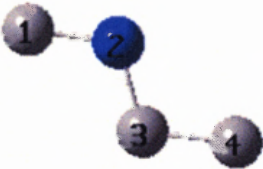
	name	Z	q	E	L	μ	α	V	A	r_{avg}
dimethylamine 124-40-3	(1) CH_3	9	0.361	-39.5526	-3.5E-06	-0.707	12.7	213.05	138.61	4.143
	(2) NH	8	-0.721	-55.5071	1.4E-03	0.525	10.6^+	137.38	75.11	3.909
	(3) CH_3 ^{y}	9	0.361	-39.5526	-3.5E-06	-0.707	12.7	213.05	138.61	4.143
	N	7	-1.104	-55.0483	1.4E-03	-0.289	9.4^{tr}	104.96	46.97	3.826
	H	1	0.383	-0.4587	-2.2E-05	-0.217	1.2	31.84	28.69	2.576
methylethylamine 624-78-2	(1) CH_3	9	0.360	-39.5445	-1.5E-06	-0.706	13.8	213.15	140.39	4.155
	(2) NH	8	-0.730	-55.5102	8.9E-05	0.520	11.0^+	136.43	67.84	3.953
	(3) CH_2	8	0.358	-38.9554	1.5E-04	-0.680	11.6	153.76	82.77	4.250
	(4) CH_3	9	0.012	-39.7434	9.7E-07	-0.277	13.9	220.32	143.23	4.159
	N	7	-1.109	-55.0510	5.5E-05	-0.285	9.6	103.65	41.09	3.866
	H	1	0.379	-0.4592	-1.1E-05	-0.221	1.4	32.57	27.24	2.601
diethylamine 109-89-7	(1) CH_3	9	0.012	-39.7384	-1.6E-06	-0.278	14.3	220.46	141.16	4.158
	(2) CH_2	8	0.357	-38.9500	-4.4E-04	-0.678	11.6	153.84	85.00	4.259
	(3) NH	8	-0.738	-55.5185	1.4E-03	0.515	10.9^+	135.15	59.13	4.021
	(4) CH_2 ^{y}	8	0.357	-38.9500	-3.7E-04	-0.678	11.6	153.84	81.69	4.254
	(5) CH_3 ^{y}	9	0.012	-39.7384	-1.6E-06	-0.278	14.3	220.46	141.16	4.158
	N	7	-1.112	-55.0589	1.5E-03	-0.286	9.2	101.86	34.43	3.915
	H	1	0.374	-0.4596	-1.7E-05	-0.224	1.7	33.23	24.52	2.625
methyl-n-propanamine 627-35-0	(1) CH_3	9	0.360	-39.5401	-3.2E-06	-0.706	13.1	213.09	141.38	4.154
	(2) NH	8	-0.730	-55.5022	-4.1E-05	0.521	11.0^+	136.38	67.52	3.954
	(3) CH_2	8	0.343	-38.9693	-1.6E-05	-0.679	11.0	152.86	73.99	4.303
	(4) CH_2	8	0.033	-39.1483	2.6E-04	-0.306	12.2	158.00	83.82	4.233
	(5) CH_3	9	-0.006	-39.7336	-1.5E-06	-0.272	13.7	221.40	141.44	4.166
	N	7	-1.109	-55.0429	-9.6E-05	-0.283	9.6^{tr}	103.62	43.26	3.864
	H	1	0.364	-0.4719	-3.9E-03	-0.186	1.4^{tr}	32.96	26.90	2.601

Table G.4 (Continued)
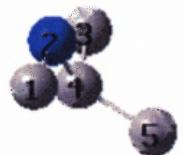
	name	Z	q	E	L	μ	α	V	A	r_{avg}
methylisopropanamine 4747-21-1	(1) CH_3	9	0.350	-39.5406	-9.3E-03	-0.728	13.1^{tr}	212.93	133.50	4.164
	(2) NH	8	-0.741	-55.5077	-3.6E-03	0.496	9.3^+	135.37	67.41	4.024
	(3) CH_3	9	0.002	-39.7543	5.2E-06	-0.286	13.9^{tr}	218.71	133.55	4.201
	(4) CH	7	0.404	-38.3159	-4.7E-03	-0.521	9.7	93.46	37.77	4.346
	(5) CH_3	9	-0.036	-39.7682	-2.5E-03	-0.265	13.9^{tr}	219.09	127.66	4.215
	N	7	-1.117	-55.0496	-3.7E-03	-0.259	8.4	101.96	40.14	3.880
	H	1	0.377	-0.4581	-1.8E-05	-0.224	0.9	33.21	26.46	2.590

Table G.5 AIM Properties for Atoms and Functional Groups: Triamines

	name	Z	q	E	L	μ	α	V	A	r_{avg}
trimethylamine 75-50-3	(1) CH_3	9	0.355	-39.5586	4.8E-06	-0.715	13.4	212.14	135.50	4.212
	(2) N	7	-1.063	-55.0715	1.9E-03	-0.298	11.8	87.58	31.07	3.878
	(3) CH_3^{y}	9	0.355	-39.5586	4.8E-06	-0.715	13.4	212.13	135.50	4.212
	(4) CH_3^{y}	9	0.355	-39.5586	4.3E-06	-0.715	13.4	212.13	135.50	4.212

Table G.6 AIM Properties for Atoms and Functional Groups: Nitriles
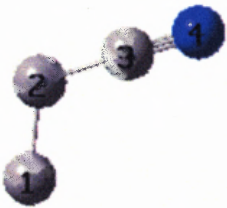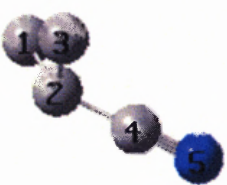
	name	Z	q	E	L	μ	α	V	A	r_{avg}
ethanenitrile 75-05-8	(1) CH_3	9	0.275	-39.6736	-7.1E-06	-0.308	12.9	210.61	146.90	4.020
	(2) C	6	0.922	-37.4735	3.5E-05	0.997	6.6	77.07	36.93	3.645
	(3) N	7	-1.197	-55.2672	-1.5E-05	0.573	9.6	168.32	116.36	3.646
propanenitrile 107-12-0	(1) CH_3	9	0.084	-39.7511	-3.6E-06	-0.256	13.5	216.27	144.31	4.119
	(2) CH_2	8	0.207	-39.0899	-8.9E-04	-0.307	11.5	153.88	90.61	4.117
	(3) C	6	0.906	-37.4761	2.9E-05	1.010	6.6	77.11	34.86	3.679
	(4) N	7	-1.198	-55.2349	-1.2E-05	0.569	9.6	169.16	119.08	3.656
butanenitrile 109-74-0	(1) CH_3	9	0.020	-39.7646	4.8E-06	-0.255	13.8	219.30	142.90	4.151
	(2) CH_2	8	0.081	-39.1516	-1.7E-03	-0.271	11.2	156.74	85.24	4.193
	(3) CH_2	8	0.189	-39.0964	-1.1E-03	-0.307	11.6	153.44	84.41	4.162
	(4) C	6	0.905	-37.4630	5.4E-05	1.009	7.1	77.22	36.57	3.683
	(5) N	7	-1.197	-55.2155	-1.2E-05	0.567	10.1	169.18	118.90	3.656
2-methylpropanenitrile 78-82-0	(1) CH_3	9	0.065	-39.7551	4.6E-06	-0.257	13.7	215.89	134.73	4.143
	(2) CH	7	0.172	-38.4835	-1.2E-03	-0.276	9.7	97.44	41.84	4.190
	(3) CH_3 [y]	9	0.065	-39.7551	4.6E-06	-0.257	13.7	215.89	134.73	4.143
	(4) C	6	0.894	-37.4837	4.7E-05	1.016	6.3	75.43	29.58	3.745
	(5) N	7	-1.198	-55.2131	-1.3E-05	0.565	9.2	169.38	118.67	3.658

Table G.7 AIM Properties for Atoms and Functional Groups: Alcohols
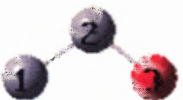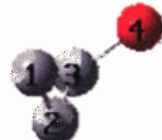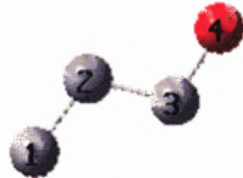
		name	Z	q	E	L	μ	α	V	A	r_{avg}
methanol 67-56-1		(1) CH$_3$	9	0.532	-39.4610	5.5E-07	-0.789	12.1	208.15	148.13	4.078
		(2) OH	9	-0.533	-75.8252	-1.9E-04	0.597	6.1$^+$	150.93	102.51	3.607
		O	8	-1.118	-75.4722	-1.3E-04	0.246	5.1	128.61	86.41	3.517
		H	1	0.586	-0.3530	-1.1E-04	-0.179	1.0	22.18	22.39	2.422
ethanol 64-17-5		(1) CH$_3$	9	0.047	-39.7479	5.1E-06	-0.285	13.6	218.10	145.91	4.138
		(2) CH$_2$	8	0.490	-38.8689	-1.3E-03	-0.757	10.7	151.63	91.73	4.160
		(3) OH	9	-0.538	-75.8106	-2.7E-04	0.646	8.7$^+$	150.56	102.51	3.607
		O	8	-1.122	-75.4563	1.8E-05	0.240	7.6	127.92	80.26	3.547
		H	1	0.584	-0.3543	-1.1E-04	-0.180	1.1	22.36	22.51	2.428
1-propanol (gauche) 71-23-8		(1) CH$_3$	9	0.014	-39.7369	-2.4E-05	-0.296	13.8	218.44	138.73	4.174
		(2) CH$_2$	8	0.048	-39.1521	1.6E-04	-0.311	11.7	157.96	90.10	4.230
		(3) CH$_2$	8	0.476	-38.8773	-1.3E-03	-0.758	10.8	151.35	80.84^{tr}	4.160^{tr}
		(4) OH	9	-0.539	-75.7988	1.0E-04	0.637	8.6$^+$	148.50	97.29	3.618
		O	8	-1.123	-75.4446	-5.8E-05	0.252	7.5	125.92	74.99	3.550
		H	1	0.585	-0.3540	-1.1E-05	-0.179	1.1	22.33	22.72	2.430
1-propanol (trans) 71-23-8-conf		(1) CH$_3$	9	-0.004	-39.7370	-7.4E-07	-0.267	14.2	221.24	142.02	4.162
		(2) CH$_2$	8	0.065	-39.1521	-3.6E-05	-0.278	11.6	156.01	92.33	4.207
		(3) CH$_2$	8	0.477	-38.8799	-7.5E-04	-0.760	10.7	151.01	84.36	4.203
		(4) OH	9	-0.539	-75.7979	-2.0E-04	0.645	8.7$^+$	150.59	102.34	3.608
		O	8	-1.122	-75.4434	-2.3E-05	0.241	7.7	127.92	80.12	3.547
		H	1	0.584	-0.3545	-1.1E-04	-0.180	1.0	22.39	22.49	2.427

Table G.7 (Continued)

		name	Z	q	E	L	μ	α	V	A	r_{avg}
2-propanol 67-63-0		(1) CH$_3$	9	0.037	-39.7546	1.0E-05	-0.290	13.8	216.97	140.26	4.180
		(2) CH	7	0.505	-38.2461	-1.1E-03	-0.658	9.0	93.24	47.45	4.289
		(3) CH$_3$	9	0.000	-39.7667	1.1E-05	-0.259	13.6	219.05	136.03	4.192
		(4) OH	9	-0.543	-75.8041	1.6E-04	0.633	8.6$^+$	149.21	95.59	3.624
		O	8	-1.122	-75.4484	3.7E-06	0.242	7.6	126.04	73.24	3.553
		H	1	0.579	-0.3556	-6.3E-05	-0.183	1.0	23.09	21.53	2.449
1-butanol (gauche) 71-36-3		(1) CH$_3$	9	-0.014	-39.7383	7.6E-07	-0.276	14.2	222.09	144.10	4.184
		(2) CH$_2$	8	0.042	-39.1392	-2.2E-05	-0.299	11.6	156.03	83.57	4.266
		(3) CH$_2$	8	0.032	-39.1658	-2.0E-05	-0.309	11.7	157.09	84.48	4.284
		(4) CH$_2$	8	0.477	-38.8715	-1.4E-03	-0.756	11.3	151.37	80.84^{tr}	4.160^{tr}
		(5) OH	9	-0.538	-75.7896	-2.2E-04	0.636	8.4$^+$	148.53	98.29	3.619
		O	8	-1.122	-75.4354	9.4E-05	0.251	7.3	125.81	75.30	3.552
		H	1	0.584	-0.3543	-5.2E-05	-0.179	1.1	22.34	22.50	2.428
1-butanol (trans) 71-36-3-conf		(1) CH$_3$	9	-0.001	-39.7353	2.4E-06	-0.275	14.1	221.34	142.28	4.165
		(2) CH$_2$	8	0.013	-39.1424	4.4E-04	-0.323	11.6	159.53	83.65	4.244
		(3) CH$_2$	8	0.048	-39.1663	-5.8E-04	-0.273	11.8	155.36	76.61	4.254
		(4) CH$_2$	8	0.477	-38.8747	-8.7E-04	-0.757	10.6	150.94	83.75	4.203
		(5) OH	9	-0.538	-75.7884	-1.9E-04	0.644	9.1$^+$	150.63	102.12	3.609
		O	8	-1.121	-75.4334	-1.1E-05	0.241	8.0	127.80	79.25	3.549
		H	1	0.583	-0.3550	-1.6E-04	-0.180	1.1	22.30	22.19	2.428
2-butanol 78-92-2		(1) CH$_3$	9	-0.012	-39.7407	-3.6E-05	-0.268	13.9	218.93	136.35	4.179
		(2) CH$_2$	8	0.027	-39.1659	-4.0E-05	-0.293	11.8	157.07	83.47	4.276
		(3) CH	7	0.492	-38.2557	-1.8E-04	-0.665	9.0	92.75	73.61	4.340
		(4) CH$_3$	9	0.036	-39.7496	9.8E-06	-0.293	14.0	216.91	134.54	4.180
		(5) OH	9	-0.544	-75.7970	-8.9E-05	0.642	8.2$^+$	147.31	89.82	3.609
		O	8	-1.124	-75.4410	1.1E-04	0.238	7.3	124.72	73.46	3.550
		H	1	0.581	-0.3561	-4.9E-05	-0.179	0.9	22.32	16.93	2.454

Table G.7 (Continued)
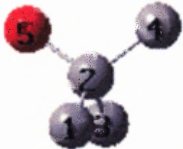
name	Z	q	E	L	μ	α	V	A	r_{avg}
(1) CH_3	9	-0.016	-39.7483	5.1E-06	-0.271	14.1	219.92	132.31	4.208
(2) CH	7	0.083	-38.5312	-1.9E-05	-0.238	9.8	97.35	50.59	4.323
(3) CH_3	9	0.004	-39.7474	-2.2E-05	-0.301	14.0	216.98	131.73	4.208
(4) CH_2	8	0.468	-38.8865	-8.8E-04	-0.758	10.6	149.90	84.36[tr]	4.203[tr]
(5) OH	9	-0.539	-75.7907	5.6E-05	0.636	8.5[+]	148.13	96.95	3.620
O	8	-1.124	-75.4365	-5.5E-05	0.254	7.5	125.48	74.37	3.551
H	1	0.585	-0.3540	-6.0E-05	-0.179	1.0	22.20	22.65	2.429
(1) CH_3	9	-0.003	-39.7721	1.9E-05	-0.272	14.0	217.22	126.12	4.198
(2) C	6	0.519	-37.6115	-6.0E-04	0.522	6.8	36.55	0.00	4.672
(3) CH_3	9	0.031	-39.7615	1.9E-05	-0.298	13.4	215.14	127.74	4.211
(4) CH_3	9	-0.003	-39.7721	1.9E-05	-0.272	13.9	217.22	126.12	4.198
(5) OH	9	-0.544	-75.7993	-2.1E-04	0.622	8.9[+]	147.32	86.18	3.646
O	8	-1.120	-75.4427	-6.7E-07	0.245	7.8	123.42	65.78	3.570
H	1	0.576	-0.3566	-1.6E-04	-0.186	1.1	23.53	20.52	2.465
(1) CH_3	9	-0.012	-39.7350	3.0E-06	-0.277	14.4	222.02	143.22	4.181
(2) CH_2	8	0.012	-39.1418	-8.6E-05	-0.324	12.1	159.70	89.05	4.276
(3) CH_2	8	0.026	-39.1549	-6.9E-05	-0.296	11.6	155.07	72.98	4.325
(4) CH_2	8	0.033	-39.1610	1.9E-04	-0.308	11.8	157.10	99.02	4.284
(5) CH_2	8	0.478	-38.8681	-8.5E-04	-0.757	11.2	151.30	80.84[tr]	4.160[tr]
(6) OH	9	-0.538	-75.7836	-2.1E-04	0.636	8.6[+]	148.55	95.46	3.617
O	8	-1.122	-75.4294	4.3E-05	0.251	7.5	125.80	73.26	3.548
H	1	0.584	-0.3544	-1.0E-04	-0.179	1.1	22.26	22.51	2.429

2-methyl-1-propanol
78-83-1

2-methyl-2-propanol
75-65-0

1-pentanol (gauche)
71-41-0

Table G.7 (Continued)
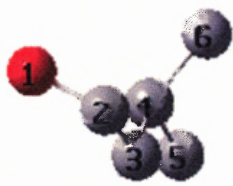
	name	Z	q	E	L	μ	α	V	A	r_{avg}
1-pentanol (trans) 71-41-0-conf	(1) CH$_3$	9	-0.010	-39.7355	2.2E-06	-0.273	14.0	221.91	141.63	4.169
	(2) CH$_2$	8	0.024	-39.1390	6.0E-04	-0.316	12.6	158.95	83.31	4.244
	(3) CH$_2$	8	-0.003	-39.1581	6.6E-04	-0.322	11.0	158.62	76.91	4.296
	(4) CH$_2$	8	0.049	-39.1616	-9.2E-05	-0.273	12.0	155.26	76.24	4.255
	(5) CH$_2$	8	0.479	-38.8714	-1.7E-05	-0.758	10.5	150.87	84.36^{tr}	4.203^{tr}
	(6) OH	9	-0.538	-75.7822	-2.7E-04	0.644	9.4$^+$	150.69	101.95	3.608
	O	8	-1.121	-75.4272	-3.9E-05	0.241	8.4	127.82	78.90	3.547
	H	1	0.583	-0.3551	-1.6E-04	-0.180	1.0	22.31	22.61	2.429
2-methyl-2-butanol 75-85-4	(1) CH$_3$	9	0.015	-39.7323	-2.3E-04	-0.306	13.6	215.86	131.21	4.176
	(2) CH$_2$	8	0.030	-39.1653	-6.4E-04	-0.305	11.8	154.79	78.31	4.309
	(3) C	6	0.505	-37.6179	-2.2E-03	0.525	7.6	36.72	0	4.413
	(4) CH$_3$	9	-0.004	-39.7698	1.4E-05	-0.276	14.2	216.59	127.11	4.235
	(5) CH$_3$	9	-0.004	-39.7719	-6.0E-05	-0.278	13.8	214.62	125.15	4.247
	(6) OH	9	-0.545	-75.7931	-2.2E-05	0.612	8.4$^+$	144.98	79.59	3.627
	O	8	-1.121	-75.4368	-1.7E-05	0.256	7.4	121.13	59.86	3.534
	H	1	0.576	-0.3562	-7.0E-05	-0.186	1.0	23.46	20.66	2.454
3-methyl-1-butanol 123-51-3	(1) OH	9	-0.539	-75.7807	-2.8E-04	0.634	8.3$^+$	148.12	91.74	3.606
	(2) CH$_2$	8	0.477	-38.8731	-5.5E-04	-0.754	15.6	148.38	80.84	4.160
	(3) CH$_2$	8	0.019	-39.1722	-1.6E-03	-0.297	13.0	156.34	77.82	4.312
	(4) CH	7	0.089	-38.5078	-1.6E-03	-0.209	7.0	95.30	32.41	4.321
	(5) CH$_3$	9	-0.025	-39.7507	6.6E-06	-0.281	13.6	220.49	131.37	4.209
	(6) CH$_3$	9	-0.028	-39.7539	-2.3E-03	-0.291	13.4	218.47	124.88	4.224
	O	8	-1.123	-75.4264	-5.1E-05	0.253	7.5^{tr}	125.45	71.19	3.532
	H	1	0.584	-0.3544	-1.7E-04	-0.179	0.8	22.28	22.35	2.425

Table G.7 (Continued)
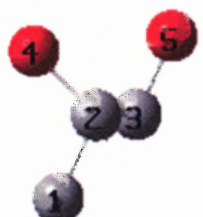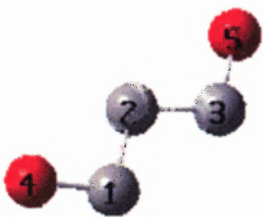
	name	Z	q	E	L	μ	α	V	A	r_{avg}
1,2-propanediol 57-55-6	(1) CH$_3$	9	0.019	-39.7608	8.6E-06	-0.250	14.0	218.41	135.70	4.178
	(2) CH	7	0.493	-38.2890	-1.6E-03	-0.636	7.9	92.04	53.30	4.292
	(3) CH$_2$	8	0.554	-38.8764	-4.3E-04	-0.724	9.1	146.13	80.84^{tr}	4.160^{tr}
	(4) OH	9	-0.541	-75.8136	5.7E-05	0.648	6.9$^+$	145.11	90.00	3.632
	(5) OH	9	-0.528	-75.8104	1.3E-04	0.712	5.0$^+$	146.07	95.27	3.566
	(4) O	8	-1.130	-75.4629	8.6E-05	0.250	6.2	122.57	67.22	3.555
	(4) H	1	0.589	-0.3507	-2.3E-05	-0.177	0.7^{tr}	22.23	22.13	2.428
	(5) O	8	-1.131	-75.4673	-1.5E-05	0.227	4.3	125.84	79.70	3.518
	(5) H	1	0.604	-0.3429	-1.1E-04	-0.166	0.7	19.95	15.23	2.428
1,3-propanediol 504-63-2	(1) CH$_2$	8	0.488	-38.8759	-3.8E-04	-0.759	10.1	150.48	85.72	4.190
	(2) CH$_2$	8	0.059	-39.1903	5.0E-04	-0.245	10.6	153.77	78.96	4.229
	(3) CH$_2$	8	0.519	-38.8644	-4.1E-04	0.726	9.6	148.56	83.31	4.179
	(4) OH	9	-0.533	-75.8082	-2.1E-04	0.650	7.2$^+$	150.00	101.53	3.603
	(5) OH	9	-0.534	-75.8068	4.0E-05	0.641	5.0$^+$	149.86	101.91	3.596
	(4) O	8	-1.118	-75.4545	8.8E-06	0.241	6.5	127.47	78.07	3.543
	(4) H	1	0.585	-0.3538	-1.6E-04	-0.178	0.7^{tr}	22.16	22.54	2.425
	(5) O	8	-1.118	-75.4531	1.1E-05	0.245	4.4	126.89	79.26	3.514
	(5) H	1	0.583	-0.3537	7.4E-05	-0.181	0.6	22.68	21.47	2.433
1-methoxy-2-propanol 107-98-2	(1) CH$_3$	9	0.536	-39.4519	6.2E-06	-0.789	12.7	207.74	141.67	4.122
	(2) O	8	-1.067	-75.4973	3.4E-04	0.029	8.0	101.38	46.22	3.564
	(3) CH$_2$	8	0.483	-38.9146	1.1E-04	-0.756	10.7	148.72	76.00	4.310
	(4) CH	7	0.535	-38.2634	-1.1E-03	-0.654	8.5	91.05	46.57	4.302
	(5) CH$_3$	9	0.042	-39.7505	6.4E-06	-0.287	13.9	217.00	140.74	4.176
	(6) OH	9	-0.530	-75.8117	-9.0E-05	0.709	8.3$^+$	145.10	88.75	3.591
	O	8	-1.132	-75.4680	-1.3E-05	0.227	7.5	125.03	73.82	3.544
	H	1	0.602	-0.3437	-1.3E-04	-0.166	0.8	19.94	15.12	2.435

Table G.8 AIM Properties for Atoms and Functional Groups: Ethers

		name	Z	q	E	L	μ	α	V	A	r_{avg}
dimethyl ether 115-10-6		(1) CH$_3$	9	0.529	-39.4608	1.1E-05	-0.809	12.6	208.19	139.54	4.110
		(2) O	8	-1.057	-75.5051	2.7E-04	0.037	8.0	106.36	58.67	3.502
		(3) CH$_3$ ^{y}	9	0.529	-39.4608	1.1E-05	-0.809	12.6	208.19	139.54	4.110
methyl ethyl ether 540-67-0		(1) CH$_3$	9	0.526	-39.4538	1.0E-05	-0.810	12.9	208.55	140.33	4.119
		(2) O	8	-1.059	-75.4967	7.0E-04	0.026	7.4	105.50	52.36	3.556
		(3) CH$_2$	8	0.488	-38.8755	1.3E-04	-0.779	10.7	150.73	82.98	4.222
		(4) CH$_3$	9	0.047	-39.7435	6.1E-06	-0.285	13.8	218.43	144.63	4.140
diethyl ether 60-29-7		(1) CH$_3$	9	0.045	-39.7388	4.6E-06	-0.286	14.1	218.676	143.170	4.133
		(2) CH$_2$	8	0.486	-38.8708	-3.1E-05	-0.778	10.8	150.849	82.682	4.214
		(3) O	8	-1.063	-75.4936	3.0E-04	0.017	8.2	104.498	46.216	3.586
		(4) CH$_2$ ^{y}	8	0.486	-38.8708	-3.1E-05	-0.778	10.8	150.849	82.682	4.213
		(5) CH$_3$ ^{y}	9	0.045	-39.7388	4.6E-06	-0.286	14.1	218.676	142.527	4.133
methyl propyl ether 557-17-5		(1) CH$_3$	9	0.526	-39.4478	6.9E-06	-0.810	13.0	208.45	140.57	4.129
		(2) O	8	-1.062	-75.4889	-1.9E-05	0.040	7.9	103.39	46.47	3.556
		(3) CH$_2$	8	0.474	-38.8848	-7.2E-05	-0.780	11.1	150.52	82.31	4.291
		(4) CH$_2$	8	0.048	-39.1495	-2.8E-05	-0.311	11.9	158.10	92.74	4.231
		(5) CH$_3$	9	0.014	-39.7359	-2.1E-05	-0.296	13.8	218.63	138.39	4.173
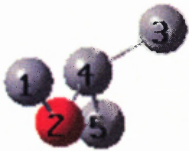
	name	Z	q	E	L	μ	α	V	A	r_{avg}
methyl isopropyl ether 598-53-8	(1) CH_3	9	0.513	-39.4524	-8.6E-03	-0.826	15.2	208.35	134.00	4.133
	(2) O	8	-1.062	-75.4789	-4.0E-03	0.043	6.7	104.78	48.08	3.569
	(3) CH_3	9	-0.007	-39.7650	-8.1E-03	-0.278	14.0	218.96	129.28	4.195
	(4) CH	7	0.494	-38.2522	-6.9E-03	-0.658	9.3	93.16	38.55	4.359
	(5) CH_3	9	0.034	-39.7543	1.2E-05	-0.293	13.8	217.23	135.03	4.180

Table G.9 AIM Properties for Atoms and Functional Groups: Aldehydes
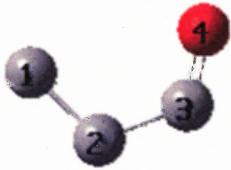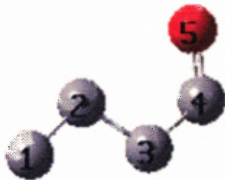
	name	Z	q	E	L	μ	α	V	A	r_{avg}
methanal 50-00-0	(1) CH_2	8	1.079	-38.4991	2.0E-05	-1.030	10.1	164.30	122.08	3.909
	(2) O	8	-1.079	-75.6569	-4.0E-06	0.507	7.3	139.09	102.50	3.436
ethanal 75-07-0	(1) CH_3	9	0.071	-39.7704	6.3E-07	-0.194	13.8	218.14	147.21	4.096
	(2) CH	7	1.036	-37.8957	-1.3E-03	-0.956	8.4	107.45	68.19	3.893
	(3) O	8	-1.107	-75.6310	1.5E-05	0.486	7.8	139.38	98.59	3.468
propanal 123-38-6	(1) CH_3	9	0.057	-39.7355	-1.3E-05	-0.283	13.5	215.88	138.13	4.145
	(2) CH_2	8	0.032	-39.1849	7.6E-05	-0.223	12.2	160.00	91.78	4.171
	(3) CH	7	1.021	-37.9046	5.6E-04	-0.967	8.2	107.36	67.60	3.903
	(4) O	8	-1.111	-75.6107	-2.3E-06	0.483	7.4	137.23	92.08	3.469
butanal 123-72-8	(1) CH_3	9	-0.003	-39.7488	1.8E-06	-0.268	14.1	221.22	142.47	4.163
	(2) CH_2	8	0.076	-39.1375	4.6E-04	-0.257	11.2	153.99	79.17	4.218
	(3) CH_2	8	0.017	-39.1961	7.2E-05	-0.221	12.2	159.29	84.82	4.214
	(4) CH	7	1.024	-37.8874	-1.8E-04	-0.962	8.6	107.49	66.60	3.907
	(5) O	8	-1.110	-75.5964	-5.4E-06	0.483	7.6	137.09	89.95	3.467

Table G.9 (Continued)
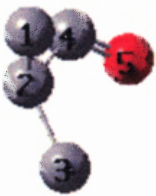
	name	Z	q	E	L	μ	α	V	A	r_{avg}
2-methylpropanal 78-84-2	(1) CH_3	9	0.021	-39.7434	2.6E-06	-0.254	13.8	218.11	134.94	4.186
	(2) CH	7	0.037	-38.5757	-4.1E-04	-0.198	10.2	100.55	68.22	4.286
	(3) CH_3	9	0.042	-39.7446	-8.8E-06	-0.286	13.4	215.13	134.79	4.190
	(4) CH	7	1.008	-37.9164	-2.4E-05	-0.972	8.3	106.09	64.14	3.988
	(5) O	8	-1.109	-75.5949	-8.4E-06	0.483	7.6	137.07	89.98	3.468

Table G.10 AIM Properties for Atoms and Functional Groups: Ketones
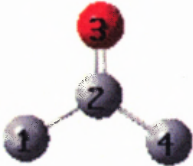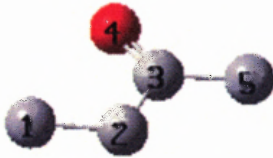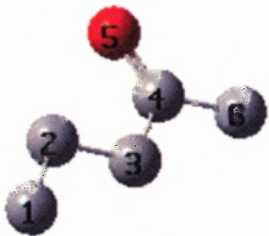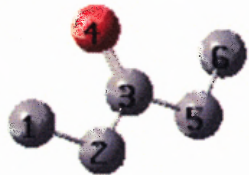
name	Z	q	E	L	μ	α	V	A	r_{avg}
propanone 67-64-1									
(1) CH_3	9	0.053	-39.7678	5.1E-06	-0.212	13.7	217.51	140.37	4.120
(2) C	6	1.020	-37.2939	-2.8E-04	0.827	6.0	48.25	27.44	3.642
(3) O	8	-1.127	-75.6159	1.7E-05	0.459	8.2	138.92	93.45	3.468
(4) $CH_3^{\,y}$	9	0.053	-39.7678	5.1E-06	-0.212	13.7	217.51	140.37	4.120
butanone 78-93-3									
(1) CH_3	9	0.053	-39.7306	-1.1E-05	-0.294	13.9	215.86	138.79	4.150
(2) CH_2	8	0.018	-39.1836	1.2E-04	-0.238	12.0	159.24	85.17	4.201
(3) C	6	1.006	-37.3035	-6.8E-04	0.838	6.1	48.12	10.05	3.613
(4) O	8	-1.129	-75.6039	-5.7E-06	0.458	7.7	136.28	83.86	3.481
(5) CH_3	9	0.051	-39.7609	6.7E-06	-0.214	13.7	217.56	139.79	4.122
2-pentanone 107-87-9									
(1) CH_3	9	-0.009	-39.7452	1.6E-06	-0.270	14.8	221.57	142.60	4.166
(2) CH_2	8	0.077	-39.1330	5.5E-04	-0.258	11.3	153.60	79.55	4.223
(3) CH_2	8	0.003	-39.1974	4.0E-05	-0.235	12.2	158.22	76.66	4.251
(4) C	6	1.007	-37.2970	-6.3E-05	0.833	6.3	47.98	9.97	3.617
(5) O	8	-1.128	-75.5937	-4.1E-07	0.457	7.4	136.15	83.97	3.483
(6) CH_3	9	0.051	-39.7559	6.1E-06	-0.214	13.9	217.64	138.92	4.123
3-pentanone 96-22-0									
(1) CH_3	9	0.052	-39.7258	-1.1E-05	-0.295	13.8	216.02	137.38	4.147
(2) CH_2	8	0.017	-39.1788	-1.7E-05	-0.238	11.8	159.11	84.25	4.198
(3) C	6	0.992	-37.3155	-2.1E-04	0.842	6.2	47.66	13.21	3.643
(4) O	8	-1.130	-75.5957	-2.5E-05	0.455	7.6	133.46	75.16	3.459
(5) $CH_2^{\,y}$	8	0.017	-39.1788	-1.7E-05	-0.238	11.8	159.11	84.25	4.198
(6) $CH_3^{\,y}$	9	0.052	-39.7258	-1.1E-05	-0.295	13.8	216.02	137.38	4.147

Table G.10 (Continued)
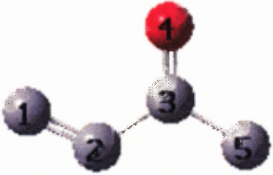
	name	Z	q	E	L	μ	α	V	A	r_{avg}
3-methyl-2-butanone 563-80-4	(1) CH$_3$	9	0.016	-39.7435	-3.7E-05	-0.263	13.7	216.47	127.45	4.195
	(2) CH	7	0.030	-38.5698	-6.6E-04	-0.213	10.3	99.62	57.68	4.313
	(3) CH$_3$	9	0.038	-39.7424	-7.6E-06	-0.297	13.6	214.74	128.17	4.192
	(4) C	6	0.994	-37.3132	-2.2E-04	0.846	6.1	47.07	15.72	3.658
	(5) O	8	-1.128	-75.5910	-4.3E-06	0.454	7.5	135.89	82.93	3.482
	(6) CH$_3$	9	0.049	-39.7579	-2.3E-05	-0.224	13.7	215.66	131.74	4.139
methyl vinyl ketone 78-94-3	(1) CH$_2$	8	0.083	-39.1325	-9.3E-06	-0.200	13.1	190.89	125.65	3.984
	(2) CH	7	-0.039	-38.6229	-1.7E-04	-0.183	12.3	138.19	72.48	4.050
	(3) C	6	1.011	-37.3048	1.4E-04	0.810	6.3	49.57	11.54	3.651
	(4) O	8	-1.115	-75.6037	1.0E-05	0.439	7.7	136.02	85.70	3.478
	(5) CH$_3$	9	0.060	-39.7639	5.1E-06	-0.213	14.2	217.18	139.90	4.118

Table G.11 AIM Properties for Atoms and Functional Groups: Carboxylic Acids
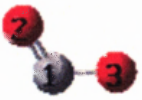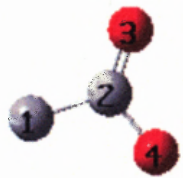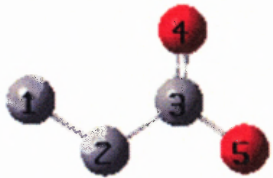
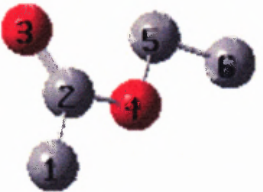|	name	Z	q	E	L	μ	α	V	A	r_{avg}
methanoic acid 64-18-6	(1) CH	7	1.702	-37.4819	8.2E-04	-0.794	6.1	90.98	62.65	3.796
	(2) O	8	-1.170	-75.7079	9.5E-06	0.538	8.0	139.71	100.02	3.470
	(3) OH	9	-0.531	-75.9902	-1.1E-04	0.832	8.6^+	145.68	104.60	3.529
	O	8	-1.153	-75.6634	-4.3E-05	0.283	7.8^{tr}	125.01	84.46	3.470
	H	1	0.622	-0.3268	-1.5E-04	-0.161	0.8	20.81	19.71	2.373
ethanoic acid 64-19-7	(1) CH_3	9	0.150	-39.7353	7.4E-06	-0.252	13.1	212.28	141.35	4.084
	(2) C	6	1.571	-36.9349	-7.3E-04	0.817	4.4	38.24	37.82	3.532
	(3) O	8	-1.186	-75.6898	2.5E-05	0.510	7.6	139.25	95.20	3.481
	(4) OH	9	-0.535	-75.9638	-7.6E-05	0.805	8.9^+	144.53	98.26	3.547
	O	8	-1.155	-75.6343	-4.1E-05	0.265	8.0	123.65	78.06	3.486
	H	1	0.619	-0.3294	-1.5E-04	-0.162	0.9	20.80	19.69	2.383
propanoic acid 79-09-4	(1) CH_3	9	0.064	-39.7365	-1.1E-05	-0.283	13.4	215.267	138.362	4.142
	(2) CH_2	8	0.103	-39.1544	6.0E-05	-0.236	11.2	154.894	86.975	4.159
	(3) C	6	1.556	-36.9474	-9.7E-04	0.834	4.8	37.886	55.880	3.517
	(4) O	8	-1.187	-75.6744	1.2E-05	0.507	7.9	137.145	86.553	3.481
	(5) OH	9	-0.537	-75.9477	-9.9E-05	0.800	8.6^+	144.384	97.377	3.549
	O	8	-1.155	-75.6177	-2.1E-05	0.266	7.6	123.472	79.175	3.488
	H	1	0.618	-0.3300	-1.5E-04	-0.163	1.0	20.858	19.567	2.385

Table G.11 (Continued)
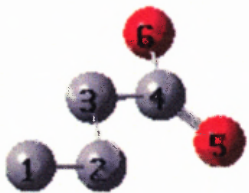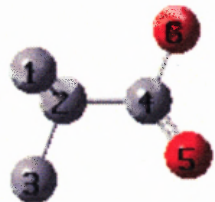
	name	Z	q	E	L	μ	α	V	A	r_{avg}
butanoic acid 107-92-6	(1) CH₃	9	-0.001	-39.7514	3.3E-06	-0.265	14.1	220.83	142.54	4.161
	(2) CH₂	8	0.081	-39.1400	6.8E-04	-0.264	11.0	153.73	79.63	4.217
	(3) CH₂	8	0.088	-39.1679	7.2E-06	-0.233	11.4	153.96	78.50	4.206
	(4) C	6	1.555	-36.9410	-1.6E-03	0.832	5.3	37.90	23.34	3.514
	(5) O	8	-1.187	-75.6633	6.4E-06	0.506	7.6	137.03	86.18	3.482
	(6) OH	9	-0.537	-75.9363	-1.0E-04	0.799	8.9⁺	144.49	97.06	3.548
	O	8	-1.154	-75.6059	-7.9E-05	0.265	7.9	123.55	77.06	3.487
	H	1	0.618	-0.3303	-1.5E-04	-0.163	1.0	20.90	19.63	2.386
2-methylpropanic acid 79-31-2	(1) CH₃	9	0.045	-39.7410	-1.1E-05	-0.269	13.7	215.91	131.95	4.172
	(2) CH	7	0.088	-38.5521	-1.2E-03	-0.219	9.8	96.82	58.36	4.257
	(3) CH₃	9	0.047	-39.7475	-7.1E-06	-0.284	13.4	214.75	128.70	4.179
	(4) C	6	1.541	-36.9634	-1.0E-03	0.845	4.6	36.43	34.01	3.553
	(5) O	8	-1.187	-75.6590	1.3E-06	0.501	7.6	137.34	85.97	3.476
	(6) OH	9	-0.536	-75.9345	5.6E-05	0.793	8.3⁺	143.92	97.80	3.558
	O	8	-1.153	-75.6041	3.9E-05	0.267	7.4	122.94	78.42	3.489
	H	1	0.617	-0.3304	-7.5E-05	-0.164	0.9	20.94	20.01	2.391

Table G.12 AIM Properties for Atoms and Functional Groups: Esters
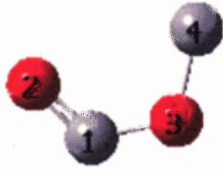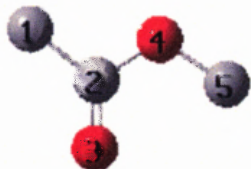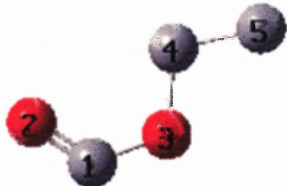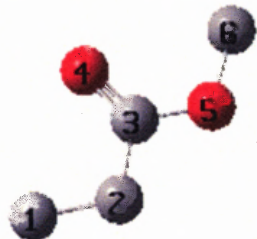
	name	Z	q	E	L	μ	α	V	A	r_{avg}
methyl methanoate 107-31-3	(1) CH	7	1.678	-37.4948	-6.9E-04	-0.807	6.6	91.93	63.11	3.815
	(2) O	8	-1.171	-75.6879	-3.4E-06	0.518	7.4	137.05	90.80	3.476
	(3) O	8	-1.083	-75.6480	3.8E-05	0.213	8.4	105.38	58.66	3.502
	(4) CH_3	9	0.576	-39.4820	5.0E-06	-0.696	11.7	199.98	134.70	4.062
methyl ethanoate 79-20-9	(1) CH_3	9	0.136	-39.7356	8.1E-06	-0.258	13.7	213.22	140.70	4.092
	(2) C	6	1.559	-36.9448	-3.1E-03	0.809	5.0	38.57	41.85	3.542
	(3) O	8	-1.184	-75.6768	6.4E-06	0.493	7.4	136.18	84.57	3.486
	(4) O	8	-1.085	-75.6271	8.8E-05	0.179	8.3	103.85	52.38	3.521
	(5) CH_3	9	0.571	-39.4743	6.8E-07	-0.713	11.7	200.59	134.33	4.071
ethyl methanoate 109-94-4	(1) CH	7	1.674	-37.4888	-1.4E-04	-0.813	6.9	92.12	62.53	3.821
	(2) O	8	-1.172	-75.6707	-6.7E-06	0.515	7.7	137.17	90.40	3.479
	(3) O	8	-1.086	-75.6380	6.1E-05	0.211	8.4	105.09	52.59	3.535
	(4) CH_2	8	0.513	-38.9060	1.6E-04	-0.652	9.9	144.35	78.41	4.151
	(5) CH_3	9	0.070	-39.7527	5.5E-06	-0.270	13.8	216.63	143.89	4.126
methyl propanoate 554-12-1	(1) CH_3	9	0.058	-39.7360	-1.2E-05	-0.287	13.3	215.83	139.08	4.146
	(2) CH_2	8	0.096	-39.1545	8.1E-05	-0.240	11.7	155.32	86.44	4.168
	(3) C	6	1.547	-36.9587	-3.1E-04	0.822	4.9	37.96	8.48	3.523
	(4) O	8	-1.184	-75.6645	-1.0E-05	0.489	7.0	133.90	76.28	3.487
	(5) O	8	-1.086	-75.6140	-1.3E-04	0.179	8.9	103.62	51.61	3.521
	(6) CH_3	9	0.570	-39.4683	5.0E-06	-0.715	11.9	200.76	134.79	4.071

Table G.12 (Continued)
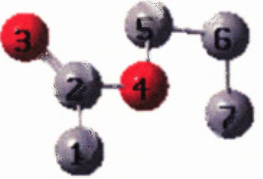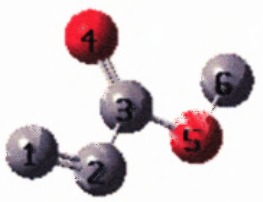
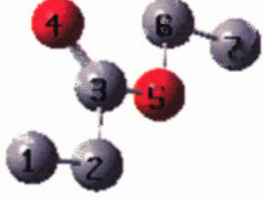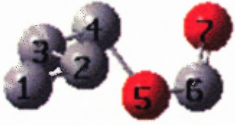|	name	Z	q	E	L	μ	α	V	A	r_{avg}
ethyl ethanoate 141-78-6	(1) CH$_3$	9	0.133	-39.7313	8.0E-06	-0.263	13.5	213.55	141.37	4.093
	(2) C	6	1.559	-36.9366	-2.7E-03	0.806	4.9	38.63	40.82	3.544
	(3) O	8	-1.190	-75.6655	1.1E-05	0.488	7.7	136.52	84.84	3.490
	(4) O	8	-1.083	-75.6195	-7.9E-05	0.181	8.4	103.70	45.87	3.559
	(5) CH$_2$	8	0.516	-38.8985	1.9E-04	-0.661	9.8	144.15	78.32	4.160
	(6) CH$_3$	9	0.062	-39.7484	5.1E-06	-0.273	13.7	217.19	144.14	4.131
propyl methanoate 110-74-7	(1) CH	7	1.673	-37.4829	7.4E-05	-0.815	6.8	92.25	61.51	3.853
	(2) O	8	-1.172	-75.6578	-9.5E-06	0.515	7.7	137.15	90.31	3.476
	(3) O	8	-1.087	-75.6275	-1.4E-04	0.220	8.5	103.45	50.22	3.549
	(4) CH$_2$	8	0.499	-38.9157	-4.8E-05	-0.654	9.9	143.93	74.69	4.210
	(5) CH$_2$	8	0.065	-39.1593	-1.1E-04	-0.293	11.7	156.94	89.02	4.217
	(6) CH$_3$	9	0.022	-39.7507	-2.1E-05	-0.277	13.8	217.93	137.80	4.168
propyl ethanoate 109-60-4	(1) CH$_3$	9	0.132	-39.7266	-6.8E-07	-0.259	13.6	213.61	141.22	4.098
	(2) C ^{tr}	6	1.559	-36.9366	-2.7E-03	0.806	4.9	38.63	9.82^{tr}	3.559^{tr}
	(3) O	8	-1.184	-75.6539	-2.1E-06	0.490	7.6	136.22	84.67	3.487
	(4) O	8	-1.089	-75.6144	-1.3E-04	0.188	8.5	101.77	38.85	3.578
	(5) CH$_2$	8	0.502	-38.9078	9.8E-06	-0.666	10.1	143.89	73.63	4.210
	(6) CH$_2$	8	0.060	-39.1560	-8.2E-05	-0.297	12.1	157.27	88.59	4.210
	(7) CH$_3$	9	0.017	-39.7468	-1.2E-04	-0.281	13.8	218.37	136.95	4.165
methyl acrylate 96-33-3	(1) CH$_2$	8	0.089	-39.1333	-7.8E-06	-0.198	13.5	190.99	124.34	3.984
	(2) CH	7	0.036	-38.5921	1.9E-04	-0.173	11.8	133.67	73.37	4.014
	(3) C	6	1.554	-36.9566	-8.7E-04	0.783	5.4	39.42	9.82	3.559
	(4) O	8	-1.176	-75.6635	-3.4E-06	0.476	7.3	134.02	77.77	3.484
	(5) O	8	-1.078	-75.6208	1.0E-04	0.180	8.7	103.41	52.08	3.518
	(6) CH$_3$	9	0.575	-39.4715	3.1E-06	-0.710	11.9	200.32	134.84	4.069

Table G.12 (Continued)
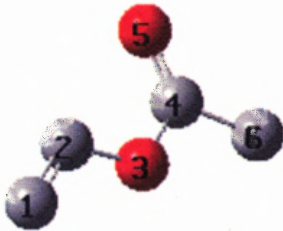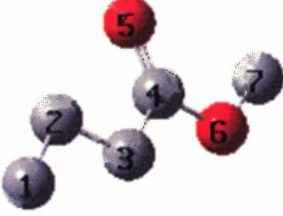
	name	Z	q	E	L	μ	α	V	A	r_{avg}
vinyl ethanoate 108-05-4	(1) CH$_2$	8	0.063	-39.1311	-1.5E-06	-0.272	14.2	198.50	132.90	4.020
	(2) CH	7	0.525	-38.2767	-1.0E-03	-0.745	10.3	118.35	68.99	3.959
	(3) O	8	-1.111	-75.6667	1.8E-05	0.169	9.2	103.89	48.44	3.528
	(4) C	6	1.543	-36.9522	-1.0E-03	0.810	5.4	38.75	9.53	3.522
	(5) O	8	-1.175	-75.6892	-1.6E-05	0.516	7.7	134.51	85.88	3.471
	(6) CH$_3$	9	0.154	-39.7297	8.1E-06	-0.256	13.7	212.04	140.44	4.085
methyl butanoate 623-42-7	(1) CH$_3$	9	-0.005	-39.7508	3.2E-06	-0.268	13.7^{tr}	221.23	142.05	4.163
	(2) CH$_2$	8	0.076	-39.1394	-3.1E-04	-0.263	11.2^{tr}	154.02	79.18	4.220
	(3) CH$_2$	8	0.081	-39.1689	-1.2E-04	-0.238	11.0^{tr}	154.24	78.40	4.216
	(4) C	6	1.542	-36.9475	-4.5E-03	0.790	5.3^{tr}	38.77	8.48	3.525
	(5) O	8	-1.195	-75.6597	-8.3E-03	0.480	7.3^{tr}	134.03	77.22	3.485
	(6) O	8	-1.089	-75.5942	-6.0E-03	0.192	7.7	104.44	52.10	3.521
	(7) CH$_3$	9	0.569	-39.4675	6.9E-04	-0.729	8.4	200.15	135.13	4.052
ethyl propanoate 105-37-3	(1) CH$_3$	9	0.057	-39.7318	-1.3E-05	-0.288	13.7	215.90	138.28	4.147
	(2) CH$_2$	8	0.093	-39.1514	3.8E-05	-0.241	11.6	155.49	85.76	4.170
	(3) C	6	1.545	-36.9530	-1.7E-03	0.823	5.3	38.16	16.60	3.532
	(4) O	8	-1.185	-75.6546	-1.3E-05	0.487	7.5	133.93	76.14	3.488
	(5) O	8	-1.088	-75.6115	6.5E-05	0.177	8.7	103.51	44.50	3.558
	(6) CH$_2$	8	0.516	-38.8921	1.9E-04	-0.666	10.0	144.41	77.89	4.162
	(7) CH$_3$	9	0.061	-39.7443	3.9E-06	-0.275	13.9	217.34	144.48	4.131
butyl methanoate 592-84-7	(1) CH$_3$	9	-0.003	-39.7510	3.2E-06	-0.270	14.2	221.23	142.62	4.174
	(2) CH$_2$	8	0.040	-39.1553	-6.0E-05	-0.297	11.5	156.13	83.93	4.256
	(3) CH$_2$	8	0.049	-39.1732	2.8E-05	-0.292	11.8	156.07	84.09	4.268
	(4) CH$_2$	8	0.499	-38.9092	-4.9E-04	-0.653	9.9	144.10	86.65	4.210
	(5) O	8	-1.087	-75.6178	4.9E-05	0.220	8.4	103.37	48.96	3.551
	(6) CH	7	1.671	-37.4793	-1.5E-03	-0.811	6.4	92.28	61.16	3.852
	(7) O	8	-1.172	-75.6483	-5.8E-06	0.514	7.8	137.15	90.48	3.476

Table G.13 AIM Properties for Atoms and Functional Groups: Fluorides

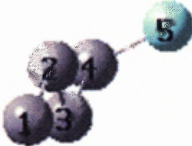|	name	Z	q	E	L	μ	α	V	A	r_{avg}
fluoromethane 593-53-3	(1) CH$_3$	9	0.630	-39.4037	5.2E-05	-0.757	11.7	203.65	150.31	3.988
	(2) F	9	-0.630	-99.7725	4.6E-06	0.145	4.8	113.48	85.49	3.204
fluoroethane 353-36-6	(1) CH$_3$	9	0.064	-39.7379	5.0E-06	-0.262	13.6	217.00	146.84	4.122
	(2) CH$_2$	8	0.572	-38.8264	-5.9E-04	-0.705	10.2	147.67	92.05	4.093
	(3) F	9	-0.636	-99.7540	1.5E-05	0.129	4.8	113.14	80.32	3.238
1-fluoropropane (gauche) 460-13-9	(1) CH$_3$	9	0.025	-39.7295	-7.3E-06	-0.278	13.5	217.95	141.68	4.163
	(2) CH$_2$	8	0.057	-39.1473	-4.7E-04	-0.291	11.5	157.67	98.47	4.214
	(3) CH$_2$	8	0.555	-38.8386	-1.1E-03	-0.707	10.2	147.66	89.42	4.142
	(4) F	9	-0.637	-99.7409	6.8E-06	0.128	4.8	111.89	76.74	3.235
1-fluoropropane (trans) 460-13-9-conf	(1) CH$_3$	9	0.010	-39.7279	2.8E-06	-0.257	13.9	220.16	142.28	4.153
	(2) CH$_2$	8	0.069	-39.1490	-4.4E-05	-0.263	11.5	156.01	87.20	4.187
	(3) CH$_2$	8	0.557	-38.8392	-2.9E-04	-0.710	9.8	147.15	83.99	4.138
	(4) F	9	-0.637	-99.7422	1.5E-05	0.131	5.3	113.13	81.42	3.238

Table G.13 (Continued)

		name	Z	q	E	L	μ	α	V	A	r_{avg}
2-fluoropropane 420-26-8		(1) CH$_3$	9	0.052	-39.7471	1.3E-05	-0.270	13.4	216.31	136.79	4.140
		(2) CH	7	0.535	-38.2280	-1.1E-03	-0.630	8.6	91.86	42.67	4.157
		(3) CH$_3$ ^{y}	9	0.052	-39.7471	1.3E-05	-0.270	13.4	216.31	136.79	4.140
		(4) F	9	-0.640	-99.7408	3.2E-05	0.112	5.0	112.15	73.48	3.253
1-fluorobutane (gauche) 2366-52-1		(1) CH$_3$	9	-0.005	-39.7335	2.3E-06	-0.271	14.2	221.34	144.22	4.178
		(2) CH$_2$	8	0.044	-39.1349	-7.2E-05	-0.293	11.6	156.24	82.61	4.256
		(3) CH$_2$	8	0.041	-39.1623	-9.1E-05	-0.291	11.8	156.82	85.59	4.267
		(4) CH$_2$	8	0.556	-38.8334	-6.7E-04	-0.707	10.1	147.64	136.50	4.144
		(5) F	9	-0.638	-99.7320	6.3E-06	0.128	5.0	111.86	76.08	3.235
1-fluorobutane (trans) 2366-52-1-conf		(1) CH$_3$	9	0.008	-39.7302	3.8E-06	-0.270	13.7	220.70	141.83	4.160
		(2) CH$_2$	8	0.018	-39.1367	-6.8E-04	-0.312	12.0	159.21	82.96	4.235
		(3) CH$_2$	8	0.053	-39.1645	6.2E-04	-0.263	11.1	155.04	79.78	4.233
		(4) CH$_2$	8	0.558	-38.8346	-3.0E-05	-0.708	10.3	147.09	85.51	4.136
		(5) F	9	-0.636	-99.7327	1.5E-05	0.130	5.2	113.15	80.38	3.237
2-fluorobutane		(1) CH$_3$	9	0.052	-39.7425	1.1E-05	-0.272	13.4	216.28	137.09	4.162
		(2) CH	7	0.518	-38.2422	-1.7E-03	-0.636	8.4	91.78	64.35	4.266
		(3) CH$_2$	8	0.044	-39.1585	-2.3E-04	-0.297	11.4	156.67	85.42	4.257
		(4) CH$_3$	9	0.025	-39.7260	-9.3E-06	-0.284	14.2	217.91	139.31	4.164
		(5) F	9	-0.641	-99.7315	1.9E-05	0.111	5.0	110.66	71.10	3.251

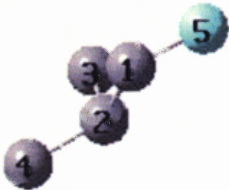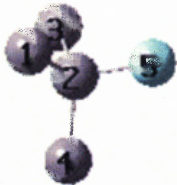|	name	Z	q	E	L	μ	α	V	A	r_{avg}
2-methyl-1-fluoropropane	(1) CH$_2$	8	0.547	-38.8485	-9.0E-05	-0.710	10.0	146.32	83.97	4.193
	(2) CH	7	0.082	-38.5324	-1.6E-04	-0.229	9.8	97.43	41.18^{tr}	4.323^{tr}
	(3) CH$_3$	9	0.013	-39.7415	-6.1E-06	-0.283	13.7	216.69	132.44	4.197
	(4) CH$_3$	9	-0.003	-39.7405	6.9E-06	-0.260	13.8	218.97	132.65	4.198
	(5) F	9	-0.639	-99.7331	9.0E-06	0.130	5.0	111.55	75.10	3.237
2-methyl-2-fluoropropane 353-61-7	(1) CH$_3$	9	0.044	-39.7556	1.8E-05	-0.281	13.5	214.81	129.09	4.182
	(2) C	6	0.510	-37.6111	4.7E-04	0.547	6.6	37.10	0.00	4.413
	(3) CH$_3$ ^{y}	9	0.044	-39.7556	1.7E-05	-0.281	13.5	214.81	129.09	4.182
	(4) CH$_3$ ^{y}	9	0.044	-39.7556	1.8E-05	-0.281	13.5	214.81	129.09	4.182
	(5) F	9	-0.641	-99.7304	4.7E-05	0.096	5.2	110.75	68.02	3.260

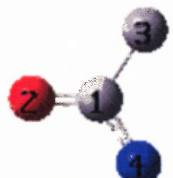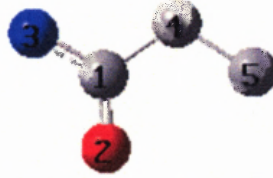**Table G.14** AIM Properties for Atoms and Functional Groups: Amides

		name	Z	q	E	L	μ	α	V	A	r_{avg}
methanamide 75-12-7		(1) CH	7	1.565	-37.5614	2.6E-04	-0.867	7.6	95.71	62.68	3.887
		(2) O	8	-1.174	-75.6616	1.6E-05	0.466	8.4	142.31	101.07	3.494
		(3) NH_2	9	-0.390	-56.1295	1.8E-05	0.778	12.2^+	183.13	124.62	3.835
		N	7	-1.328	-55.3123	1.1E-04	0.077	9.7	128.19	70.63	3.766
		H	1	0.475	-0.4050	-2.5E-05	-0.185	1.2	27.64	26.69	2.480
		H	1	0.464	-0.4123	-2.6E-05	-0.188	1.3	27.92	26.92	2.490
ethanamide 60-35-5		(1) C	6	1.498	-36.9921	-4.2E-04	0.748	5.2	39.28	8.77	3.593
		(2) O	8	-1.188	-75.6431	2.9E-05	0.440	8.5	141.78	95.34	3.506
		(3) CH_3	9	0.077	-39.7605	4.8E-06	-0.251	13.7	215.73	140.25	4.111
		(4) NH_2	9	-0.388	-56.1036	-2.6E-05	0.728	12.0^+	181.48	117.80	3.852
		N	7	-1.312	-55.2780	1.4E-05	0.043	9.6	125.82	66.10	3.776
		H	1	0.469	-0.4083	-2.3E-05	-0.187	1.2	27.88	26.51	2.491
		H	1	0.455	-0.4173	-2.2E-05	-0.191	1.2	28.48	25.17	2.508
propanamide 79-05-0		(1) C	6	1.493	-36.9939	-4.3E-04	0.767	1.4	38.64	8.12	3.585
		(2) O	8	-1.187	-75.6251	3.9E-06	0.436	7.2	139.02	86.63	3.506
		(3) NH_2	9	-0.400	-56.1043	-3.5E-05	0.732	12.0^+	181.29	117.37	3.856
		(4) CH_2	8	0.030	-39.1789	-3.6E-05	-0.272	12.0^{tr}	158.33	85.47	4.195
		(5) CH_3	9	0.063	-39.7341	-1.9E-05	-0.303	11.1	215.28	137.69	4.149
		N	7	-1.330	-55.2806	6.3E-05	0.042	9.6^{tr}	125.91	64.82	3.783
		H	1	0.472	-0.4075	-2.5E-05	-0.186	1.2^{tr}	27.78	25.88	2.488
		H	1	0.458	-0.4162	-2.4E-05	-0.190	1.2^{tr}	28.43	24.93	2.499

	name	Z	q	E	L	μ	α	V	A	r_{avg}
	(1) C	6	1.484	-36.9967	-5.0E-04	0.769	5.0	37.36	11.23	3.681
	(2) O	8	-1.191	-75.6128	-3.8E-06	0.432	8.0	139.02	84.19	3.507
	(3) CH$_2$	8	0.021	-39.1896	-8.0E-04	-0.287	12.0	157.03	80.12	4.229
butanamide	(4) CH$_2$	8	0.074	-39.1340	8.7E-06	-0.316	11.5	155.03	83.13	4.193
541-35-5	(5) CH$_3$	9	0.001	-39.7504	-1.2E-04	-0.278	13.7	219.57	133.40	4.148
	(6) NH$_2$	9	-0.390	-56.0879	2.1E-06	0.733	12.6$^+$	182.01	115.64	3.858
	N	7	-1.315	-55.2629	4.3E-05	0.050	10.0	126.27	63.88	3.776
	H	1	0.468	-0.4087	-1.9E-05	-0.187	1.3	27.99	26.11	2.491
	H	1	0.457	-0.4163	-1.6E-04	-0.191	1.3	28.26	25.13	2.516
	(1) C	6	1.483	-37.0092	-3.2E-04	0.775	0.6	37.01	3.85	3.662
	(2) O	8	-1.193	-75.6087	3.2E-06	0.425	6.5	139.20	82.82	3.490
	(3) NH$_2$	9	-0.396	-56.0953	2.8E-05	0.741	12.8$^+$	181.74	116.20	3.880
2-methylpropanamide	(4) CH$_3$	9	0.037	-39.7432	-8.7E-06	-0.289	11.9	216.28	133.76	4.186
563-83-7	(5) CH	7	0.032	-38.5741	3.4E-04	-0.283	1.0	98.62	35.82	4.353
	(6) CH$_3$	9	0.037	-39.7432	-1.0E-05	-0.289	11.8	216.28	135.04	4.186
	N	7	-1.326	-55.2722	9.2E-05	0.055	9.6^{tr}	126.53	66.49	3.771
	H	1	0.460	-0.4152	-2.3E-05	-0.190	1.7	27.93	24.09	2.473
	H	1	0.470	-0.4080	-2.1E-05	-0.187	1.5^{tr}	27.90	26.03	2.484

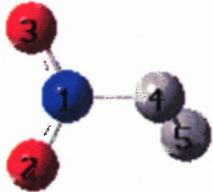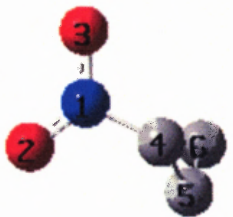**Table G.15** AIM Properties for Atoms and Functional Groups: Nitros

	name	Z	q	E	L	μ	α	V	A	r_{avg}
nitromethane 75-52-5	(1) N	7	0.433	-54.2732	9.0E-04	0.736	6.9	51.17	13.62	3.490
	(2) O	8	-0.475	-75.2409	3.6E-05	-0.268	6.9	124.64	89.13	3.400
	(3) O ^{y}	8	-0.475	-75.2409	3.6E-05	-0.268	6.9	124.64	89.13	3.400
	(4) CH$_3$	9	0.518	-39.5831	6.6E-06	-0.546	11.6	200.02	139.93	4.045
nitroethane 79-24-3	(1) N	7	0.428	-54.2710	-9.1E-06	0.738	7.4	50.48	26.11	3.515
	(2) O	8	-0.484	-75.2213	2.8E-05	-0.272	6.7	123.49	82.47	3.417
	(3) O	8	-0.476	-75.2157	3.5E-05	-0.270	7.0	124.74	89.44	3.415
	(4) CH$_2$	8	0.422	-39.0183	-1.7E-04	-0.526	10.3	145.36	83.74	4.115
	(5) CH$_3$	9	0.109	-39.7474	-9.2E-06	-0.260	12.9	212.50	137.31	4.118
1-nitropropane 108-03-2	(1) N	7	0.425	-54.2603	-5.4E-05	0.738	7.2	48.51	7.18	3.519
	(2) O	8	-0.486	-75.2046	3.3E-05	-0.272	6.6	123.32	80.83	3.403
	(3) O	8	-0.477	-75.1994	3.2E-05	-0.270	6.9	124.86	87.55	3.426
	(4) CH$_2$	8	0.407	-39.0265	-6.8E-04	-0.531	10.4	144.89	78.75	4.110
	(5) CH$_2$	8	0.093	-39.1563	-4.4E-05	-0.275	11.1	153.57	84.21	4.180
	(6) CH$_3$	9	0.037	-39.7603	-8.2E-05	-0.247	13.3	216.06	132.44	4.142
2-nitropropane 79-46-9	(1) N	7	0.429	-54.2683	5.8E-04	0.739	7.4	49.22	8.45	3.561
	(2) O	8	-0.488	-75.2061	2.3E-05	-0.274	6.8	121.65	76.76	3.392
	(3) O	8	-0.479	-75.1987	3.5E-05	-0.272	6.7	124.88	88.06	3.417
	(4) CH$_3$	9	0.088	-39.7568	-3.6E-07	-0.261	13.2	211.96	135.62	4.155
	(5) CH	7	0.362	-38.4239	-6.2E-05	-0.483	8.6	90.59	35.16	4.241
	(6) CH$_3$ ^{y}	9	0.088	-39.7568	1.1E-06	-0.261	13.2	211.96	135.62	4.155

352

Table G.16 (Continued)

name	Z	q	E	L	μ	α	V	A	r_{avg}
hydrogen 1333-74-0									
(1) H	1	0.000	-0.5755	-6.7E-07	-0.103	2.5	56.64	57.01	2.687
(2) H [y]	1	0.000	-0.5755	-6.7E-07	-0.103	2.5	56.64	57.01	2.687
water 7732-18-5									
(1) O	8	-1.202	-75.4687	5.4E-05	0.366	7.7	151.66	118.44	3.502
(2) H	1	0.601	-0.3412	-2.4E-03	-0.170	0.9	21.64	22.47	2.417
(3) H [y]	1	0.601	-0.3412	-2.4E-03	-0.170	0.6	21.64	22.47	2.417
ammonia 7664-41-7									
(1) N	7	-1.232	-55.0447	5.9E-05	-0.033	9.4	150.71	102.17	3.740
(2) H	1	0.411	-0.4333	-2.7E-05	-0.215	1.5	31.04	29.70	2.549
(3) H [y]	1	0.411	-0.4333	-2.1E-05	-0.215	1.5	31.04	29.70	2.549
(4) H [y]	1	0.411	-0.4333	-2.1E-05	-0.215	1.5	31.04	29.70	2.549
nitrogen 7727-37-9									
(1) N	7	0.000	-54.7075	3.1E-08	-0.610	5.9	132.52	98.59	3.471
(2) N [y]	7	0.000	-54.7075	3.1E-08	-0.610	5.9	132.52	98.59	3.471

Table G.16 (Continued)

	name	Z	q	E	L	μ	α	V	A	r_{avg}
nitrous oxide 10024-97-2	(1) N	7	0.272	-54.6844	8.5E-06	0.505	6.5	75.84	42.22	3.429
	(2) N	7	0.094	-54.4516	2.0E-05	-0.828	6.9	127.34	97.70	3.495
	(3) O	8	-0.365	-75.2565	1.2E-05	-0.253	6.5	122.25	102.03	3.348
neon 7440-01-9	(1) Ne	10	0.000	-128.3552	-4.3E-05	0.000	2.0	105.29	107.83	2.929
oxygen 7782-44-7	(1) O	8	0.000	-75.0099	-1.6E-07	-0.416	4.4	115.14	91.27	3.280
	(2) O ^{y}	8	0.000	-75.0099	-1.6E-07	-0.416	4.4	115.14	91.27	3.280
hydrogen cyanide 74-90-8	(1) C	6	0.941	-37.4391	1.6E-04	1.066	6.5	91.98	60.54	3.598
	(2) N	7	-1.150	-55.3218	-1.6E-05	0.600	8.6	168.13	114.61	3.619
	(3) H	1	0.209	-0.5159	-1.0E-05	-0.124	1.7	39.18	37.09	2.517

REFERENCES

Abdurahmanov, A. A., Rahimova, R. A., & Imanov, L. M. (1970). Microwave Spectrum of Normal Propyl Alcohol. *Physics Letters A, 32,* 123-124.

Abrams, D. S., & Prausnitz, J. M. (1975). Statistical thermodynamics of Liquid Mixtures: A New Expression for the Excess Gibbs Energy of Partly or Completely Miscible Systems. *AIChE Journal, 21,* 116-128.

AIM2000 1.0 [Computer Software]. (1998). Bielefeld, Germany: University of Applied Sciences.

AIMPAC [Computer Software], (1982). Hamilton, Canada: McMaster University.

Alonso, C., Montero, E. A., Chamorro, C. R., Segovia, J. J., Martín, M. C., and Villamañán, M. A. Vapor-Liquid Equilibrium of Octane-Enhancing Additives in Gasolines 5. Total Pressure Data and G^E for Binary and Ternary Mixtures Containing 1,1-Dimethylpropyl Methyl Ether (TAME), 1-Propanol, and n-Hexane at 313.15 K. *Fluid Phase Equilibria, 212,* 81-95.

Amos, A. T., & Hall, G. G. (1961). Single Determinant Wave Functions. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 263,* 483-493.

Anisimov, M. A., & Sengers, J. V. (2000) Critical Region. In J. V. Sengers, R. F. Kayser, C. J. Peters & H. J. White Jr. (Eds.), *Equations of State for Fluids and Fluid Mixtures* (pp. 381-434). New York: Elsevier.

Bader, R. F. W., Carroll, M. T., Cheeseman, J. R., & Chang, C. (1987). Properties of Atoms in Molecules: Atomic Volumes. *Journal of the American Chemical Society, 109,* 7968-7979.

Bader, R. F. W., Keith, T. A., Gough, K. M., & Laidig, K. E. (1992). Properties of Atoms in Molecules: Additivity and Transferability of Group Polarizabilities. *Molecular Physics, 75,* 1167-1189.

Bader, R. F. W., & Bayles, D. (2000). Properties of Atoms in Molecules: Group Additivity. *Journal of Physical Chemistry A, 104,* 5579-5589.

Bader, R. F. W. (1990). *Atoms in Molecules: A Quantum Theory.* Oxford: Clarendon.

Bader, R. F. W., & Becker, P. (1988). Transferability of Atomic Properties and the Theorem of Hohenberg and Kohn. *Chemical Physics Letters, 148,* 452-458.

Barker, J. A., & Smith, F. (1954). Statistical Thermodynamics of Associated Solutions. *Journal of Chemical Physics, 22*, 375-380.

Bartha, F., Kapuy, O., Kozmutza, C., & Van Alsenoy, C. (2003). Analysis of Weakly Bound Structures: Hydrogen Bond and the Electron Density in a Water Dimer. *Journal of Molecular Structure (Theochem), 666-667*, 117-122.

Becke, A. D. (1992). Density-functional Thermochemistry. III. The Role of Exact Exchange. *Journal of Chemical Physics, 98*, 5648-5652.

Benson, S. W. (1976). *Thermochemical Kinetics* (2nd ed.). New York: Wiley.

Besler, B. H., Merz, K. M., & Kollman, P. A. (1990). Atomic Charges Derived from Semiempirical Methods. *Journal of Computational Chemistry, 11*, 431-439.

Biegler-König, F. W., Bader, R. F. W., & Tang, T. (1982). Calculation of the Average Properties of Atoms in Molecules. II. *Journal of Computational Chemistry, 3*, 317-328.

Biegler-König, F. W. (2000). Calculation of Atomic Integration Data. *Journal of Computational Chemistry, 21*, 1040-1048.

Bondi, A. (1964). van der Waals Volumes and Radii. *Journal of Physical Chemistry, 68*, 441-451..

Born, M., & Mayer, J. E. (1932). Zur Gittertheorie der Ionenkristalle. *Zeitschrift für Physik, 75*, 1.

Boys, S. F. (1950). Electronic Wave Functions. I. A General Method of Calculation for the Stationary States of Any Molecular System. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 200*, 542-554.

Boys, S. F. (1960). The Integral Formulae for the Variational Solution of the Molecular Many-Electron Wave Equations in Terms of Gaussian Functions with Direct Electronic Correlation. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 258*, 402-411.

Brdarski, S., & Karlström, G. (1998). Modeling of the Exchange Repulsion Energy. *Journal of Physical Chemistry A, 102*, 8182-8192.

Breneman, C. M., & Wiberg, K. B. (1990). Determining Atom-Centered Monopoles from Molecular Electrostatic Potentials. The Need for High Sampling Density in Formamide Conformational Analysis. *Journal of Computational Chemistry, 11*, 361-373.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983). CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *Journal of Computational Chemistry, 4,* 187-217.

Browne, J. C., & Poshusta, R. D. (1962). Quantum-Mechanical Integrals over Gaussian Atomic Orbitals. *Journal of Chemical Physics, 36,* 1933-1937.

Buckingham, A. D. (1967). Permanent and Induced Molecular Moments and Long-Range Intermolecular Forces. *Advances in Chemical Physics, 12,* 107-143.

Buckingham, R. A., & Corner, J. (1947). Tables of Second Virial and Low-Pressure Joule-Thompson Coefficients for Intermolecular Potentials with Exponential Repulsions. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 189,*_118-129.

Carnahan, N. F., & Starling, K. E. (1969). Equation of State for Nonattracting Rigid Spheres. *Journal of Chemical Physics, 51,* 635-636.

Chen, B., Potoff, J. J., & Siepmann, J. I. (2001). Monte Carlo Calculations for Alcohols and Their Mixtures with Alkanes. Transferable Potentials for Phase Equilibria. 5. United-Atom Descriptions of Primary, Secondary, and Tertiary Alcohols. *Journal of Physical Chemistry B, 105,* 3093-3104.

Chirlian, L. E., & Francl, M. M. (1987). Atomic Charges Derived from Electrostatic Potentials: A Detailed Study. *Journal of Computational Chemistry, 8,* 894-905.

Computational Chemistry Comparison and Benchmark Database. (2004). *NIST Standard Reference Database 101.* Retrieved March 7, 2005, from http://srdata.nist.gov.

Constantinou, L., & Gani, R. (1994). New Group Contribution Method for Estimating Properties of Pure Compounds. *AIChE Journal, 40,* 1697-1710.

Cortés-Guzmán, F., & Bader, R. F. W. (2003). Transferability of Group Energies and Satisfaction of the Virial Theorem. *Chemical Physics Letters, 379,* 183-192.

Cox, S. R., & Williams, D. E. (1981). Representation of the Molecular Electrostatic Potential by a Net Atomic Charge Model. *Journal of Computational Chemistry, 2,* 304-323.

Curtiss, L. A., Frurip, D. J., & Blander, M. (1979). Studies of Molecular Association in H_2O and D_2O Vapors by Measurement of Thermal Conductivity. *Journal of Chemical Physics, 71,* 2703-2711.

Dahlhoff, G., Pfenning, A., Hammer, H., & van Oorschot, M. (2000). Vapor-Liquid Equilibria in Quarternary Mixtures of Dimethyl Ether + *n*-Butane + Ethanol + Water. *Journal of Chemical and Engineering Data, 45,* 887-892.

Dewar, M. J. S., Zoebisch, E. G., Healy, E. F., & Stewart, J. J. P. (1985). AM1: A New General Purpose Quantum Mechanical Molecular Model. *Journal of the American Chemical Society, 107,* 3902-3909.

Dipole Moments. (2004). In D. R. Lide (Ed.), *CRC Handbook of Chemistry and Physics* (85th ed.) (pp. 9-45 to 9-51). New York: CRC Press.

Donohue, M. D., & Prausnitz, J. M. (1975). Combinatorial Entropy of Mixing Molecules that Differ in Size and Shape: Simple Approximation for Binary and Multicomponent Mixtures. *Canadian Journal of Chemistry, 53,* 1586.

Dunning, T. H. (1989). Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *Journal of Chemical Physics, 90,* 1007-1023.

Dymond, J. H., & Smith, E. B. (1969). *The Virial Coefficients of Gases: A Critical Compilation.* New York: Oxford.

Engkvist, O., Åstrand, P., & Karlström, G. (2000). Accurate Intermolecular Potentials Obtained from Molecular Wave Functions: Bridging the Gap between Quantum Chemistry and Molecular Simulations. *Chemical Reviews, 100,* 4087-4108.

Flory, P. J. (1942). Thermodynamics of High Polymer Solutions. *Journal of Chemical Physics, 10,* 51-61.

Fredenslund, A., Jones, R. L., & Prausnitz, J. M. (1975). Group Contribution Estimation of Activity Coefficients in Nonideal Liquid Mixtures. *AIChE Journal, 21,* 1086-1099.

Gaussian 98W [Computer Software]. (1998). Pittsburgh, PA: Gaussian, Inc.

GaussView 2.1 [Computer Software]. (2000). Pittsburgh, PA: Gaussian, Inc.

Gmehling, J., Li, J., & Schiller, M. (1993). A Modified UNIFAC Model. 2. Present Parameter Matrix and Results for Different Thermodynamic Properties. *Industrial Engineering and Chemistry Research, 32,* 178-193.

Graña, A. M., & Mosquera, R. A. (1999). The Transferability of the Carbonyl Group in Aldehydes and Ketones. *Journal of Chemical Physics, 110,* 6606-6616.

Graña, A. M., & Mosquera, R. A. (2000). Transferability in Aldehydes and Ketones. II. Alkyl Chains. *Journal of Chemical Physics, 113,* 1492-1500.

Gray, C. G., & Gubbins, K. E. (1984). *Theory of Molecular Fluids, Volume 1.* Oxford: Clarendon Press.

Guggenheim, E. A. (1944a). Statistical Thermodynamics of Mixtures with Zero Energies of Mixing. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 183,* 203-212.

Guggenheim, E. A. (1944b). Statistical Thermodynamics of Mixtures with Non-Zero Energies of Mixing. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 183,* 213-227

Guggenheim, E. A. (1952). *Mixtures.* Oxford, Clarendon.

Hart, J. R., & Rappé, A. K. (1992). van der Waals Functional Forms for Molecular Simulation. *Journal of Chemical Physics, 97,* 1109-1115.

Hayes, I. C., & Stone, A. J. (1984). Matrix Elements Between Determinantal Wavefunctions of Non-Orthogonal Orbitals. *Molecular Physics, 53,* 83-105.

Hazma, A. & Mayer, I. (2001a). Overlap Repulsion with Löwdin's Pairing Theorem. *International Journal of Quantum Chemistry, 82,* 53-59.

Hazma, A. & Mayer, I. (2001b). Overlap Repulsion with Löwdin's Pairing Theorem II. The Leading Term. *International Journal of Quantum Chemistry, 82,* 105-112.

Head-Gordon, M. (1996). Quantum Chemistry and Molecular Processes. *Journal of Physical Chemistry, 100,* 13213-13225.

Hehre, W. J., Radom, L., Schleyer, P. v. R., & Pople, J. A. (1986). *Ab Initio Molecular Orbital Theory.* New York: Wiley.

Hill, T. L. (1960). *An Introduction to Statistical Mechanics.* Reading, MA: Addison Wesley.

Hirschfelder, J. O., Curtiss, C. F., & Bird, R. B. (1964). *Molecular Theory of Gases and Liquids.* New York: Wiley.

Hohenberg, P., & Kohn, W. (1964). Inhomogeneous Electron Gas. *Physical Review, 136,* B864-B871.

Huggins, M. L. (1942). Theory of Solutions of High Polymers. *Journal of the American Chemical Society, 64,* 1712-1719.

Jensen, J. (1996). Modeling Intermolecular Exchange Integrals between Nonorthogonal Molecular Orbitals. *Journal of Chemical Physics, 104,* 7795-7796.

Jeziorski, B., Moszynski, R., & Szalewicz, K. (1994). Perturbation Theory Approach to Intermolecular Potential Energy Surfaces of van der Waals Complexes. *Chemical Reviews, 94,* 1887-1930.

Johnson, B. G., Gill, P. M. W., & Pople, J. A. (1993). The Performance of a Family of Density Functional Methods. *Journal of Chemical Physics, 98,* 5612-5626.

Jorgensen, W. L., Madura, J. D., & Swenson, C. J. (1984). Optimized Intermolecular Potential Functions for Liquid Hydrocarbons. *Journal of the American Chemical Society, 106,* 6638-6646.

Kairys, V., & Jensen, J. H. (1999). Evaluation of the Charge Penetration Energy Between Non-Orthogonal Molecular Orbitals using the Spherical Gaussian Overlap Approximation. *Chemical Physics Letters, 315,* 140-144.

Kehiaian, H. V., Grolier, J. E., & Benson, G. C. (1978). Thermodynamics of Organic Mixtures. A Generalized Quasichemical Theory in Terms of Group Surface Interactions. *Journal de Chimie Physique, 75,* 1031-1048.

Kihara, T. (1978). *Intermolecular Forces.* New York: Wiley.

Klamt, A. (1995). Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena. *Journal of Physical Chemistry, 99,* 2224-2235.

Klamt, A., Jonas, V., Bürger, T., & Lohrenz, J. C. W. (1998). Refinement and Parameterization of COSMO-RS. *Journal of Physical Chemistry A, 102,* 5074-5085.

Klamt, A., Krooshof, G. J. P., & Taylor, R. (2002). COSMOSPACE: Alternative to Conventional Activity-Coefficient Models. *AIChE Journal, 48,* 2332-2349.

Knowles, P. J., & Meath, W. J. (1986a). Non-Expanded Dispersion Energies and Damping Functions for Ar_2 and Li_2. *Chemical Physics Letters, 124,* 164-171.

Knowles, P. J., & Meath, W. J. (1986b). Non-Expanded Dispersion and Induction Energies, and Damping Functions, for Molecular Interactions with Application to HF...He. *Molecular Physics, 59,* 965-984.

Knowles, P. J., & Meath, W. J. (1987). A Separable Method for the Calculation of Dispersion and Induction Energy Damping Functions with Applications to the Dimers Arising from He, Ne, and HF. *Molecular Physics, 60,* 1143-1158.

Knox, D. E. (1982). *Thermodynamic Mixture Properties and Lattice Theories.* Unpublished doctoral dissertation, Rensselaer Polytechnic Institute, Troy, New York.

Knox, D. E., Van Ness, H. C., & Hollinger, H. B. (1984) A Model for Representation of GE and HE. *Fluid Phase Equilibria, 15,* 267-285.

Knox, D. E. (1987). A One-Parameter Group-Contribution Model for Liquid Mixtures. *Journal of Solution Chemistry, 6,* 625-634.

Kohn, W., & Sham, L. J. (1965). Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review, 140,* A1133-A1138.

Kozmutza, C., Varga, I., & Udvardia, L. (2003). Comparison of the Extent of Hydrogen Bonding in H_2O-H_2O and H_2O-CH_4 Systems. *Journal of Molecular Structure (Theochem), 666-667,* 95-97.

Kryachko, E., & Scheiner, S. (2004). CH•••F Hydrogen Bonds, Dimers of Fluoromethanes. *Journal of Physical Chemistry A, 108,* 2527-2535.

Levenspiel, O. (1996). *Understanding Engineering Thermo.* Upper Saddle River, NJ: Prentice Hall.

Levine, I. N. (2000). *Quantum Chemistry* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Lias, S. G. (2004). Ionization Energies of Gas-Phase Molecules. In D. R. Lide (Ed.), *CRC Handbook of Chemistry and Physics* (85th ed.) (pp. 10-167 to 10-182). New York: CRC Press.

Lin, S. T., & Sandler, S. I. (1999). Infinite Dilution Activity Coefficients from Ab Initio Solvation Calculations. *AIChE Journal, 45,* 2606-2618 .

Lin, S. T., & Sandler, S. I. (1999). Prediction of Octanol-Water Partition Coefficients Using a Group Contribution Solvation Model. *Industrial Engineering and Chemistry Research, 38,* 4081-4091.

Lin, S., & Sandler, S. I. (2000). Multipole Corrections To Account for Structure and Proximity Effects in Group Contribution Methods: Octanol-Water Partition Coefficients. *Journal of Physical Chemistry A, 104,* 7099-7105.

Lin, S., & Sandler, S. I. (2002). A Priori Phase Equilibrium Prediction from a Segment Contribution Solvation Model. *Industrial Engineering and Chemistry Research, 41,* 899-913.

Lorenzo, L., & Mosquera, R. A. (2002). Approximate Transferability in Alkanes Revisited. *Chemical Physics Letters, 356,* 305-312.

Löwdin, P. (1962). Band Theory, Valence Bond, and Tight-Binding Calculations. *Journal of Applied Physics, 33,* 251-280.

Mandado, M., Mosquera, R. A., & Graña, A. M. (2002). Approximate Transferability in Alkanols. *Journal of Molecular Structure (Theochem), 584,* 221-234.

Mandado, M., Vila, A., Graña, A. M., Mosquera, R. A., & Cioslowski, J. (2003). Transferability of Energies of Atoms in Organic Molecules. *Chemical Physics Letters, 371,* 739-743.

Martin, M. G., & Siepmann, J. I. (1998). Transferable Potentials for Phase Equilibria. 1. Unites-Atom Descriptions of *n*-Alkanes. *Journal of Physical Chemistry B, 102,* 2567-2577.

Martinez, G. M. (1995). Flory and Guggenheim Lattice Statistics Reinterpreted and Extended to Include Molecular Clustering and Chemical Dissociation. *Journal of Chemical Physics, 103,* 9813-9824.

Marrero, J., & Gani, R. (2001). Group-Contribution Based Estimation of Pure Component Properties, *Fluid Phase Equilibria, 183-184,* 183-208.

Marrero-Morejon, J., & Pardillo-Fontdevilla, E. (1999). Estimation of Pure Compound Properties using Group-Interaction Contributions. *AIChE Journal, 45,* 615-621

Maurer, G. M., & Prausnitz, J. M. (1978). On the Derivation and Extension of the UNIQUAC Equation. *Fluid Phase Equilibria, 2,* 91-99.

Mayer, I. (1997). Simple Proof of the Pairing Theorem. *International Journal of Quantum Chemistry, 63,* 31-33.

McDermott, C., & Ashton, N. (1977). Note on the Definition of Local Composition. *Fluid Phase Equilibria, 1,* 33-35.

McQuarrie, D. A. (2000). *Statistical Mechanics.* Sausalito, CA: University Science Books.

Miller, T. M. (2004). Atomic and Molecular Polarizabilities. In D. R. Lide (Ed.), *CRC Handbook of Chemistry and Physics* (85th ed.) (pp. 10-167 to 10-182). New York: CRC Press.

Miyamoto, H., & Watanabe, K. (2000). A Thermodynamic Property Model for Fluid-Phase Propane. *International Journal of Thermophysics, 21,* 1045-1072.

Møller, C., & Plesset, M. S. (1934). Note on an Approximation Treatment for Many-Electron Systems. *Physical Review, 46,* 618-622.

Momany, F. A. (1978). Determination of Partial Atomic Charges from Ab Initio Molecular Electrostatic Potentials. Application to Formamide, Methanol, and Formic Acid. *Journal of Physical Chemistry, 82,* 592-601.

Mulliken, R. S. (1955). Electronic Population Analysis on LCAO-MO Molecular Wave Functions. I. *Journal of Chemical Physics, 23,* 1833-1846.

Murrell, J. N., Carter, S., Farantos, S. C., Huxley, P., & Varandas, A. J. C. (1984). *Molecular Potential Energy Functions.* New York: Wiley.

Panayiotou, C. (2003a). Equation-of-State Models and Quantum Mechanics Calculations. *Industrial Engineering and Chemistry Research, 42,* 1495-1507.

Panayiotou, C. (2003b). The QCHB Model of Fluids and Their Mixtures. *Journal of Chemical Thermodynamics, 35,* 349-381.

Perdew, J. P., Burke, K., & Wang, Y. (1996). Generalized Gradient Approximation for the Exchange-Correlation Hole of a Many-Electron System. *Physical Review B, 54,* 16533-16539.

Poling, B. E., Prausnitz, J. M., & O'Connell, J. P. (2001). *The Properties of Gases and Liquids* (5th ed.). New York: McGraw-Hill.

Quiñónez, P. B., Vila, A., Graña, A. M., & Mosquera, R. A. (2003). AIM Study on the Influence of Fluorine Atoms on the Alkyl Chain. *Chemical Physics, 287,* 227-236.

Rappé, A. K., & Bernstein, E. R. (2000). Ab Initio Calculation of Nonbonded Interactions: Are We There Yet? *Journal of Physical Chemistry A, 104,* 6117-6128.

Rice, P., & El-Nikheli, A. (1995). Isothermal Vapour-Liquid Equilibrium Data for the Systems *n*-Pentane with *n*-Hexane, *n*-Octane, and *n*-Decane. *Fluid Phase Equilibria, 107,* 257-267.

Rosen, N. (1931). Calculation of Interaction between Atoms with *s*-Electrons. *Physical Review, 38,* 255-276.

Sanchez, I. C., & Lacombe, R. H. (1976). An Elementary Molecular Theory of Classical Fluids. Pure Fluids. *Journal of Physical Chemistry, 80,* 2352-2362.

Sandler, S. I. (1994). Thermophysical Properties: What Have We Learned Recently, and What Do We Still Need to Know? *International Journal of Thermophysics, 15,* 1013-1035.

Sayegh, S. G., & Vera, J. H. Lattice-Model Expressions for the Combinatorial Entropy of Liquid Mixtures: A Critical Discussion. *Chemical Engineering Journal, 19*, 1-10.

Seo, J., Lee, J., & Kim, H. (2000). Isothermal Vapor-Liquid Equilibria for Ethanol and *n*-Pentane System at the Near Critical Region. *Fluid Phase Equilibria, 172*, 211-219.

Sigfridsson, E., & Ryde, U. (1998). Comparison of Methods for Deriving Atomic Charges from the Electrostatic Potentials and Moments. *Journal of Computational Chemistry, 19*, 377-395.

Smirnova, N. A., & Victorov, A. I. (1987). Thermodynamic Properties of Pure Fluids and Solutions from the Hole Group-Contribution Model. *Fluid Phase Equilibria, 34*, 235-263.

Smith, J. M., Van Ness, H. C., & Abbott, M. M. (1996). *Introduction to Chemical Engineering Thermodynamics* (5th ed.). New York: McGraw-Hill.

Stanley, H. E. (1971). *Introduction to Phase Transitions and Critical Phenomena.* New York: Oxford.

Staverman, A. J. (1950). The Entropy of High Polymer Solutions. *Recl. Trav. Chim. Pays-Bais, 69*, 163-174.

Stefanov, B. B., & Cioslowski, J. (1995). An Efficient Approach to Calculation of Zero-Flux Atomic Surfaces and Generation of Atomic Integration Data. *Journal of Computational Chemistry, 16*, 1394-1404.

Stewart, J. J. P. (1989). Optimization of Parameters for Semiempirical Methods I. Method. *Journal of Computational Chemistry, 10*, 209-220.

Stone, A. J. (1996). *The Theory of Intermolecular Forces.* Oxford: Clarendon.

Stubbs, J. M., Potoff, J. J., & Siepmann, J. I. (2004). Transferable Potentials for Phase Equilibria. 6. United-Atom Description for Ethers, Glycols, Ketones, and Aldehydes. *Journal of Physical Chemistry B, 108*, 17596-17605.

Szabo, A., & Ostlund, N. S. (1989). *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory.* Mineola, NY: Dover.

Taimoori, M., & Panayiotou, C. (2001). The Non-Random Distribution of Free Volume in Fluids: Non-Polar Systems. *Fluid Phase Equilibria, 192*, 155-169.

Teodorescu, M., & Rasmussen, P. (2001). High-Pressure Vapor-Liquid Equilibria in the Systems Nitrogen + Dimethyl Ether, Methanol + Dimethyl Ether, Carbon Dioxide + Dimethyl Ether + Methanol, and Nitrogen + Dimethyl Ether + Methanol. *Journal of Chemical and Engineering Data, 46*, 640-646.

Tompa, H. (1956). Polymer Solutions. London: Butterworths.

Ungerer, P., Beauvals, C., Delhommelle, J., Boutin, A., Rousseau, B., & Fuchs, A. H. (2000) Optimization of the Anisotropic Intermolecular Potential for n-Alkanes. *Journal of Chemical Physics, 112,* 5499-5510.

Unlu, O., Gray, N. H., Gerek, Z. N., & Elliot, J. R. (2004). Transferable Step Potentials for the Straight-Chain Alkanes, Alkenes, Alkynes, Ethers, and Alcohols. *Industrial Engineering and Chemistry Research, 43,* 1788-1793.

Vila, A., Carballo, E., & Mosquera, R. A. (2000). AIM Study of the Transferability of the Oxygen Atom in Linear Ethers. *Canadian Journal of Chemistry, 78,* 1535-1543.

Vila, A., & Mosquera, R.A. (2001). Transferability in Alkyl Monoethers. II. Methyl and Methylene Fragments. *Journal of Chemical Physics, 115,* 1264-1273.

Wang, C., Li, H., Zhu, L., & Han, S. (2001). Isothermal and Isobaric Vapor + Liquid Equilibria of N,N-Dimethylformamide + n-Propanol + n-Butanol. *Fluid Phase Equilibria, 189,* 119.127.

Wawrzyniak, P. K., Panek, J., Latajkaa, Z., & Lundell, J. (2004). Theoretical Study of the Complex Between Formic Acid and Argon. *Journal of Molecular Structure, 691,* 115-122.

Wilson, G. M. (1964). Vapor-Liquid Equilibrium. XI. A New Expression for the Excess Free Energy of Mixing. *Journal of the American Chemical Society, 86,* 127-130.

Wu, H. S., & Sandler, S. I. (1991a). Use of ab Initio Quantum Mechanics Calculations in Group Contribution Methods. 1. Theory and the Basis for Group Identifications. *Industrial Engineering and Chemistry Research, 30,* 881-889.

Wu, H. S., & Sandler, S. I. (1991b). Use of ab Initio Quantum Mechanics Calculations in Group Contribution Methods. 2. Test of New Groups in UNIFAC. *Industrial Engineering and Chemistry Research, 30,* 889-897.

Yeganegi, M., Pylypenko, D., Hon, A., Choi, C., Zsoldos, Z., Chass, G. A., & Csizmadia, I. G. (2003). Intermolecular Interactions of Small Biologically Active Molecules: Acetone, Methylamine and Water; Methyl Phosphate, Water and Divalent Ions; Phenol and Water; N-Ac-L-Gly-NH-Me and Water. *Journal of Molecular Structure (Theochem), 666-667,* 99-107.