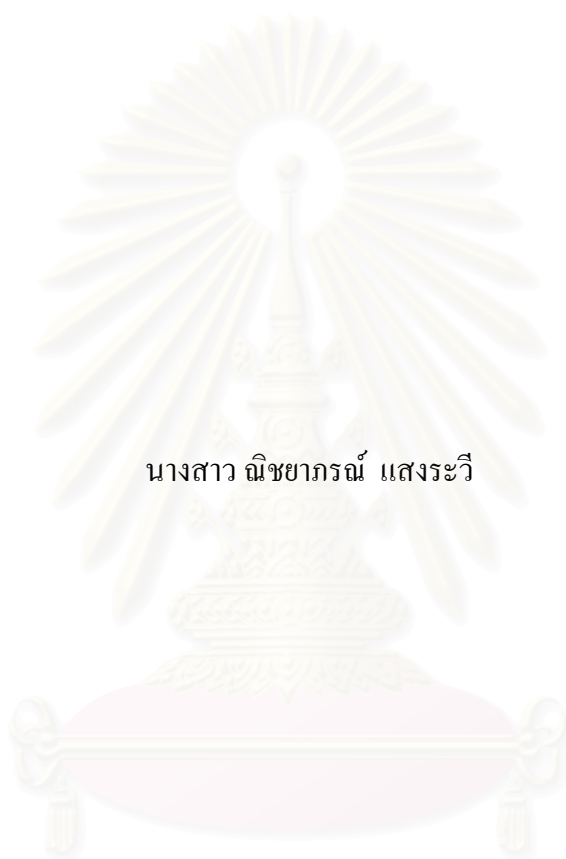


การเขียนโปรแกรมฟอร์แทรนสำหรับการปรับเส้นโค้งและการปรับปรุงคุณภาพ
ของฟังก์ชันศักระหว่างโมเดล



นางสาว ณิชยาภรณ์ แสงระวี

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต

สาขาวิชาวิทยาการคณนา ภาควิชาคณิตศาสตร์

คณะวิทยาศาสตร์ จุฬาลงกรณ์มหาวิทยาลัย

ปีการศึกษา 2544

ISBN 974-03-1436-8

ลิขสิทธิ์ของจุฬาลงกรณ์มหาวิทยาลัย

FORTRAN PROGRAMMING FOR CURVE FITTING AND
IMPROVING QUALITY OF INTERMOLECULAR
PAIR POTENTIAL FUNCTION



Miss Nichayaporn Sangrawee

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

A Thesis Submitted in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Computational Science

Department of Mathematics

Faculty of Science

Chulalongkorn University

Academic Year 2001

ISBN 974-03-1436-8

Thesis Title Fortran Programming for Curve Fitting and Improving Quality
of Intermolecular Pair Potential Function
By Miss Nichayaporn Sangrawee
Field of study Computational Science
Thesis Advisor Associate Professor Supot Hannongbua, Ph.D.

Accepted by the Faculty of Science, Chulalongkorn University in Partial
Fulfillment of the Requirements for the Master's Degree

..... Deputy Dean for Administrative Affairs
(Associate Professor Pipat Karntiang, Ph.D.) Acting Dean, Faculty of Science

Thesis Committee

..... Chairman
(Associate Professor Vudhichai Parasuk, Ph.D.)

..... Thesis Advisor
(Associate Professor Supot Hannongbua, Ph.D.)

..... Member
(Krung Sinapiromsaran, Ph.D.)

ณิชากรณ์ แสงระวี : การเขียนโปรแกรมฟอร์แทรนสำหรับการปรับเส้นโค้งและการปรับปรุงคุณภาพของฟังก์ชันศักย์คู่ระหว่างโมเลกุล (FORTRAN PROGRAMMING FOR CURVE FITTING AND IMPROVING QUALITY OF INTERMOLECULAR PAIR POTENTIAL FUNCTION). อ. ที่ปรึกษา : รศ. ดร. สุพจน์ หารหนองบัว, 95 หน้า.
ISBN 974-03-1436-8.

ได้พัฒนาโปรแกรม MRQFIT เวอร์ชัน 1.0 ขึ้น เพื่อใช้สำหรับการปรับและปรับปรุงคุณภาพของฟังก์ชันศักย์คู่ระหว่างโมเลกุล ซึ่งเป็น โปรแกรมที่เขียนด้วยภาษาฟอร์แทรนของ FORTRAN-77 บนระบบยูนิกซ์ ในขั้นตอนของการปรับเส้นโค้ง พลังงานเสถียรของสารเชิงซ้อน ซึ่งคำนวณบนพื้นฐานของการคำนวณเคมีควอนตัม จะนำมาฟิตโดยใช้วิธีเลเวนเบิร์ก-มาร์ควาร์ดต์ ซึ่งเป็นวิธีการประดิษฐ์ฟังก์ชันที่ไม่เป็นเชิงเส้นในหลายมิติโดยวิธีกำลังสองน้อยที่สุด ในขั้นตอนของการปรับปรุงคุณภาพของฟังก์ชันนั้น ฟังก์ชันจะถูกทดสอบคุณภาพโดยใช้วิธีของเบเวริดจ์และตรวจสอบหาหลุมพลังงานที่อาจเกิดขึ้นได้ในฟังก์ชัน ในการศึกษาโปรแกรมนี้ได้พัฒนาและทดสอบโดยเลือกใช้ระบบของ Li^+-NH_3 และ NH_3-NH_3 ซึ่งจะเห็นว่าพลังงานที่คำนวณจากฟังก์ชันและจากวิธีการทางเคมีควอนตัมมีความสอดคล้องกันอย่างชัดเจน โดยเฉพาะอย่างยิ่งในบริเวณที่พลังงานเป็นลบ โปรแกรมนี้สามารถใช้สำหรับการปรับเส้นโค้งและการปรับปรุงคุณภาพของฟังก์ชันศักย์คู่ระหว่างโมเลกุล

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

ภาควิชาคณิตศาสตร์..... ลายมือชื่อนิสิต

สาขาวิชาวิทยาการคณนา..... ลายมือชื่ออาจารย์ที่ปรึกษา

ปีการศึกษา ..2544..... ลายมือชื่ออาจารย์ที่ปรึกษาร่วม

4172283323 : MAJOR MATHEMATICS

KEY WORD : POTENTIAL FUNCTION / LEVENBERG-MARQUARDT ALGORITHM /
BEVERIDGE ALGORITHM

NICHAYAPORN SANGRAWEE : FORTRAN PROGRAMMING FOR CURVE
FITTING AND IMPROVING QUALITY OF INTERMOLECULAR PAIR
POTENTIAL FUNCTION. THESIS ADVISOR : ASSOC. PROF. SUPOT
HANNONGBUA, Ph.D., 95 pp. ISBN 974-03-1436-8.

The program MRQFIT version 1.0 has been developed in order to fit and improve quality of intermolecular pair potential function. This program implements in Fortran language using FORTRAN-77 on UNIX-system. In the curve fitting step, the stabilization energies of the complex which have been calculated by means of quantum chemical calculations are fitted using a multidimensional non-linear least-squares procedure by the Levenberg-Marquardt algorithm. In the step of function improvement, the quality of the function is examined using Beveridge algorithm, and then, the false minima which may be available in the function are searched. In this study, the $\text{Li}^+\text{-NH}_3$ and $\text{NH}_3\text{-NH}_3$ systems are used to develop and test the program. It is clearly seen that the energies obtained from the function are in good agreement with the SCF energies, especially for the attractive regions. The program can be used to fit interaction energies to a functional form and improve quality of the function between the two molecules.

Department ... Mathematics Student's signature

Field of study Computational Science. Advisor's signature

Academic year 2001 Co-advisor's signature

ACKNOWLEDGEMENTS

This thesis was completely finished with the excellent suggestions from my advisor, Associate Professor Dr. Supot Hannongbua. I therefore would like to express my sincerest gratitude for giving me his guiding, advising, understanding and encouraging. I am very obliged to Associate Professor Dr. Vudhichai Parasuk and Dr. Krung Sinapiromsaran for their careful reading for improving this manuscript and valuable suggestions as thesis committee.

I also would like to gratefully acknowledge to Graduate school of Chulalongkorn University for granting a research scholarship and the Austrain-Thai Center for Computer Assisted Chemical Education and Research as computer resource supplements and other facilities.

My absolute acknowledgement is dedicated to my parents and sisters for their understanding, encouragement and support throughout the entire study.



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

CONTENTS

	Pages
ABSTRACT IN THAI.....	iv
ABSTRACT IN ENGLISH.....	v
ACKNOWLEDGEMENTS.....	vi
LIST OF FIGURES.....	x
LIST OF TABLES.....	xiii
CHAPTER 1 INTRODUCTION.....	1
1.1 Computational Science.....	1
1.2 Computer Simulations in Chemistry.....	3
1.3 Objectives of This Study.....	4
1.4 Scope of This Study.....	5
CHAPTER 2 THE INTERMOLECULAR POTENTIAL FUNCTION: HEART OF THE SIMULATION.....	6
2.1 Total Interaction Energy of the Simulated System.....	6
2.2 Pair Potential.....	8
2.3 Analytical Form of Potential Functions.....	11
CHAPTER 3 NONLINEAR LEAST SQUARES.....	14
3.1 Introduction.....	14
3.2 Nonlinear Models.....	17
3.3 Levenberg-Marquardt Method.....	21

	Pages
CHAPTER 4 DETAIL OF CALCULATIONS.....	25
4.1 Development of Intermolecular Potential Function.....	25
4.1.1 Selection of Representative Geometries of the Pairs.....	26
4.1.2 The SCF Calculations.....	27
4.2 Program Implementation.....	28
4.2.1 PART 1: Input Elementary Data.....	29
4.2.2 PART 2: Selection of the Functional Form.....	32
4.2.3 PART 3: Set an Initial Guess for the Parameters.....	33
4.2.4 PART 4: Fitting Energies to the Function.....	35
4.2.5 PART 5: Testing the Quality of the Function.....	40
4.2.6 PART 6: Search for False Minima of the Function.....	42
4.2.7 PART 7: Write Output.....	44
4.3 Specification of the systems for this study.....	45
CHAPTER 5 RESULTS AND DISCUSSION.....	49
5.1 Development of Lithium Ion - Ammonia Potential Function.....	49
5.1.1 Input Elementary Data.....	49
5.1.2 Selection of the Functional Form.....	52
5.1.3 Set an Initial Guess for the Parameters.....	52
5.1.4 Fitting Energies to the Functional Form.....	53
5.1.5 Testing the Quality of the Function.....	55
5.1.6 Search for False Minima of the Function.....	58
5.2 Development of Ammonia - Ammonia Potential Function.....	67
5.2.1 Input Elementary Data.....	67
5.2.2 Selection of the Functional Form.....	67

	Pages
5.2.3 Set an Initial Guess for the Parameters.....	67
5.2.4 Fitting Energies to the Functional Form.....	68
5.2.5 Testing the Quality of the Function.....	70
5.2.6 Search for False Minima of the Function.....	73
REFERENCES.....	83
APPENDICES.....	86
Appendix I MRQFIT version 1.0 (Program Manual).....	87
Appendix II Output Files for MRQFIT Program.....	90
CURRICULUM VITAE.....	95

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

LIST OF FIGURES

Figures	Pages
1.1 The connection between experiment, theory, and computer simulation [1]	2
2.1 The general procedure for constructing the intermolecular potential function (Steps 2 - 4 are executed more than once).....	7
2.2 The Lennard-Jones potential.....	8
2.3 The Lennard-Jones potential as a sum of attractive (Ar^{-6}) and repulsive (Br^{-12}) components.....	9
2.4 (a) the hard-sphere potential, (b) the square-well potential and (c) the soft-sphere potential with repulsion parameter $\kappa=1$	11
4.1 The procedure for constructing the potential function by means of quantum chemical calculations.....	26
4.2 The sequence of MRQFIT program version 1.0.....	29
4.3 The structure of input elementary data.....	30
4.4 The flowchart of part 1; Input Elementary Data.....	32
4.5 The flowchart of part 2; Selection of the Functional Form.....	33
4.6 The flowchart of part 3; Set an Initial Guess for the Parameters.....	34
4.7 The flowchart of part 4; Fitting Energies to the Function.....	39
4.8 The flowchart of part 5; Testing the Quality of the Function.....	42
4.9 The flowchart of part 6; Search for False Minima of the Function.....	44
4.10 The flowchart of part 7; Write Output.....	45
4.11 Definition of geometries variable (θ, ϕ) for the configuration of ammonia-lithium ion.....	46
4.12 Definition of geometries variable (θ, ϕ) for the ammonia- ammonia configuration.....	48
5.1 An input file for the $\text{Li}^+\text{-NH}_3$ system (see Figure 4.3).....	51

Figures	Pages
5.2 The stabilization energies obtained from the DZP-ab initio (ΔE_{SCF}) and from the fitting analytical potential (ΔE_{FIT}), which the derivatives calculated using central different technique, with the parameter given in Table 5.4 ($\phi = 0^\circ$ see Figure 4.11).....	63
5.3 The stabilization energies obtained from the DZP-ab initio (ΔE_{SCF}) and from the fitting analytical potential (ΔE_{FIT}), which the derivatives calculated analytically, with the parameter given in Table 5.5 ($\phi = 0^\circ$ see Figure 4.11).....	64
5.4 Comparison of the stabilization energies from the DZP-ab initio (ΔE_{SCF}) and the potential function (ΔE_{FIT}), which the derivatives calculated using central different technique, with the final values of the fitting parameters given in Table 5.4.....	65
5.5 Comparison of the stabilization energies from the DZP-ab initio (ΔE_{SCF}) And the potential function (ΔE_{FIT}), which the derivatives calculated analytically, with the final values of the fitting parameters given in Table 5.5.....	66
5.6 The stabilization energies obtained from the DZP-ab initio (ΔE_{SCF}) and from the fitting analytical potential (ΔE_{FIT}), which the derivatives calculated using central different technique, with the parameter given in Table 5.10 ($\theta = 150^\circ$ see Figure 4.12).....	78
5.7 The stabilization energies obtained from the DZP-ab initio (ΔE_{SCF}) and from the fitting analytical potential (ΔE_{FIT}), which the derivatives calculated analytically, with the parameter given in Table 5.11 ($\theta = 150^\circ$ see Figure 4.12).....	79

Figures	Pages
5.8 Comparison of the stabilization energies from the DZP-ab initio (ΔE_{SCF}) and the potential function (ΔE_{FIT}), which the derivatives calculated using central different technique, with the final values of the fitting parameters given in Table 5.10.....	80
5.9 Comparison of the stabilization energies from the DZP-ab initio (ΔE_{SCF}) and the potential function (ΔE_{FIT}), which the derivatives calculated analytically, with the final values of the fitting parameters given in Table 5.11.....	81



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

LIST OF TABLES

Tables	Pages
4.1 Cartesian coordinates (in Angstroms) and atomic net charges (in a.u.) of ammonia molecule obtained from <i>ab initio</i> calculations using DZP basis set	47
5.1 The selected trajectories, in term of θ and ϕ angles (see Figure 4.11), for lithium ion around ammonia.....	50
5.2 The optimized parameters (the derivatives were approximated using the central different technique) representing interaction of N and H atoms of ammonia with lithium ion (energy and r are in kcal.mol ⁻¹ and angstroms, respectively).....	53
5.3 The optimized parameters (the derivatives were calculated directly) representing interaction of N and H atoms of ammonia with lithium ion (energy and r are in kcal.mol ⁻¹ and angstroms, respectively).....	54
5.4 The final optimized parameters (the derivatives were approximated using the central different technique) representing interaction of N and H atoms of ammonia with lithium ion (energy and r are in kcal.mol ⁻¹ and angstroms, respectively).....	56
5.5 The final optimized parameters (the derivatives were calculated analytically) representing interaction of N and H atoms of ammonia with lithium ion (energy and r are in kcal.mol ⁻¹ and angstroms, respectively)..	57
5.6 The energies from <i>ab initio</i> calculations (ΔE_{SCF}) and from the pair potential function which the derivatives calculated using central different technique (ΔE_{FIT}^1) and calculated analytically (ΔE_{FIT}^2) (energy and r are in kcal.mol ⁻¹ and angstroms, respectively).....	59

Tables	Pages
5.7 Lithium ion-ammonia interaction energies from <i>ab initio</i> calculation (ΔE_{SCF}) and from the pair potential function which the derivatives calculated using central different technique (ΔE_{FIT}^1) and calculated analytically (ΔE_{FIT}^2), for (a) $\theta = 90$ and $\phi = 0$, (b) $\theta = 60$ and $\phi = 0$, (c) $\theta = 30$ and $\phi = 0$ (energy and r are in kcal.mol ⁻¹ and angstroms, respectively).....	60
5.8 The optimized parameters (the derivatives were approximated using the central different technique) representing interaction of atoms between the two ammonia molecules (energy and r are in kcal.mol ⁻¹ and angstroms, respectively).....	69
5.9 The optimized parameters (the derivatives were calculated analytically) representing interaction of atoms between the two ammonia molecules (energy and r are in kcal.mol ⁻¹ and angstroms, respectively).....	70
5.10 The final optimized parameters (the derivatives were approximated using the central different technique) representing interaction of atoms between the two ammonia molecules (energy and r are in kcal.mol ⁻¹ and angstroms, respectively).....	72
5.11 The final optimized parameters (the derivatives were calculated analytically) representing interaction of atoms between the two ammonia molecules (energy and r are in kcal.mol ⁻¹ and angstroms, respectively)...	73
5.12 The energies from <i>ab initio</i> calculations (ΔE_{SCF}) and from the pair potential function which the derivatives calculated using central different technique (ΔE_{FIT}^1) and calculated analytically (ΔE_{FIT}^2) (energy and r are in kcal.mol ⁻¹ and angstroms, respectively).....	75

Tables

Pages

5.13 Ammonia-ammonia interaction energies from <i>ab initio</i> calculation (ΔE_{SCF}) and from the pair potential function which the derivatives calculated using central different technique (ΔE_{FIT}^1) and calculated analytically (ΔE_{FIT}^2), for (a) $\theta = 150$ and $\phi = 30$, (b) $\theta = 150$ and $\phi = 60$ (energy and r are in kcal.mol ⁻¹ and angstroms, respectively).....	76
--	----



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

CHAPTER 1

INTRODUCTION

1.1 Computational Science

It is generally agreed that rapid development in the field of computer technology plays a significant role in many fields of science. In the past, physical properties were characterized by experiments and theories. In experiment, a system is subjected to measurements and results. In theory, a model of the system is developed in form of a set of mathematical equations. Computational science techniques are used to simulate physical events and to process large amounts of generated or collected data. The use of simulation in research is now established as a third basic methodology of doing scientific research, joining theory and experiments.

As such, computational science must be distinguished from computer science. Computer science focuses on the computer itself; it is the science and engineering of computer systems, including hardware and software. On the other hand, computational science refers to the application of computers solving problems in science and engineering. In addition to a particular scientific knowledge or engineering application, it includes mathematics (mathematical modeling, numerical analysis), computer science (computer architecture, programming, networks), and scientific visualization.

Computational science refers to the knowledge and techniques required to perform computer simulations instead of experiments for evaluating interesting informations of systems, particularly, those systems for which experiments are prohibitive. Due to the rapid development of computer technology, computer

simulations have been expanded at a tremendous speed and application of these and other simulation methods have penetrated in almost every field of mathematics, physics, chemistry, biology, other applied science and engineering fields. Results of the simulations may also be compared with those of real experiments. In the first place, this is a test of the underlying model used in a computer simulation. Eventually, if the model is acceptable, the simulator hopes the model will offer insights to the experimentalist, and assist in the interpretation of new results. This dual roles of simulation, as a bridge between models and theoretical predictions on one hand, and between models and experimental results on the other, is illustrated in Figure 1.1. Because of this connecting role, and the way in which simulations are conducted and analysed, these techniques are often termed “computer experiments”.

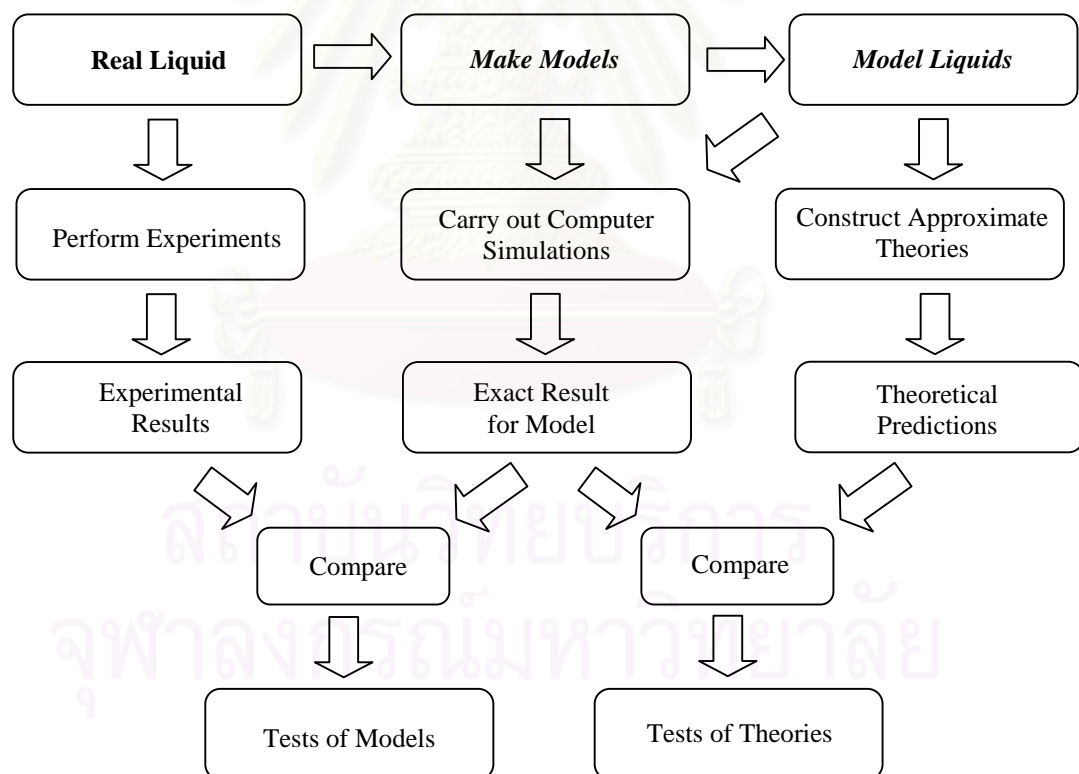


Figure 1.1 The connection between experiment, theory, and computer simulation [1].

1.2 Computer Simulations in Chemistry

Quantum chemical investigations, which still have strong limits in the size of molecular system and the related computational time, have been a useful tool for predicting configurations of molecules, stabilization energies, excitation energies, force constants, and other physical data. Although they often produce data in good agreement with experiment, this approach is not reflect all properties of a condensed system with large amounts of particles.

Models of chemical systems and specialized theoretical chemistry are mainly constructed in the form of quantum chemical calculations and statistical mechanics simulations. These tools are widely applicable for many systems, especially for investigation of microscopic properties and lead directly and perhaps easily to a set of interesting results or macroscopic properties of the system and can fill the gap between theory and experiment.

Computer simulation methods, such as Monte Carlo practically introduced by Metropolis [2], and Molecular Dynamics introduced by Alder [3], are tools for studying statistical and dynamical properties of rather large ensembles of molecules based on accurate potentials derived from quantum chemical calculations. Both simulation methods can be used to evaluate structural properties of systems, included some data unacquirable by experimental techniques. Therefore, theoretical investigations of the liquid state and solutions have rapidly developed and numerous studies based on these methods for solution have been published [4-18].

1.3 Objectives of This Study

As already mentioned, computer simulations is a powerful tool for developing our understanding of molecular systems at the atomic/molecular level. However, there is major limitation preventing computer simulations from fulfilling their true potential in providing scientific insights and aiding the industrial development of new materials. The problem is the accuracy of the fundamental input into the simulations, i.e., the intermolecular model potential, which quantifies the forces acting between the particles. In other word, due to the shortage of reasonable intermolecular potentials, the development in the field of solution research by means of computer simulations seem to be slower than it should be according to technological advances.

Nowadays, many simulations are undertaken in order to model real systems. The simulations seek to produce results that are in agreement with experiment. The first stage in such a computer simulation is to find a model for the intermolecular interactions in the chosen system, which is sufficiently realistic to give acceptable results. Since the quality of the results from the simulations definitely depend on the quality of the input, it is important to be able to assess the accuracy of an intermolecular potential. The accuracy of intermolecular potentials is well known for the rare gases. Much current work on deriving accurate intermolecular potentials is concentrating on small rigid polyatomic molecules. There are, of course, difficulties to do so since there are no generally reliable simple procedures for developing intermolecular potentials.

To improve this situation, it is the main objectives of this study, i.e., to develop part of the semi-automatic Fortran programming for the development of intermolecular pair potential function. In continuation from the first part which generates the stabilization energies of the complex, molecules and/or ions, in all

possible configurations [19], this work fits those interaction energies to a functional form and improves quality of the function.

FORTTRAN (FORmula TRANslation) has been the most common programming language for computer simulation, in view of its suitability for scientific applications, with a variety of mathematical functions, and the availability of FORTRAN compilers producing efficient code. In this study, FORTRAN-77 has been used to develop the program [20-21].

1.4 Scope of This Study

Since the development of the potential function by means of quantum chemical calculations are divided into four consecutive steps, namely, (i) selection of representative geometries of the pairs; (ii) performance the *ab initio* calculations; (iii) fitting of pair interaction energies to a functional form and (iv) improving quality of the function, the scope of this work is to write the second part of the semi-automatic Fortran programming regarding to steps (iii) and (iv). The $\text{Li}^+\text{-NH}_3$ and $\text{NH}_3\text{-NH}_3$ systems are used to develop and test the performance of the program.

This thesis is organized into 5 Chapters. After brief review of Computational Science, Computer Simulations in Chemistry, objectives and scope of this study in Chapter 1, theoretical background of the intermolecular potential function, and Nonlinear least squares are presented in Chapters 2 and 3, respectively. Then, detailed calculations, results and discussion are given in Chapters 4 and 5, respectively.

CHAPTER 2

THE INTERMOLECULAR POTENTIAL FUNCTION: HEART OF THE SIMULATION

2.1 Total Interaction Energy of the Simulated System

The Monte Carlo (MC) and Molecular Dynamics (MD) simulation techniques have been well established and are used by numerous scientists at present. The most important factor, the results of calculations, depends undoubtedly on the potential functions used. At the same time, it is necessary to evaluate energy and force as rapidly as possible in the course of the simulation. A potential function describes the interaction energy among particles in that system. In general, the total configurational interaction energy of the N particles system can be written as

$$\Delta E = \sum_i^N v_1(r_i) + \sum_i^N \sum_{j>i}^N v_2(r_i, r_j) + \sum_i^N \sum_{j>i}^N \sum_{k>j>i}^N v_3(r_i, r_j, r_k) + \dots \quad (2.1)$$

where the first term in the right hand side of equation (2.1), $v_1(r_i)$, represents the effect of an external field. The second term, v_2 called the pair potential, is the term with highest. The third term, v_3 , called three-body connecting interaction that sometime becomes very significant in case of condensed systems. In most of published simulation, however, they were not included, due to the requirement of the large computer time. Four-body (and higher) terms are expected to be small in comparison to v_2 and v_3 . In practice, one assumes to represent ΔE of the system by the pair potential only, known as the pair-wise additive approximation, whereas the remaining terms are often referred to as non-additive corrections. The basic idea is that the total configurational energy of the many particles system is approximately written as a sum of the individual interaction between the particles.

The general procedures for developing the *ab initio* intermolecular potential function consist of four steps as illustrated in Figure 2.1.

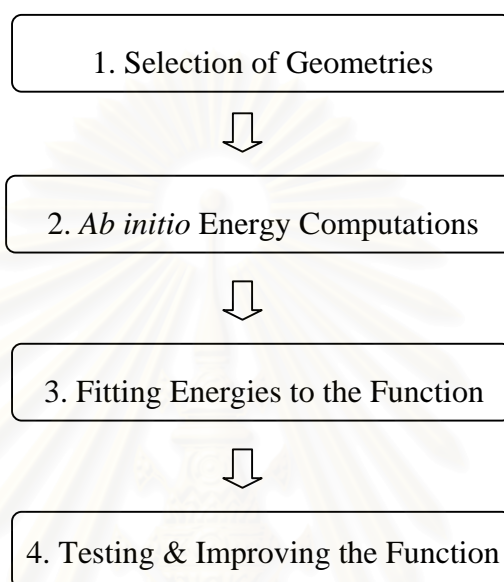


Figure 2.1 The general procedure for constructing the intermolecular potential function (Steps 2 - 4 are executed more than once).

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

2.2 Pair Potential

As mentioned above, majorities of simulations approximate ΔE simply by the pair potential term. So far, there have been several types of pair potentials used in computer simulations. The Lennard-Jones 6-12 pair potential [22], is one of the commonly used. It has the form as given bellows;

$$v^{LJ}(r) = 4\varepsilon \left[-\left(\frac{\sigma}{r}\right)^6 + \left(\frac{\sigma}{r}\right)^{12} \right] \quad (2.2).$$

The function contains only two adjustable parameters: the collision diameter σ (the separation for which the energy is zero) and the well depth ε . Physical meaning of these parameters are graphically illustrated in Figure 2.2.

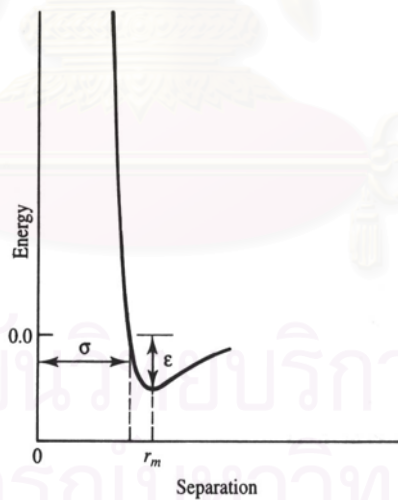


Figure 2.2 The Lennard-Jones potential.

The Lennard-Jones equation may also be expressed in term of the separation at which the energy passes through a minimum, r_m . At this separation, the first partial derivative of the energy with respect to the internuclear distance is zero (i.e. $\partial v / \partial r = 0$) from which it can easily be shown that $r_m = 2^{1/6}\sigma$. We can thus also write Lennard-Jones 6-12 potential function as follows,

$$v^{LJ}(r) = \epsilon \left[-2 \left(\frac{r_m}{r} \right)^6 + \left(\frac{r_m}{r} \right)^{12} \right] \quad (2.3)$$

or

$$v^{LJ}(r) = -\frac{A}{r^6} + \frac{B}{r^{12}} \quad (2.4)$$

where A is equal to $2\epsilon r_m^6$ (or $4\epsilon\sigma^6$) and B is equal to ϵr_m^{12} (or $4\epsilon\sigma^{12}$).

The Lennard-Jones potential is characterised by an attractive part that varies as r^{-6} and a repulsive part as r^{-12} . These two components are drawn in Figure 2.3.

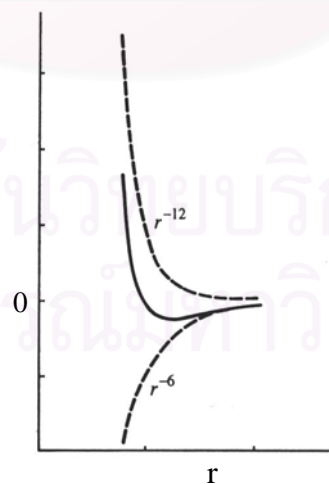


Figure 2.3 The Lennard-Jones potential as a sum of attractive (Ar^{-6}) and repulsive (Br^{-12}) components.

The other forms of pair potential, which are very simple and convenient to implement in the computer simulation and the liquid-state theory, are:

a) The hard-sphere potential

$$v^{\text{HS}}(r) = \begin{cases} \infty & (r < \sigma) \\ 0 & (r \geq \sigma) \end{cases} \quad (2.5)$$

b) The square-well potential

$$v^{\text{SW}}(r) = \begin{cases} \infty & (r < \sigma_1) \\ -\varepsilon & (\sigma_1 \leq r < \sigma_2) \\ 0 & (r \geq \sigma_2) \end{cases} \quad (2.6)$$

c) The soft-sphere potential

$$v^{\text{SS}}(r) = \varepsilon \left(\frac{\sigma}{r} \right)^\kappa = ar^{-\kappa} \quad (2.7)$$

where κ is a parameter, often chosen to be an integer. The soft-sphere potentials contain no attractive part.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

These three potentials are graphically shown in Figure 2.4.

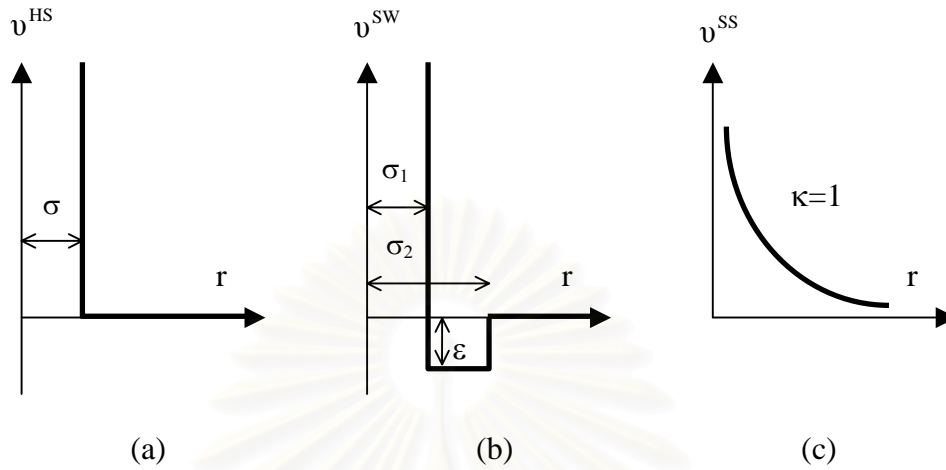


Figure 2.4 (a) the hard-sphere potential, (b) the square-well potential and (c) the soft-sphere potential with repulsion parameter $\kappa = 1$.

2.3 Analytical Form of Potential Functions

Among the analytical forms of potential function, if the data points can be fitted into many forms, the simplest function is normally chosen. One of the necessary criteria is that the selected function must be differentiable; otherwise it will not be able to be used in the molecular dynamics simulations in which force acting on each particle has to be derived from the function and its derivative. Given two molecules M and N, the pair interaction potential ΔE_{MN} might contain explicitly both angular and radial dependency. Practically, one normally use only the radial function depending on the interatomic distance r_{ij} with i and j being atoms belonging to molecules M and N, respectively. For accuracy and flexibility reasons, one would like to use a fairly long series of terms. On the other hand, the longer the series, the larger the number of associated fitting parameters, and the larger the number of machine cycles required to compute interaction energies in the simulation.

The two general forms of the potential functions are expressed as:

$$(1) \Delta E_{MN} = \sum_i^k \sum_j^l \left(-\frac{A_{ij}}{r_{ij}^x} + \frac{B_{ij}}{r_{ij}^y} + \frac{q_i q_j}{r_{ij}} + Z_{ij} \right); x \neq y \quad (2.8)$$

$$(2) \Delta E_{MN} = \sum_i^k \sum_j^l \left(-\frac{A_{ij}}{r_{ij}^x} + B_{ij} \exp(-C_{ij} r_{ij}) + \frac{q_i q_j}{r_{ij}} + Z_{ij} \right) \quad (2.9)$$

where k and l are the number of atoms on molecules M and N , respectively, A_{ij} , B_{ij} and C_{ij} are fitting parameters which represent interaction between atom i of molecule M and atom j of molecule N , and q_i , q_j are the atomic net charges of each atom i and atom j of molecules M and N , respectively. These charges were normally taken from Mulliken populations analysis in the quantum chemical calculations [23]. In additions, although x and y are normally set at 6 and 12, respectively, however, these values can be adjust. ΔE_{MN} is the interaction energy. The first two terms formally describe the short-range attractive and repulsive interactions, respectively. The third term describes the long-range Coulomb interaction, and the last terms, which has the form as in equation (2.10), will be added in order to obtain a good statistical fitting.

$$Z_{ij} = 0 + \frac{E_{ij}}{r_{ij}^m} + \frac{F_{ij}}{r_{ij}^n} + \frac{G_{ij}}{r_{ij}^o} + \dots \quad (2.10)$$

where E_{ij} , F_{ij} , and G_{ij} are fitting parameters; m , n and o are integers which differ from x and y .

Among the 2 forms, equation (2.8) is normally used for the pair of weak interactions (the stabilization energy is higher than $-20 \text{ kcal.mol}^{-1}$) such as ligand-

ligand or ligand-solvent interactions. On the other hand, equation (2.9) is suitable for the pair of strong interactions such as ion-ion, ion-ligand or ion-solvent.



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

CHAPTER 3

NONLINEAR LEAST SQUARES

3.1 Introduction

Given a set of observations, one often wants to condense and summarize the data by fitting it to a “model” that depends on adjustable parameters. Frequently, the model is simply a convenient class of functions, such as polynomials or Gaussians, and the fit supplies the appropriate coefficients. Otherwise, the model’s parameters come from some underlying theory that the data are supposed to satisfy such as coefficients of rate equations in a complex network of chemical reactions, or orbital elements of a binary star. Modeling can also be used as a kind of constrained interpolation, where one wants to extend a few data points into a continuous function, but with some underlying idea of what that function should look like.

The basic approach in all cases is usually the same: One chooses or designs a *figure-of-merit function* (“merit function”, for short) that measures the agreement between the data and the model with a particular choice of parameters. The merit function is conventionally arranged so that small values represent close agreement. The parameters of the model are then adjusted to achieve a minimum in the merit function, yielding *best-fit parameters*. The adjustment process is thus a problem of minimization in many dimensions.

There are important issues that go beyond the finding of best-fit parameters. Data are generally not exact. They are subjected to *measurement errors*. Thus, typical data never exactly fit the model that is being used, even when that model is correct. One needs the means to assess whether or not the model is appropriate, that is, one needs to test the *goodness-of-fit* against some useful statistical standard.

One usually also needs to know the accuracy with which parameters are determined by the data set. In other word, one needs to know the likely errors of the best-fit parameters.

The important message one wants to deliver is that fitting of parameters is not the end-all of parameter estimation. To be genuinely useful, a fitting procedure should provide (i) parameters, (ii) error estimates on the parameters, and (iii) a statistical measurement of goodness-of-fit. When the third item suggests that the model is an unlikely match to the data, then items (i) and (ii) are probably worthless.

Chi-Square Fitting

Suppose that one fits N data points (x_i, y_i) $i = 1, \dots, N$, to a model that has M adjustable parameters a_j , $j = 1, \dots, M$. The model predicts a functional relationship between the measured independent and dependent variables,

$$y(x) = y(x; a_1 \dots a_M) \quad (3.1)$$

where the dependence on the parameters is indicated explicitly on the right-hand side.

One wants to minimize to get fitted values for the a_j 's. The first thing that comes to mind is the familiar least-squares fit,

$$\text{minimize over } a_1 \dots a_M : \sum_{i=1}^N [y_i - y(x_i; a_1 \dots a_M)]^2 \quad (3.2).$$

If each data point (x_i, y_i) has its own, known standard deviation σ_i , then the model parameters are obtained by minimizing the quantity

$$\chi^2 = \sum_{i=1}^N \left[\frac{y_i - y(x_i; a_1 \dots a_M)}{\sigma_i} \right]^2 \quad (3.3)$$

called the “chi-square.”

To whatever extent the measurement errors actually *are* normally distributed, the quantity χ^2 is corresponding to a sum of N squares of normally distributed quantities, each normalized to unit variance. Once one has adjusted the $a_1 \dots a_M$ to minimize the value of χ^2 , the terms in the sum are not all statistically independent. For models that are linear in the a 's, however, it turns out that the probability distribution for different values of χ^2 at its minimum can nevertheless be derived analytically, and is the *chi-square distribution for $N - M$ degrees of freedom*.

In some cases the uncertainties associated with a set of measurements are not known in advance, and considerations related to χ^2 fitting are used to derive a value for σ . If one assumes that all measurements have the same standard deviation, $\sigma_i = \sigma$, and that the model fit well, then one can proceed by first assigning an arbitrary constant σ to all points, next fit for the model parameters by minimizing χ^2 and finally recomputing,

$$\sigma^2 = \sum_{i=1}^N [y_i - y(x_i)]^2 / N \quad (3.4).$$

If one takes the derivative of equation (3.3) with respect to the parameters a_k , one obtains equations that must hold at the chi-square minimum,

$$0 = \sum_{i=1}^N \frac{[y_i - y(x_i)]}{\sigma_i^2} \frac{(\partial y(x_i; \dots, a_k, \dots))}{\partial a_k} \quad (3.5)$$

Equation (3.5) is, in general, a set of M nonlinear equations for the M unknown a_k .

Multidimensional Fits

If one measures a single variable y as a function of more than one variable say, a vector of variables \mathbf{x} , then basis functions will be functions of a vector, $X_1(\mathbf{x}), \dots, X_M(\mathbf{x})$. The χ^2 merit function is now

$$\chi^2 = \sum_{i=1}^N \left[\frac{y_i - \sum_{k=1}^M a_k X_k(x_i)}{\sigma_i} \right]^2 \quad (3.6).$$

3.2 Nonlinear Models

One now considers fitting when the model depends *nonlinearly* on the set of M unknown parameters a_k , $k = 1, 2, \dots, M$. One uses the same approach as in previous section, namely to define a χ^2 merit function and determine best-fit parameters by its minimization. With nonlinear dependences, however, the minimization must proceed iteratively. Given trial values for the parameters, one

develops a procedure that improves the trial solution. The procedure is then repeated until χ^2 stops (or effectively stops) decreasing.

Sufficiently close to the minimum, one expects the χ^2 function to be well approximated by a quadratic form, which one can write as

$$\chi^2(\mathbf{a}) \approx \gamma - \mathbf{d} \cdot \mathbf{a} + \frac{1}{2} \mathbf{a} \cdot \mathbf{D} \cdot \mathbf{a} \quad (3.7)$$

where \mathbf{d} is an M – vector (Gradient vector) and \mathbf{D} is an $M \times M$ matrix (Hessian matrix).

If the approximation is good, one knows how to jump from the current trial parameters \mathbf{a}_{cur} to the minimizing ones \mathbf{a}_{min} in a single leap, namely

$$\mathbf{a}_{\text{min}} = \mathbf{a}_{\text{cur}} + \mathbf{D}^{-1} \cdot [-\nabla\chi^2(\mathbf{a}_{\text{cur}})] \quad (3.8).$$

On the other hand, (3.7) might be a poor local approximation to the shape of the function that we are trying to minimize at \mathbf{a}_{cur} . In that case, one can do is taken a step down the gradient, as in the steepest descent method. In other words,

$$\mathbf{a}_{\text{next}} = \mathbf{a}_{\text{cur}} - \text{constant} \times \nabla\chi^2(\mathbf{a}_{\text{cur}}) \quad (3.9)$$

where the constant is small enough and the downhill direction is not exhausted.

To use (3.8) or (3.9), one must be able to compute the gradient of the χ^2 function at any set of parameters \mathbf{a} . To use (3.8) one also needs the matrix \mathbf{D} , which is the second derivative matrix (Hessian matrix) of the χ^2 merit function, at any \mathbf{a} . One knows exactly the form of χ^2 since it is based on a model function that one has specified. Therefore the Hessian matrix is known to us. Thus one is free to use (3.8) whenever cares to do so.

Calculation of the Gradient and Hessian

The fitted model is

$$y = y(\mathbf{x}; \mathbf{a}) \quad (3.10)$$

and the χ^2 merit function is

$$\chi^2(\mathbf{a}) = \sum_{i=1}^N \left[\frac{y_i - y(\mathbf{x}_i; \mathbf{a})}{\sigma_i} \right]^2 \quad (3.11).$$

The gradient of χ^2 with respect to the parameters \mathbf{a} , which will be zero at the χ^2 minimum, has components

$$\frac{\partial \chi^2}{\partial a_k} = -2 \sum_{i=1}^N \frac{[y_i - y(\mathbf{x}_i; \mathbf{a})]}{\sigma_i^2} \frac{\partial y(\mathbf{x}_i; \mathbf{a})}{\partial a_k} \quad k=1,2,\dots,M \quad (3.12)$$

Taking an additional partial derivative gives

$$\frac{\partial^2 \chi^2}{\partial a_l \partial a_k} = 2 \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[\frac{\partial y(\mathbf{x}_i; \mathbf{a})}{\partial a_k} \frac{\partial y(\mathbf{x}_i; \mathbf{a})}{\partial a_l} - [y_i - y(\mathbf{x}_i; \mathbf{a})] \frac{\partial^2 y(\mathbf{x}_i; \mathbf{a})}{\partial a_l \partial a_k} \right] \quad (3.13)$$

It is conventional to remove the factors of 2 by defining

$$\beta_k \equiv -\frac{1}{2} \frac{\partial \chi^2}{\partial a_k} \quad \alpha_{kl} \equiv \frac{1}{2} \frac{\partial^2 \chi^2}{\partial a_l \partial a_k} \quad (3.14)$$

Making $[\alpha] = \frac{1}{2} \mathbf{D}$ in equation (3.8), in terms of which that equation can be rewritten as the set of linear equations

$$\sum_{l=1}^M \alpha_{kl} \delta_{al} = \beta_k \quad (3.15).$$

This set is solved for the increments δ_{al} that, added to the current approximation, give the next approximation. In the context of least-squares, the

matrix $[\alpha]$, equal to one-half times the Hessian matrix, is usually called the *curvature matrix*.

Equation (3.9), the steepest descent formula, translates to

$$\delta_{a_l} = \text{constant} \times \beta_l \quad (3.16).$$

Note that the components α_{kl} of the Hessian matrix (3.13) depend both on the first derivatives and on the second derivatives of the basis functions with respect to their parameters.

Second derivatives occur because the gradient (3.12) already has dependence on $\partial y / \partial a_k$, so the next derivative simply must contain terms involving $\partial^2 y / \partial a_l \partial a_k$. The second derivative term can be dismissed when it is zero, or small enough to be negligible when compared to the term involving the first derivative. It also has an additional possibility of being ignorable small in practice: The term multiplying the second derivative in equation (3.13) is $[y_i - y(x_i; \mathbf{a})]$. For a successful model, this term should just be the random measurement error of each point. This error can have either sign, and should in general be uncorrelated with the model. Therefore, the second derivative terms tend to cancel out when summed over i .

Inclusion of the second-derivative term can in fact be destabilizing if the model fits badly or is contaminated by outlier points that are unlikely to be offset by compensating points of opposite sign. From this point on, one will always use as the definition of α_{kl} as

$$\alpha_{kl} = \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[\frac{\partial y(x_i; \mathbf{a})}{\partial a_k} \frac{\partial y(x_i; \mathbf{a})}{\partial a_l} \right] \quad (3.17).$$

One should understand that minor (or even major) fiddling with $[\alpha]$ has no effect at all on what final set of parameters \mathbf{a} is reached, but affects only the iterative

route that is taken in getting there. The condition at the χ^2 minimum that $\beta_k = 0$ for all k is independent of how $[\alpha]$ is defined.

3.3 Levenberg-Marquardt Method

The Levenberg-Marquardt [24] is one of the most widely used non-linear curve fitting methods. It is very useful for finding solutions to complex fitting problems. Its utility is the way in which the fit method moves smoothly between the two extremes of the steepest descent method and the inverse-Hessian method [25-31], for finding the next step. The algorithm uses the method of steepest descent to determine the direction and then step size when the results are far from the minimum. But as the solution approaches the minimum, the algorithm switches to the Hessian matrix for determining the step in order to zero in on the best fit.

The method is based on two elementary, but important, insights. Consider the “constant” in equation (3.16). There is no information about the answer in the gradient that tells only the slope, not how far that slope extends. Marquardt’s first insight is that the components of the Hessian matrix give *some* information about the order-of-magnitude scale of the problem.

The quantity χ^2 is nondimensional, i.e., is a pure number; this is evident from its definition (3.11). On the other hand, β_k has the dimensions of $1/a_k$, which may well be dimensional. The constant of proportionality between β_k and δa_k must therefore have the dimensions of a_k^2 . Scan the components of $[\alpha]$ and one sees that there is only one obvious quantity with these dimensions, and that is $1/\alpha_{kk}$, the reciprocal of the diagonal element. So that must set the scale of the constant. But that scale might itself be too big. So let’s divide the constant by some (nondimensional) fudge factor λ , with the possibility of setting $\lambda \gg 1$ to cut down the step. In other words, replace equation (3.16) by

$$\delta \mathbf{a}_l = \frac{1}{\lambda \alpha_{ll}} \beta_l \quad \text{or} \quad \lambda \alpha_{ll} \delta \mathbf{a}_l = \beta_l \quad (3.18).$$

It is necessary that α_{ll} be positive, but this is guaranteed by definition (3.17) another reason for adopting that equation.

Marquardt's second insight is that equations (3.18) and (3.15) can be combined if we define a new matrix α' by the following prescription

$$\begin{aligned} \alpha'_{jj} &\equiv \alpha_{jj}(1+\lambda) \\ \alpha'_{jk} &\equiv \alpha_{jk} \quad (j \neq k) \end{aligned} \quad (3.19)$$

and then replace both (3.18) and (3.15) by

$$\sum_{l=1}^M \alpha'_{kl} \delta \mathbf{a}_l = \beta_k \quad (3.20)$$

When λ is very large, the matrix α' is forced into being *diagonally dominant*, so equation (3.20) goes over to be identical to (3.18). On the other hand, as λ approaches zero, equation (3.20) goes over to (3.15).

Given an initial guess for the set of fitted parameters \mathbf{a} , the recommended Marquardt recipe is as follows:

- Compute $\chi^2(\mathbf{a})$,
- Pick a modest value for λ , say $\lambda = 0.001$,
- (*) Solve the linear equations (3.20) for $\delta \mathbf{a}$ and evaluate $\chi^2(\mathbf{a} + \delta \mathbf{a})$,
- If $\chi^2(\mathbf{a} + \delta \mathbf{a}) \geq \chi^2(\mathbf{a})$ then *increase* λ by a factor of 10 (or any other substantial factor) and go back to (*).

else *decrease* λ by a factor of 10, update the trial solution $\mathbf{a} \leftarrow \mathbf{a} + \delta \mathbf{a}$,

and go back to (*).

Also necessary is a condition for stopping. Iterating to convergence (to machine accuracy or to the round off limit) is generally wasteful and unnecessary since the minimum is best only a statistical estimate of the parameters \mathbf{a} . A change in the parameters that changes χ^2 by an amount $\ll 1$ is *never* statistically meaningful.

Outright failure by a zero pivot is possible, but unlikely. More often, a small pivot will generate a large correction, which is then rejected, the value of λ being then increased. For sufficiently large λ the matrix $[\alpha']$ can have no small pivots. Thus the method does tend to stay away from zero pivots.

These considerations suggest that, in practice, one might as well stop iterating on the first or second occasion that χ^2 decreases by a negligible amount, say either less than 0.01 absolutely or (in case round off prevents that being reached) some fractional amount like 10^{-3} . Don't stop after a step where χ^2 *increases*: That only shows that λ has not yet adjusted itself optimally.

Once the acceptable minimum has been found, one wants to set $\lambda = 0$ and compute the matrix

$$[\mathbf{C}] \equiv [\alpha]^{-1} \quad (3.21)$$

which is the estimated covariance matrix of the standard errors in the fitted parameters \mathbf{a} .

One hopes that this algorithm will converge to the desired solution, but it could also end up in a local minimum or diverge. This depends on the initial value. Since the Levenberg-Marquardt method is an iterative process that requires the initial input, any interpretations from the final results are only as good as the initial

input. Therefore the use of incorrect function forms can fit the data fine, but using the “fit parameters” for further interpretation would be erroneous.



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

CHAPTER 4

DETAIL OF CALCULATIONS

4.1 Development of Intermolecular Potential Function

As mentioned before that the following steps are required to develop the potential function by means of quantum chemical calculations:

- (i) selection of representative geometries of the pairs,
- (ii) the *ab initio* calculations,
- (iii) fitting of the computed interaction energies to the functional form,
- and (iv) improving quality of the function which consists of 2 steps:
 - testing the quality of the function
 - search for false minima of the function.

This procedure is schematically illustrated in Figure 4.1.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

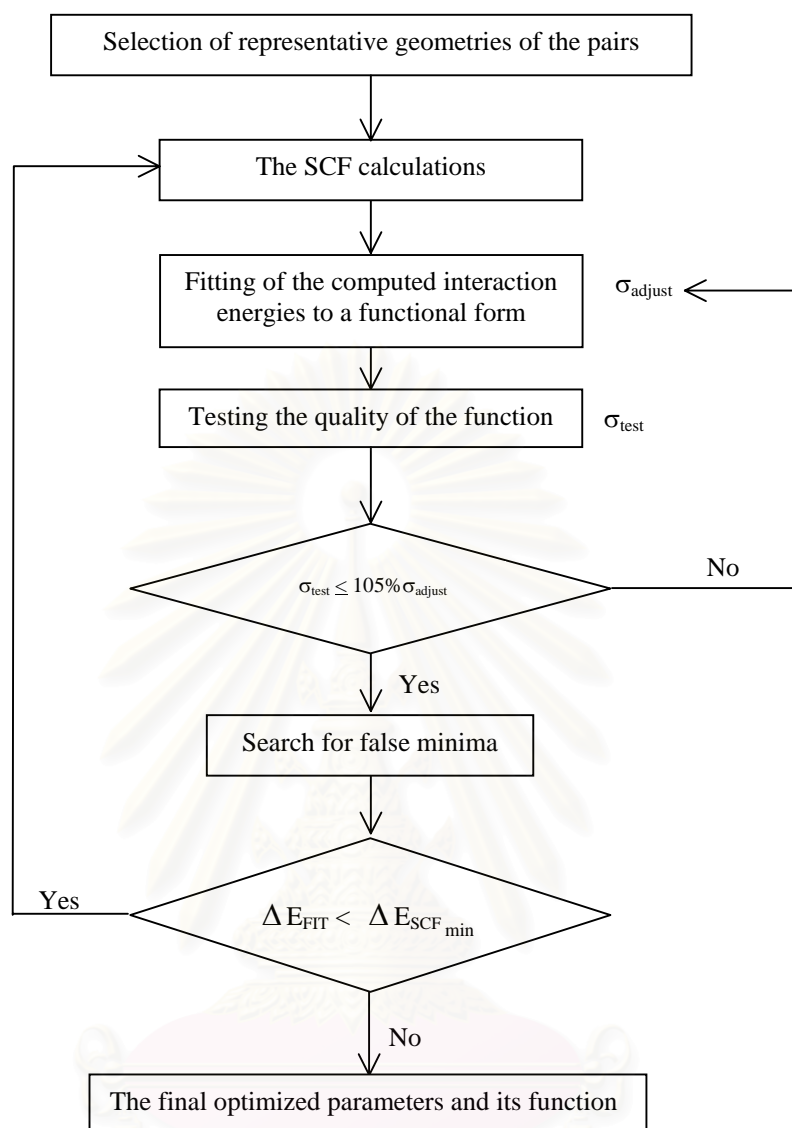


Figure 4.1 The procedure for constructing the potential function by means of quantum chemical calculations.

4.1.1 Selection of Representative Geometries of the Pairs

The pair geometries were selected with regard to chemically representative conformations. The first molecule was fixed at the origin of the cartesian coordinate system and the second molecule was placed at numerous positions around the first molecule. The distance between atoms of two molecules should not be too close to

each other because this may lead to strong repulsive interaction. The symmetry of the molecules can reduce the repetition and save the calculated time of the interaction energies in the next step. Both molecules were treated as rigid throughout the calculations.

4.1.2 The SCF Calculations

In the SCF calculations, the interaction energy ΔE_{SCF} between two molecules is determined as the difference of the supersystem energy (E_{12}) and sum of the subsystem energies (E_1, E_2)

$$\Delta E_{\text{SCF}}(\text{kcal / mol}) = [E_{12} - (E_1 + E_2)] \times 627.5 \quad (4.1)$$

where E_{12} , E_1 and E_2 are total energy (in atomic units) obtained from quantum chemical calculations.

In this study, part of the semi-automatic Fortran program has been developed to construct the intermolecular pair potential function in the steps (iii) and (iv), fitting of pair interaction energies to a functional form and improving quality of the function. Details of the program are given in the next section.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

4.2 Program Implementation

The program, “**MRQFIT**” version 1.0, has been developed and its source code has been written in Fortran language.

The program manual is given in Appendix I.

Program Strategies

For easy understanding, the program can be separated into seven parts:

1. Input Elementary Data,
2. Selection of the Functional Form,
3. Setting an Initial Guess of the Fitting Parameters,
4. Fitting Energies to the Function,
5. Testing the Quality of the Function,
6. Search for False Minima,
7. Write Output.

The flowchart of **MRQFIT** program is shown in Figure 4.2.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

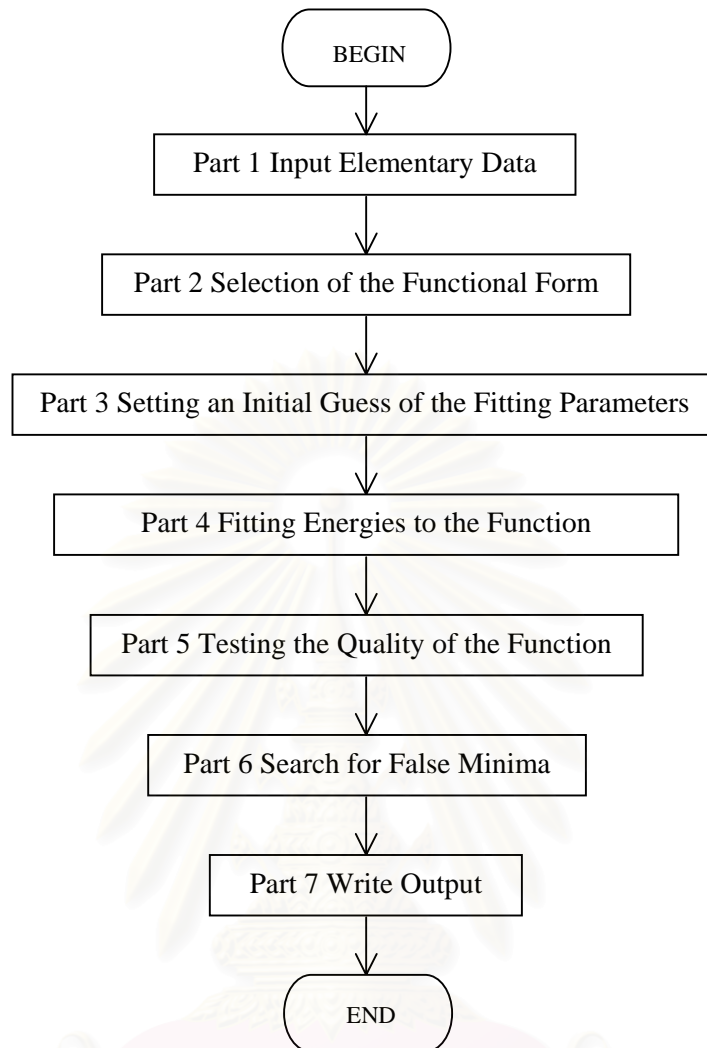


Figure 4.2 The sequence of MRQFIT program version 1.0.

4.2.1 PART 1: Input Elementary Data

This part consists of two steps. The first step uses subroutine **Indata** to obtain input elementary data from user. The input elementary data is composed of four sections.

The structure of input is shown in Figure 4.3.

fname natom1 natom2 maxiter nset	}	Section 1
charge1 charge2	}	Section 2
pair12 sset	}	Section 3
comol1	}	Section 4
thetha phi ener comol2	}	Section 5

Figure 4.3 The structure of input elementary data.

The variables of input elementary data are defined as follows:

Section 1:

fname: Data filename, *i.e.*, imrqfit

natom1: Number of atoms of molecule 1

natom2: Number of atoms of molecule 2

maxiter: Maximum number of iterations

nset: Number of types of subfunctions

Section 2:

charge1: Atomic net charge of each atom of molecule 1

charge2: Atomic net charge of each atom of molecule 2

Section 3:

pair12, sset: A positive interger which represents atom i and atom j of two molecules and specify types of subfunctions

Section 4:

comol1: Cartesian coordinates of molecule 1

Section 5:

thetha, phi: Angles of each trajectories (see Figure 4.11, 4.12)

ener, comol2: Interaction energies and corresponding cartesian coordinates of molecule 2

The second step, section 5 of input elementary data, is separated into two files using subroutine **Isplit**. The first file is for fitting and the second one is for testing the function.

The criteria to separate file which implements in this program is to keep approximately one-fifth of the data for testing. Therefore, the trajectories are selected by choosing one and skiping four trajectories. The selected trajectories (tested set) which should be between 10 to 20 percent of all the data and corresponding cartesian coordinates are used for testing and the rests are used for fitting the function. The procedure of this part is shown in Figure 4.4.

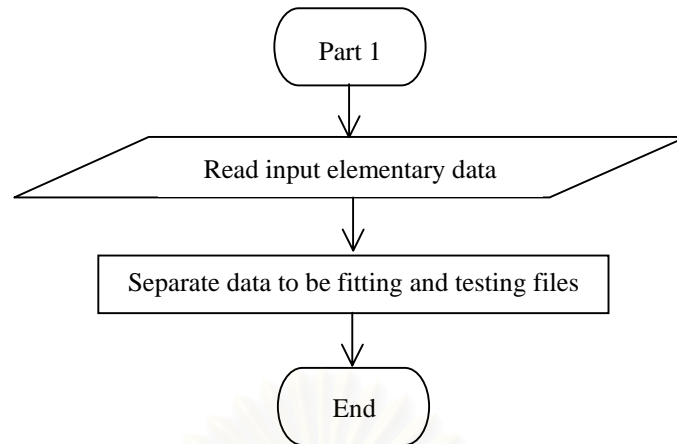


Figure 4.4 The flowchart of part 1; Input Elementary Data.

4.2.2 PART 2: Selection of the Functional Form

This part is designed for the user to select the function form suitable for the studying system. Power of denominator of each term can be also assigned. The functional form which available in the program are:

$$1. \Delta E_{MN} = \sum_i^m \sum_j^n \left(-\frac{A_{ij}}{r_{ij}^x} + \frac{B_{ij}}{r_{ij}^y} + \frac{q_i q_j}{r_{ij}} \right) \quad (4.2)$$

$$2. \Delta E_{MN} = \sum_i^m \sum_j^n \left(-\frac{A_{ij}}{r_{ij}^x} + \frac{B_{ij}}{r_{ij}^y} + \frac{q_i q_j}{r_{ij}} + \frac{C_{ij}}{r_{ij}^z} \right) \quad (4.3)$$

$$3. \Delta E_{MN} = \sum_i^m \sum_j^n \left(-\frac{A_{ij}}{r_{ij}^x} + B_{ij} \exp(-C_{ij} r_{ij}) + \frac{q_i q_j}{r_{ij}} \right) \quad (4.4)$$

$$4. \Delta E_{MN} = \sum_i^m \sum_j^n \left(-\frac{A_{ij}}{r_{ij}^x} + B_{ij} \exp(-C_{ij} r_{ij}) + \frac{q_i q_j}{r_{ij}} + \frac{D_{ij}}{r_{ij}^y} \right) \quad (4.5)$$

$$5. \Delta E_{MN} = \sum_i^m \sum_j^n \left(-\frac{A_{ij}}{r_{ij}^x} + B_{ij} \exp(-C_{ij} r_{ij}) + \frac{q_i q_j}{r_{ij}} + D_{ij} \exp(-E_{ij} r_{ij}) \right) \quad (4.6)$$

$$6. \Delta E_{MN} = \sum_i^m \sum_j^n \left(-A_{ij} \exp(-B_{ij} r_{ij}) + C_{ij} \exp(-D_{ij} r_{ij}) + \frac{q_i q_j}{r_{ij}} \right) \quad (4.7)$$

This procedure executes by subroutine **Selectf**. It starts with reading the selected choice (1-6), then, sets the power of denominator of each term to yield the functional form.

The procedure of this part is shown in Figure 4.5.

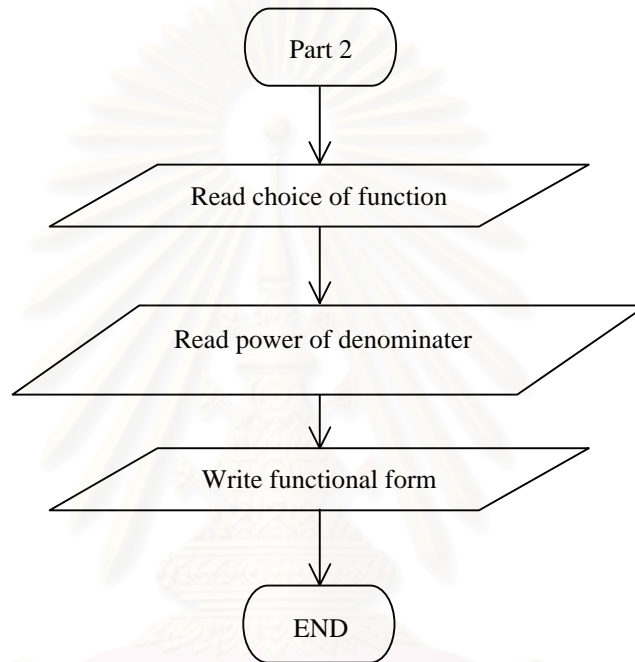


Figure 4.5 The flowchart of part 2; Selection of the Functional Form.

4.2.3 PART 3: Set an Initial Guess for the Parameters

This part is designed to set an initial guess for the parameters. The procedure is given in subroutines **Ranpar** and **Ran**.

This part begins with reading the functional form (the output of subroutine **Selectf**), then set the initial parameters which are corresponding to each function.

Initial guess for the parameters are generated from the power of denominator of each parameters using the following expression,

$$a(i) = (10.0 + (40.0 * \mathbf{Ran}(\text{iseed}))) * (1.5)^R \quad (4.8)$$

where $a(i)$ is i th initial parameter, R is its power of denominator. The subroutine **Ran** assigns real number between 0.0 and 1.0, and initial `iseed` is 1.

If the parameter is the power of exponential term, its initial value was assigned as a real number in an interval between 0.0 and 1.0,

$$a(i) = \mathbf{Ran}(\text{iseed}) \quad (4.9).$$

The main procedure of this part is shown in Figure 4.6.

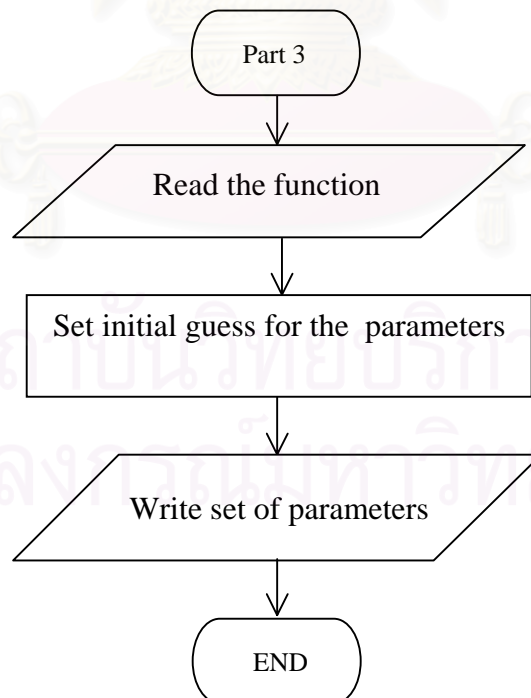


Figure 4.6 The flowchart of part 3; Set an initial Guess for the Parameters.

4.2.4 PART 4: Fitting Energies to the Function

This part is designed to fit the function by using a multidimensional non-linear least-squares procedure, namely the Marquardt-Levenberg algorithm (See Chapter 3). This will minimize the χ^2 merit function until constancy of the fitting parameters is reached.

In many applications, the data points will not be equally important, because some data points are known to be more important than others. The more important the data point, the larger the weight. To improve the fit, weighting factor is included in the fitting process. If the error in the j th data point is approximately e_j then choose $w_j = 1/e_j$, where w_j are the weights. Thus the smaller the error, the larger the weight. Weighting can improve the statistical properties of the solution. Of course, the parameters obtained from the model will change as the weighting factor changes.

In this study, weighting factor has been given to the most important configuration (which corresponds to the lowest SCF energy points). In addition, highly repulsive configurations are also excluded from the process.

The χ^2 merit function that will be used to minimize to get *goodness-of-fit* in the program is:

$$\chi^2(j) = \sum_{j=1}^{nconf} [(\Delta E_{SCF(j)} - \Delta E_{FIT(j)})^2 * w(j)] / df \quad (4.10)$$

where $w(j)$, $\Delta E_{SCF(j)}$ and $\Delta E_{FIT(j)}$ are weight value, SCF energy and the energy which calculate from the function of j th data point, respectively, $nconf$ is number of configurations and df is number of degree of freedom which is expressed as:

$$df = nconf - (npar + ncut) \quad (4.11)$$

where $npar$ is number of parameter which will be used to adjust in fitting procedure and $ncut$ is number of SCF energy data points which are excluded from the process.

Input of this part consists of:

1. An initial guess for the set of fitting parameters (the output from subroutine **Ranpar**),
2. Functional form (the output from subroutine **Selectf**).

The following is the steps inside this part:

1. Read an initial guess of fitting parameters (\mathbf{a}),
2. Compute $\chi^2(\mathbf{a})$,
3. Set $\lambda = 0.001$,
4. Solve $\delta\mathbf{a}$ and evaluate $\chi^2(\mathbf{a} + \delta\mathbf{a})$ by following conditions:
 - If $\chi^2(\mathbf{a} + \delta\mathbf{a}) \geq \chi^2(\mathbf{a})$ then set $\lambda = \lambda * 10$ and go back to (4).
 - else update the trial solution $\mathbf{a} \leftarrow \mathbf{a} + \delta\mathbf{a}$ and go back to (4).

The step 4 will be repeated until consistency of the fitting parameters is reached.

The convergence criteria which implements in this program is to use the following equation:

$$\|\mathbf{a}_{\text{new}} - \mathbf{a}_{\text{old}}\| \leq \|0.01 * \mathbf{a}_{\text{old}}\| \text{ and } \chi^2(\mathbf{a}) \leq 0.5 \quad (4.12)$$

where \mathbf{a}_{new} and \mathbf{a}_{old} are the current and previous parameters, respectively.

Note that warning message will be given in the output file if the convergence is not reached.

Thus, from the above procedure, the output consists of:

1. Set of parameters,
2. Standard deviation yielded from adjusting the function (σ_{adjust}),
3. Number of iterations.

Beside subroutines **Ranpar** and **Selectf**, this part also uses:

- Subroutine **Gradhs** to calculate matrix $[\alpha]$ (equation 3.17) and vector β (equation 3.12 and 3.14).
- Subroutine **Dev** to calculate vector of derivatives (term $\frac{\partial y(x_i; a)}{\partial a_k}$ in equation 3.12).

Note that the program uses two different methods to compute derivative. They are the analytical derivative computation and the approximate derivative computation using central difference technique [32-33].

To calculate analytically, the partial derivative with respect to the parameters of two general forms of the potential functions (see equations 4.2 and 4.4) are expressed as:

$$\frac{\partial \Delta E_{MN}}{\partial A_{ij}} = \sum_i^m \sum_j^n \left(-\frac{1}{r_{ij}^x} \right) \quad (4.13)$$

$$\frac{\partial \Delta E_{MN}}{\partial B_{ij}} = \sum_i^m \sum_j^n \left(\frac{1}{r_{ij}^y} \right) \quad (4.14)$$

and

$$\frac{\partial \Delta E_{MN}}{\partial A_{ij}} = \sum_i^m \sum_j^n \left(-\frac{1}{r_{ij}^x} \right) \quad (4.15)$$

$$\frac{\partial \Delta E_{MN}}{\partial B_{ij}} = \sum_i^m \sum_j^n \left(\exp(-C_{ij} r_{ij}) \right) \quad (4.16)$$

$$\frac{\partial \Delta E_{MN}}{\partial C_{ij}} = \sum_i^m \sum_j^n \left((B_{ij} * (-r_{ij})) * \left(\exp(-C_{ij} r_{ij}) \right) \right) \quad (4.17).$$

Considering the partial derivative with respect to A_{ij} and B_{ij} of the first form and A_{ij} of second form, only r_{ij} are involved. Because we fixed the distance in each calculation, therefore the derivative yields the same value. Considering the first form, this means that ΔE_{MN} is a linear function in term of A_{ij} and B_{ij} .

- Subroutine **Guassj**, which performs with Gauss-Jordan method [32-33], is used to calculate the matrix solution ($\delta \mathbf{a}$ in equation 3.20).

- Subroutine **Statva** to calculate the statistical value.

- Subroutine **Funct** to return the model function.

The procedure of this part is shown in Figure 4.7.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

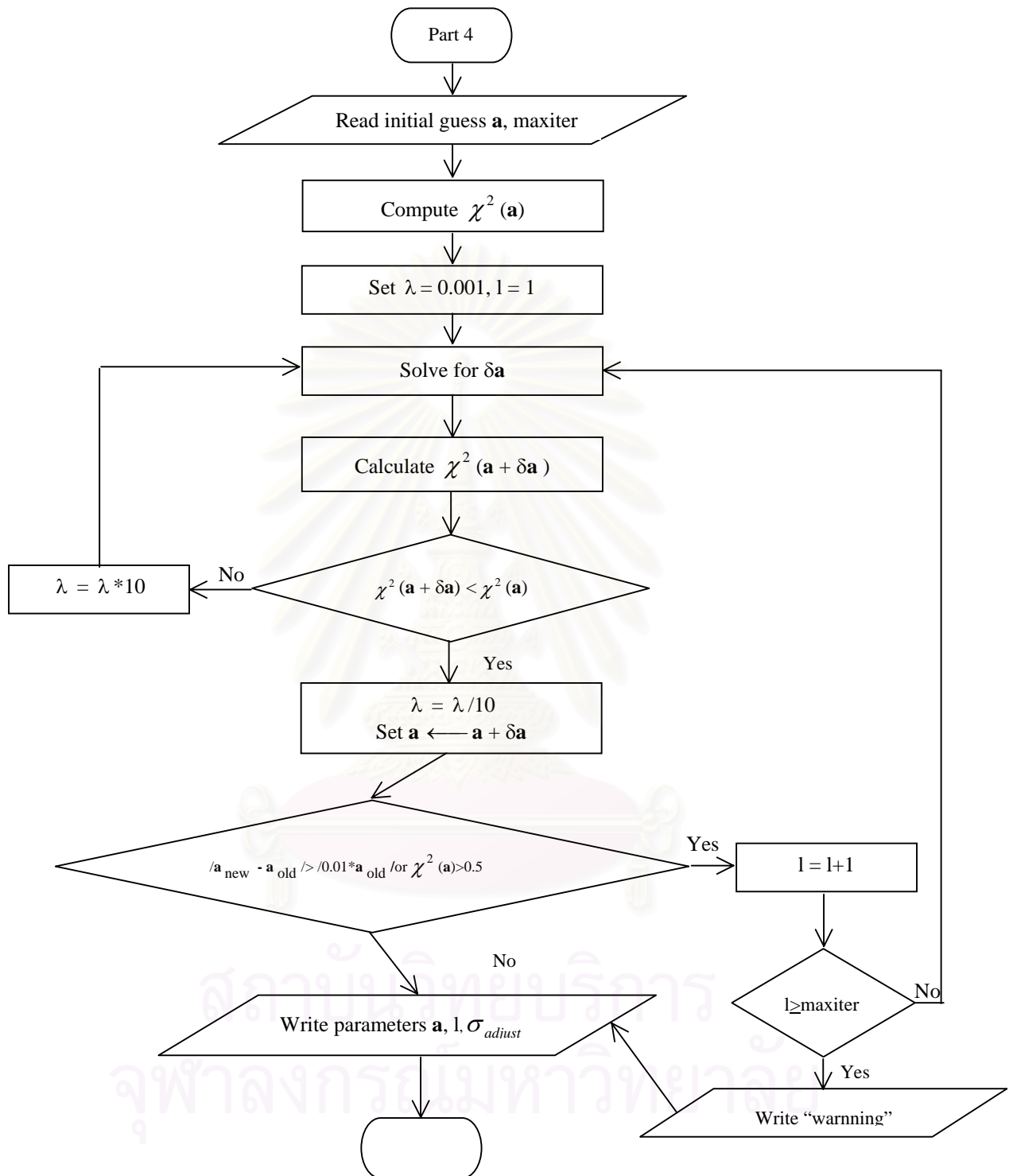


Figure 4.7 The flowchart of part 4; Fitting Energies to the Function.

4.2.5 PART 5: Testing the Quality of the Function

This part is designed to test the obtained function. The predictive capabilities of the potential functions are tested according to the procedure suggested by Beveridge et al. [6], i.e., namely, the SCF energies (outside the original set) are calculated and compared with the values predicted by the function. These points are then included in the fitting procedure and the whole process was repeated until constancy of the fitting parameters, fluctuated within a range of $\pm 5\%$, is obtained and a sufficiently low standard deviation is reached.

This part is examined using subroutine **Testf**.

Input of this procedure consists of :

1. The pair potential function (output from part 4),
2. Standard deviation of the fit (σ_{adjust} , also output from part 4),
3. The configurations (Cartesian coordinate of molecule 2 and corresponding interaction energy which yields from quantum chemical calculations ΔE_{SCF}), which have never been used in the fit (tested set).

The following is the steps inside this procedure:

1. Calculate the energy points using the pair potential function (ΔE_{FIT}),
2. Compare all ΔE_{SCF} and ΔE_{FIT} values and their standard deviation from each other (σ_{test}) by following conditions:
 - If $\sigma_{test} > 105\%$ of σ_{adjust} then go back to step of fitting function and included these points in the fitting procedure and the whole processes are repeated.
 - else go to the next step to search for false minima of the function.

Thus, from the above procedure, the output consists of:

1. Standard deviation for the tested data set (σ_{test}),
2. ΔE_{SCF} and ΔE_{FIT} of the tested configurations.

The procedure of this part is shown in Figure 4.8.



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

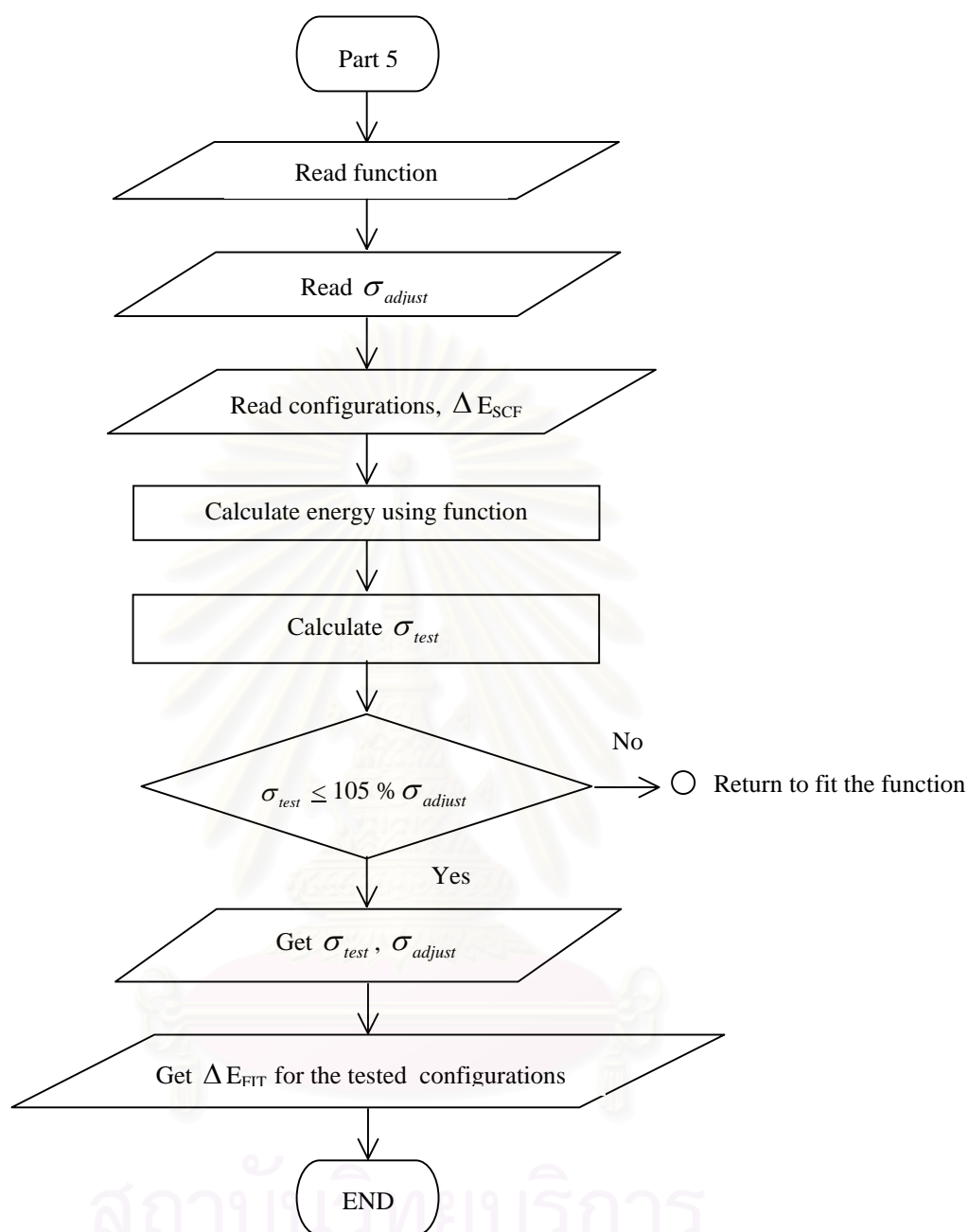


Figure 4.8 The flowchart of part 5; Testing the Quality of the Function.

4.2.6 PART 6: Search for False Minima of the Function

In this part, false minima that may be available in the function will be searched. This checking procedure is also essential as unwanted minima may take place in the function.

The procedure for this part is in subroutines **Sfmin** and **Genco**.

Input of this procedure consists of :

1. The configurations, which different from the previous (the output from subroutine **Genco**),
2. The pair potential function (the output from part 4).

The following steps are in this procedure:

1. Calculate the energy points for as many configurations of the dimer as possible using the pair potential function (ΔE_{FIT}),
2. Search for the false minima of the function by following conditions:
 - If the function contains artificial minima then go back to step of calculate SCF energy. Then these points are included in the fitting procedure and the whole processes are repeated.else write the pair potential function to the output file.

Note that artificial minima are the energy points which are deeper or higher than those yielded from quantum chemical calculations at the nearest configuration.

Thus, from above procedure, the output consists of the final fitting parameters and its function.

The procedure of this part is shown in Figure 4.9.

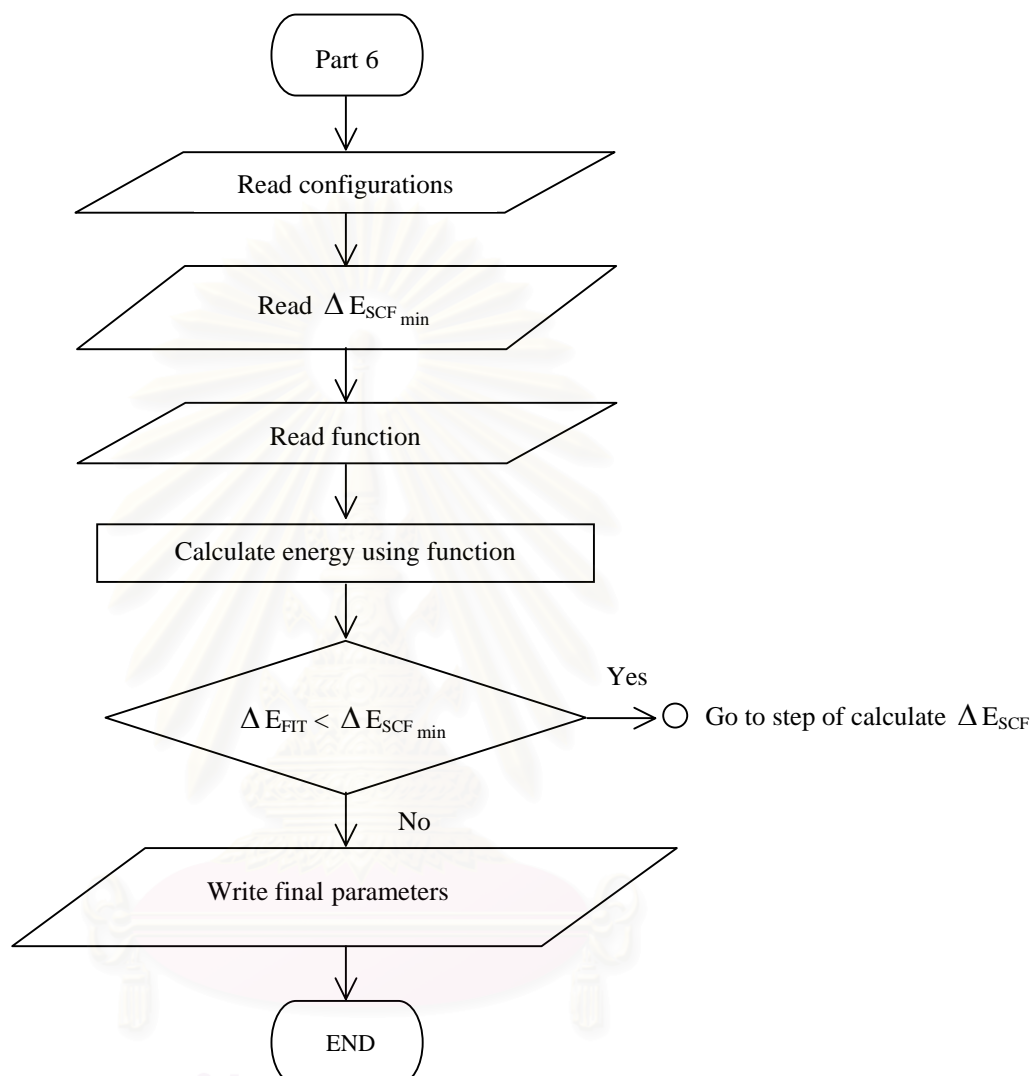


Figure 4.9 The flowchart of part 6; Search for False Minima of the Function.

4.2.7 PART 7: Write Output

This part is designed to write output of the calculations. The output consists of three files. The first file contains statistical values and adjusted parameters every iterations, the second one collects final values of the fitting parameters. The third file

keeps interaction energy of all configurations both those yield from quantum chemical calculations (ΔE_{SCF}) and the pair potential function (ΔE_{FIT}).

The example of output files are given in Appendix II.

The procedure of this part is shown in Figure 4.10.

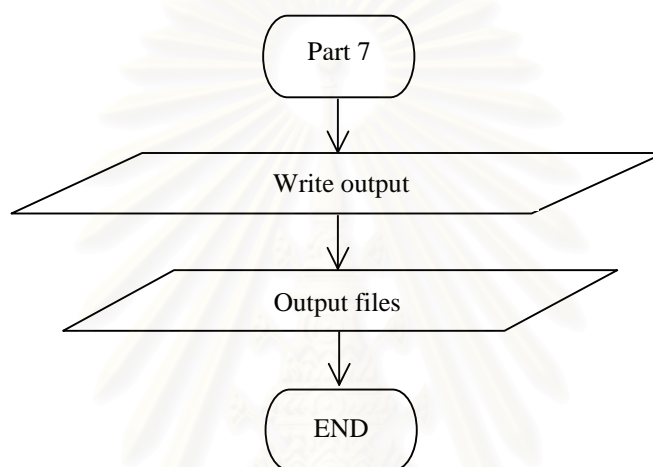


Figure 4.10 The flowchart of part 7; Write Output.

4.3 Specification of the systems for this study

In this study, the $\text{Li}^+\text{-NH}_3$ and $\text{NH}_3\text{-NH}_3$ systems are used to develop and test the performance of the program.

In the $\text{Li}^+\text{-NH}_3$ system, the nitrogen atom of the ammonia molecule was placed at the origin of cartesian coordinate and the lithium ion was placed at numerous positions within the space around ammonia, where $0^\circ \leq \theta \leq 180^\circ$ and $0^\circ \leq \phi \leq 120^\circ$ in steps of 30° and $1.20 \leq r \leq 10.00 \text{ \AA}$, when r is distance between the nitrogen atom of ammonia molecule and the lithium ion. With this consideration 735 configurations of the lithium ion have been generated. The corresponding interaction energies were calculated using DZP basis set (Double Zeta plus Polarization

Function) with BSSE corrections. Both molecules were treated as rigid throughout the calculations.

The graphic drawing represented the definition of geometries variable for configurations of ammonia-lithium ion was shown in Figure 4.11.

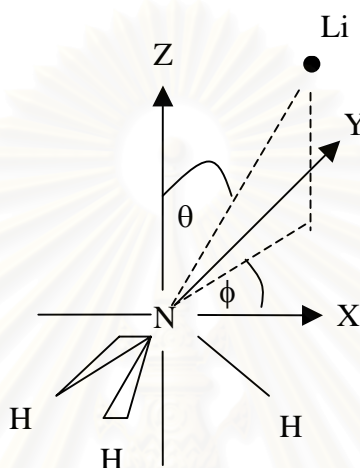


Figure 4.11 Definition of geometries variable (θ, ϕ) for the configuration of ammonia-lithium ion.

The atomic coordinates of ammonia molecule and the atomic net charges obtained from the Mulliken population analysis which are used in the Coulombic term in the pair potential were collected in Table 4.1. Due to the symmetry reason, all three hydrogen atoms of NH_3 are treated equivalently. Therefore, only two types of subfunctions are available in the Li^+-NH_3 system, i.e., those for Li^+-N and Li^+-H pairs.

Table 4.1 Cartesian coordinates (in Angstroms) and atomic net charges (in a.u.) of ammonia molecule obtained from *ab initio* calculations using DZP basis set.

Atom	X	Y	Z	Atomic Charge
N	0.000000	0.000000	0.000000	-0.74207
H	0.937760	0.000000	-0.381470	0.24736
H	-0.468880	-0.812120	-0.381470	0.24736
H	-0.468880	0.812120	-0.381470	0.24736

For the NH₃-NH₃ system, the nitrogen atom of the first molecule was placed at the origin of cartesian coordinate and the second molecule was placed at numerous positions within the space around the first molecule, where $0^\circ \leq \theta \leq 180^\circ$ and $0^\circ \leq \phi \leq 60^\circ$ in steps of 30° and $2.30 \leq r \leq 10.00 \text{ \AA}$, when r is distance between the nitrogen atoms of two molecules. Here, orientation of the second ammonia around its nitrogen atom using the same conditions for r , θ and ϕ , has been also taken into consideration. With this consideration 819 configurations of the second molecule have been generated. The corresponding interaction energies were calculated using DZP basis set (Double Zeta plus Polarization Function) with BSSE corrections. Both molecules were treated as rigid throughout the calculations.

The graphic drawing represented the definition of geometries variable for the configurations of ammonia-ammonia was shown in Figure 4.12.

● = Nitrogen atom of the second molecule

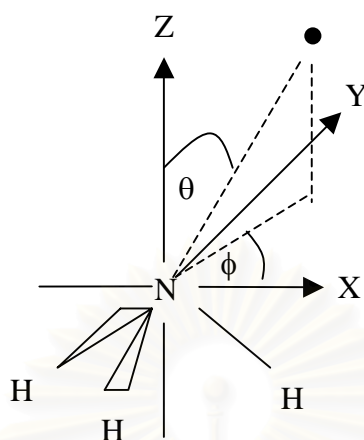


Figure 4.12 Definition of geometries variable (θ, ϕ) for the ammonia- ammonia configuration.

Due to the symmetry reason, all three hydrogen atoms of NH_3 are treated equivalently. Therefore, only three types of subfunctions are available in the NH_3 - NH_3 system, i.e., N-N, N-H, H-H pairs.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

CHAPTER 5

RESULTS AND DISCUSSION

In this Chapter, characteristics of the newly developed Fortran program, case studied for the lithium ion-ammonia and ammonia- ammonia potentials have been reported. Discussed has been made in terms of optimized parameters, the statistical values, number of iteration cycle, etc.

5.1 Development of Lithium Ion - Ammonia Potential Function

5.1.1 Input Elementary Data

In this part, after user inputs elementary data, the program will count and report the number of all trajectories, SCF data points and the corresponding cartesian coordinates which will be included in the development of the function. In this example, the fitting procedure contains 35 trajectories and 735 SCF energy points in the geometries mentioned in Figure 4.11. Then, one-fifth of the trajectory, taking one every five lines, will be kept for testing.

The selected 7 trajectories as well as the corresponding 147 SCF energies are shown in Table 5.1. Here, the other 28 trajectories with 588 SCF energy points were used for fitting.

Table 5.1 The selected trajectories, in term of θ and ϕ angles (see Figure 4.11), for lithium ion around ammonia.

Trajectory	θ	ϕ
1	0	0
2	30	0
3	60	0
4	90	0
5	120	0
6	150	0
7	180	0

An input file for the $\text{Li}^+\text{-NH}_3$ system is shown in Figure 5.1.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

```

imrqfit
  4      1
500     2

-0.74207  0.24736  0.24736  0.24736
1.

  1,1,      1
  2,1,      2
  3,1,      2
  4,1,      2

  0.000000  0.000000  0.000000
  0.937760  0.000000 -0.381470
-0.468880 -0.812120 -0.381470
-0.468880  0.812120 -0.381470

theta =    0  Phi =    0
:
 55.044      0.000000  0.000000  1.300000
 -6.330      0.000000  0.000000  1.500000
-21.772      0.000000  0.000000  1.600000
-31.266      0.000000  0.000000  1.700000
:
theta =    90  Phi =    90
:
 -0.829      0.000000  2.549999  0.000000
 -0.978      0.000000  2.599999  0.000000
 -1.067      0.000000  2.649999  0.000000
 -1.108      0.000000  2.699999  0.000000
:
theta =   180  Phi =   120
:
 -1.169      0.000000  0.000000 -2.400000
 -1.267      0.000000  0.000000 -2.450000
 -1.283      0.000000  0.000000 -2.500000
 -1.233      0.000000  0.000000 -2.549999
:

```

Figure 5.1 An input file for the $\text{Li}^+\text{-NH}_3$ system (see Figure 4.3).

In this program, the input elementary data using free format.

5.1.2 Selection of the Functional Form

Here, the energy points for the lithium ion-ammonia dimer, with choice 3 of the functional form, $x=6$ (see equation 4.4), were displayed, that is

$$\Delta E_{\text{NH}_3-\text{Li}} = \sum_i^k \sum_j^l \left(-\frac{A_{ij}}{r_{ij}^x} + B_{ij} \exp(-C_{ij}r_{ij}) + \frac{q_i q_j}{r_{ij}} \right) \quad (5.1)$$

where $k = 4$ (number of atoms of ammonia molecule) and $l = 1$ (number of atom of Li^+), A_{ij} , B_{ij} and C_{ij} are fitting parameters, r_{ij} is the distance between the i th atom of ammonia and lithium ion, $q_i = -0.74207$ and 0.24736 atomic unit for atoms N and H of ammonia and $q_j = +1$ atomic unit for Li^+ .

5.1.3 Set an Initial Guess for the Parameters

In this part, the program random set of an initial guess of the fitting parameters. In this example, equation (5.1) contains 6 parameters namely A_{NLi} , B_{NLi} , C_{NLi} , A_{HLi} , B_{HLi} , C_{HLi} . The first 3 parameters represents interaction between atom N and Li^+ while the last 3 parameters for the H and Li^+ pair.

Note that due to symmetry reason, all 3 hydrogen atoms of NH_3 are represented by the same set of fitting parameters.

- The initial guess of parameters

$$A_{\text{NLi}} = 224.434555$$

$$B_{\text{NLi}} = 23.7004414$$

$$C_{\text{NLi}} = 0.13874417$$

$$A_{\text{HLi}} = 402.141846$$

$$B_{\text{HLi}} = 37.5042381$$

$$C_{\text{HLi}} = 0.91231015$$

5.1.4 Fitting Energies to the Functional Form

With 588 ammonia-lithium interaction energies (see Section 5.1.1), the pair lower than 10 kcal.mol^{-1} were fitted to the analytical function, equation 5.1. Note that repulsive interaction higher than 10 kcal.mol^{-1} are excluded from the fit because it is assumed that those configurations should be rarely detected in the simulation.

In this program, the vector of derivatives (term $\frac{\partial y(x_i; a)}{\partial a_k}$ in equation 3.12) has been calculated by the two methods. The first method, the derivatives were approximated using the central different technique which the second one, the derivatives were calculated analytically.

The optimized parameters of fit in which using the first method are given in the Table 5.2.

Table 5.2. The optimized parameters (the derivatives were approximated using the central different technique) representing interaction of N and H atoms of ammonia with lithium ion (energy and r are in kcal.mol^{-1} and angstroms, respectively).

Atom	A ($\text{\AA}^6 \text{ kcal.mol}^{-1}$)	B (kcal.mol^{-1})	C (kcal.mol^{-1})
NL _i	-0.10083546E+04	-0.36473591E+03	0.11088778E+01
HL _i	-0.15106658E+04	-0.75820667E+05	0.43180332E+01

Characteristics of the fit in which the derivatives were approximated using the central different technique:

Minimum energy included in the fit	=	-40.83	kcal.mol ⁻¹ .
Maximum energy included in the fit	=	10.00	kcal.mol ⁻¹ .
Maximum absolute residual	=	4.54	kcal.mol ⁻¹ .
Minimum absolute residual	=	0.001	kcal.mol ⁻¹ .
Average absolute residual	=	0.61	kcal.mol ⁻¹ .
Number of iteration	=	47	cycle.
Standard deviation (σ_{fit1})	=	0.196	kcal.mol ⁻¹ .

The optimized parameters of fit in which using the second method are given in the Table 5.3.

Table 5.3. The optimized parameters (the derivatives were calculated analytically) representing interaction of N and H atoms of ammonia with lithium ion (energy and r are in kcal.mol⁻¹ and angstroms, respectively).

Atom	A ($\text{\AA}^6 \text{kcal.mol}^{-1}$)	B (kcal.mol ⁻¹)	C (kcal.mol ⁻¹)
NL _i	-0.80015872E+03	-0.21806948E+03	0.10526058E+01
HL _i	0.58478275E+03	0.5649.9844E+04	0.2.9654104E+01

Characteristics of the fit in which the derivatives were calculated analytically:

Minimum energy included in the fit	=	-40.83	kcal.mol ⁻¹ .
Maximum energy included in the fit	=	10.00	kcal.mol ⁻¹ .
Maximum absolute residual	=	16.91	kcal.mol ⁻¹ .
Minimum absolute residual	=	0.005	kcal.mol ⁻¹ .
Average absolute residual	=	1.19	kcal.mol ⁻¹ .

Number of iteration	=	53	cycle.
Standard deviation (σ_{fit1})	=	0.41	kcal.mol ⁻¹ .

5.1.5 Testing the Quality of the Function

The analytical potential function obtained from 588 SCF data points was carefully tested in order to make sure that it is possible to represent interactions of all lithium ion positions and orientations relative to ammonia molecule.

In this part, the 147 ammonia-lithium configurations (with the corresponding 147 ΔE_{SCF}), which are different from the first 588 configurations, were used to test the quality of the function. The energy points (ΔE_{FIT}) due to the tested configurations were calculated using the pair potential with the optimal parameter for the two fitted methods shown in Table 5.2 and 5.3.

The standard deviations for the tested data set (σ_{test}) have been calculated for both methods, the values of 0.195 and 0.42 kcal.mol⁻¹ have, respectively yielded.

Characteristics of the test in which the energy points (ΔE_{FIT}) were calculated using the pair potential with the optimal parameter shown in Table 5.2:

Minimum energy included in the test	=	-4.928	kcal.mol ⁻¹ .
Maximum energy included in the test	=	10.00	kcal.mol ⁻¹ .
Maximum absolute residual	=	3.15	kcal.mol ⁻¹ .
Minimum absolute residual	=	0.017	kcal.mol ⁻¹ .
Average absolute residual	=	0.607	kcal.mol ⁻¹ .
Standard deviation (σ_{test})	=	0.195	kcal.mol ⁻¹ .

Characteristics of the test in which the energy points (ΔE_{FIT}) were calculated using the pair potential with the optimal parameter shown in Table 5.3:

Minimum energy included in the test =	-4.928	kcal.mol ⁻¹ .
Maximum energy included in the test=	10.00	kcal.mol ⁻¹ .
Maximum absolute residual	=	8.46 kcal.mol ⁻¹ .
Minimum absolute residual	=	0.024 kcal.mol ⁻¹ .
Average absolute residual	=	1.34 kcal.mol ⁻¹ .
Standard deviation (σ_{test})	=	0.42 kcal.mol ⁻¹ .

The 147 ΔE_{SCF} were, then, added to the fit and the standard deviations ($\sigma_{fit 2}$) of 0.194 and 0.43 kcal.mol⁻¹, respectively observed. The final optimal parameters are given in table 5.4 and 5.5 for the method 1 and 2, respectively.

Table 5.4. The final optimized parameters (the derivatives were approximated using the central different technique) representing interaction of N and H atoms of ammonia with lithium ion (energy and r are in kcal.mol⁻¹ and angstroms, respectively).

Atom	A (\AA^6 kcal.mol ⁻¹)	B (kcal.mol ⁻¹)	C (kcal.mol ⁻¹)
NL _i	-0.10277754E+04	-0.37715706E+03	0.11240189E+01
HL _i	-0.14943341E+04	-0.73518057E+05	0.43047872E+01

Characteristics of the fit in which the derivatives were approximated using the central different technique:

Minimum energy included in the fit	=	-40.83	kcal.mol ⁻¹ .
Maximum energy included in the fit	=	10.00	kcal.mol ⁻¹ .
Maximum absolute residual	=	4.97	kcal.mol ⁻¹ .
Minimum absolute residual	=	0.4E-03	kcal.mol ⁻¹ .
Average absolute residual	=	0.60	kcal.mol ⁻¹ .
Number of iteration	=	48	cycle.
Standard deviation ($\sigma_{fit 2}$)	=	0.194	kcal.mol ⁻¹ .

Table 5.5. The final optimized parameters (the derivatives were calculated analytically) representing interaction of N and H atoms of ammonia with lithium ion (energy and r are in kcal.mol⁻¹ and angstroms, respectively).

Atom	A (Å ⁶ kcal.mol ⁻¹)	B (kcal.mol ⁻¹)	C (kcal.mol ⁻¹)
NL _i	-0.78853383E+03	-0.24192296E+03	0.10348343E+01
HL _i	0.43054531E+03	0.21654474E+04	0.24836065E+01

Characteristics of the fit in which the derivatives were calculated analytically:

Minimum energy included in the fit	=	-40.83	kcal.mol ⁻¹ .
Maximum energy included in the fit	=	10.00	kcal.mol ⁻¹ .
Maximum absolute residual	=	21.14	kcal.mol ⁻¹ .
Minimum absolute residual	=	0.3E-03	kcal.mol ⁻¹ .
Average absolute residual	=	1.25	kcal.mol ⁻¹ .
Number of iteration	=	67	cycle.
Standard deviation ($\sigma_{fit 2}$)	=	0.43	kcal.mol ⁻¹ .

It can be seen from the characteristic data that both methods are well fit, especially the first method, with σ_{fit1} of 0.196 and 0.41 kcal.mol⁻¹. Moreover, the quality of the fit is confirmed by the testing procedure. The standard deviations for the tested data set (σ_{test}) for both methods are 0.195 and 0.42 kcal.mol⁻¹, respectively. After inclusion of the test set, the σ_{fit2} for both methods are 0.194 and 0.43 kcal.mol⁻¹, respectively.

5.1.6 Search for False Minima of the Function

Now come to the last step of the development of the potential function, false minima which lower than the global minimum of the pair was searched. In this step, subroutine **Genco** will be used to generated 2322 ammonia-lithium configurations, which different from those included in the fitted and tested sets, where $0^\circ \leq \theta \leq 150^\circ$ and $0^\circ \leq \phi \leq 60^\circ$ in steps of 10° and $1.20 \leq r \leq 10.00 \text{ \AA}$ (see Figure 4.11). Then, the 2322 energy points have been calculated using the pair potential function.

The results show that among the generated 2322 data points, all of them are higher than the global minimum of the pair (-40.83 kcal.mol⁻¹). This means that no artificial minima were found in the fitted function.

The minimum interaction energy between ammonia and lithium is found at the distance of 2.0 Å, $\theta = 0$, $\phi = 0$ as given in Table 5.6 (ΔE_{FIT} which the derivatives calculate using central different technique (ΔE_{FIT}^1) and calculated analytically (ΔE_{FIT}^2)). Some of the computed energies with different values of θ and ϕ are shown in Table 5.7 (a-c) and plotted in Figure 5.2 and Figure 5.3, respectively.

Table 5.6. The energies from *ab initio* calculations (ΔE_{SCF}) and from the pair potential function which the derivatives calculated using central different technique (ΔE_{FIT}^1) and calculated analytically (ΔE_{FIT}^2) (energy and r are in kcal.mol⁻¹ and angstroms, respectively).

Distance	ΔE_{SCF}	ΔE_{FIT}^1	ΔE_{FIT}^2
1.6000	-22.933	-23.485	-27.970
1.7000	-32.262	-33.715	-35.931
1.8000	-37.558	-38.762	-39.713
1.9000	-40.115	-40.724	-40.784
<u>2.0000</u>	-40.834	-40.830	-40.738
2.1000	-40.347	-39.857	-39.588
2.2000	-39.099	-38.238	-37.912
2.3000	-37.397	-36.272	-35.954
2.4000	-35.451	-34.141	-33.867
2.5000	-33.401	-31.963	-31.750
2.6000	-31.337	-29.811	-29.665
3.0000	-23.812	-22.134	-22.233
3.4000	-18.005	-16.333	-16.572
3.8000	-13.780	-12.175	-12.465
4.2000	-10.787	-9.2376	-9.524
5.2000	-6.5240	-5.0967	-5.293
6.2000	-4.3980	-3.2107	-3.319
8.2000	-2.3830	-1.6756	-1.702

Table 5.7. Lithium ion-ammonia interaction energies from *ab initio* calculation (ΔE_{SCF}) and from the pair potential function which the derivatives calculated using central different technique (ΔE_{FIT}^1) and calculated analytically (ΔE_{FIT}^2), for (a) $\theta = 90$ and $\phi = 0$, (b) $\theta = 60$ and $\phi = 0$, (c) $\theta = 30$ and $\phi = 0$ (energy and r are in kcal.mol⁻¹ and angstroms, respectively).

(a) $\theta = 90$ and $\phi = 0$

Distance	ΔE_{SCF}	ΔE_{FIT}^1	ΔE_{FIT}^2
1.6000	23.604	23.604	23.604
1.7000	12.233	12.233	12.233
1.8000	4.8870	5.2584	11.127
1.9000	0.26600	1.4556	7.261
2.0000	-2.5020	-0.57173	3.817
2.1000	-4.0320	-1.8079	0.873
2.2000	-4.7430	-2.6909	-1.540
2.3000	-4.9280	-3.3928	-3.442
2.4000	-4.7870	-3.9677	-4.876
2.5000	-4.4580	-4.4253	-5.900
2.6000	-4.0340	-4.7644	-6.578
3.0000	-2.3070	-5.0375	-6.996
3.4000	-1.1780	-4.2169	-5.738
3.8000	-0.57000	-3.1199	-4.223
4.2000	-0.25100	-2.1527	-2.944
5.2000	0.17000E-01	-0.72141	-1.070
6.2000	0.61000E-01	-0.19956	-0.357
8.2000	0.56000E-01	0.11937E-01	-0.21884E-01

(b) $\theta = 60$ and $\phi = 0$

Distance	ΔE_{SCF}	ΔE_{FIT}^1	ΔE_{FIT}^2
1.6000	-4.2320	0.14441	1.383
1.7000	-12.605	-10.418	-8.293
1.8000	-17.471	-16.176	-14.278
1.9000	-20.026	-19.203	-17.916
2.0000	-21.059	-20.641	-20.011
2.1000	-21.118	-21.127	-21.062
2.2000	-20.568	-21.025	-21.395
2.3000	-19.655	-20.554	-21.231
2.4000	-18.545	-19.845	-20.726
2.5000	-17.346	-18.986	-19.990
2.6000	-16.131	-18.036	-19.106
3.0000	-11.774	-14.023	-15.056
3.4000	-8.6110	-10.471	-11.325
3.8000	-6.4490	-7.7164	-8.393
4.2000	-4.9780	-5.7054	-6.232
5.2000	-2.9630	-2.8808	-3.150
6.2000	-1.9930	-1.6766	-1.809
8.2000	-1.0870	-0.81299	-0.843

(c) $\theta = 30$ and $\phi = 0$

Distance	ΔE_{SCF}	ΔE_{FIT}^1	ΔE_{FIT}^2
1.6000	-18.264	-16.892	-20.176
1.7000	-27.085	-27.311	-28.812
1.8000	-32.107	-32.669	-33.287
1.9000	-34.575	-35.020	-35.225
2.0000	-35.340	-35.565	-35.599
2.1000	-34.993	-35.021	-35.008
2.2000	-33.940	-33.830	-33.830
2.3000	-32.463	-32.267	-32.309
2.4000	-30.755	-30.509	-30.602
2.5000	-28.944	-28.669	-28.816
2.6000	-27.116	-26.820	-27.017
3.0000	-20.453	-20.049	-20.392
3.4000	-15.369	-14.803	-15.200
3.8000	-11.718	-10.997	-11.386
4.2000	-9.1480	-8.2964	-8.644
5.2000	-5.5180	-4.4995	-4.715
6.2000	-3.7220	-2.7952	-2.910
8.2000	-2.0210	-1.4412	-1.468

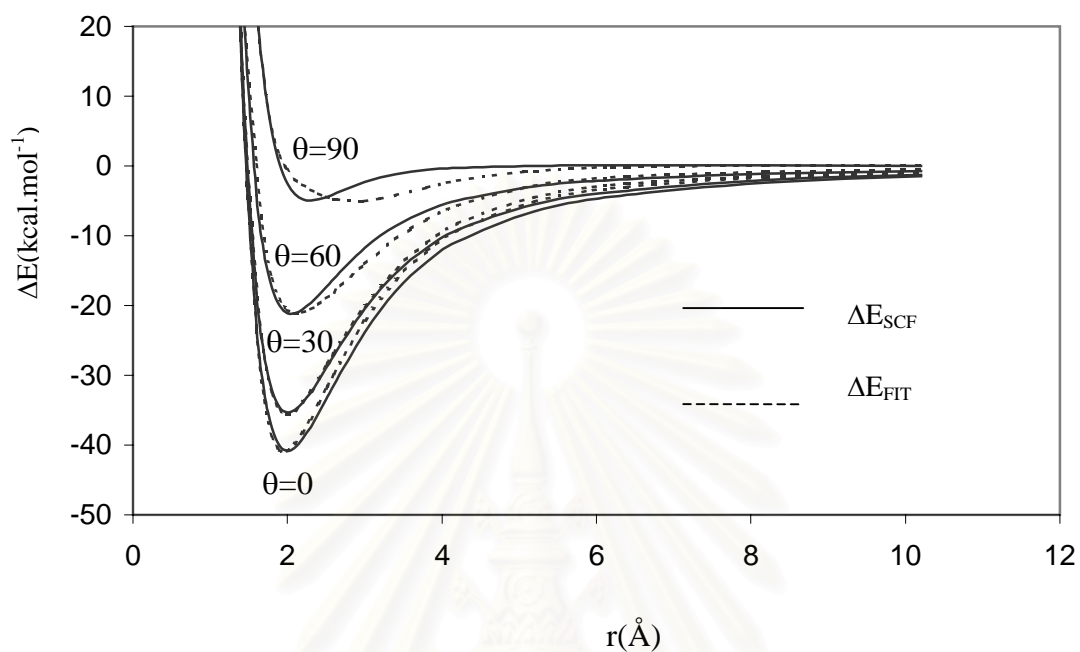


Figure 5.2 The stabilization energies obtained from the DZP-ab initio (ΔE_{SCF}) and from the fitting analytical potential (ΔE_{FIT}), which the derivatives calculated using central different technique, with the parameter given in Table 5.4 ($\phi = 0^\circ$ see Figure 4.11).

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

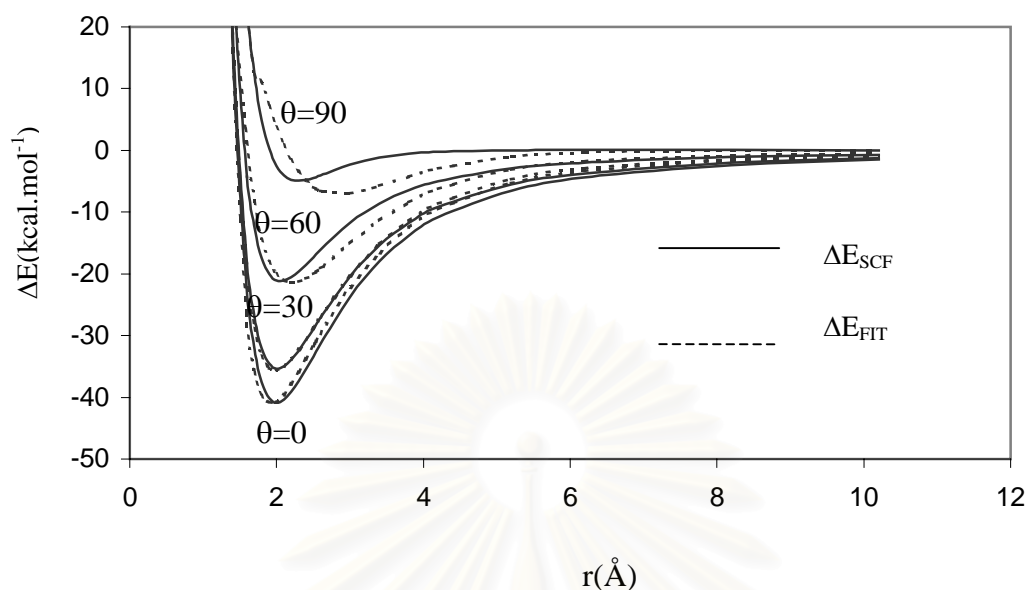


Figure 5.3 The stabilization energies obtained from the DZP-ab initio (ΔE_{SCF}) and from the fitting analytical potential (ΔE_{FIT}), which the derivatives calculated analytically, with the parameter given in Table 5.5 ($\phi = 0^\circ$ see Figure 4.11).

It is clearly seen that distance to the energy minimum obtained from the two sources takes place at the same position. Good agreement was yielded, especially in the area around the minimum. In addition, all data points were, again, compared in Figure 5.4 and Figure 5.5. It is clearly seen that the energies, obtained from the functions are good agreement with the SCF energies, especially for the attractive regions, which is important for prediction of the simulation results.

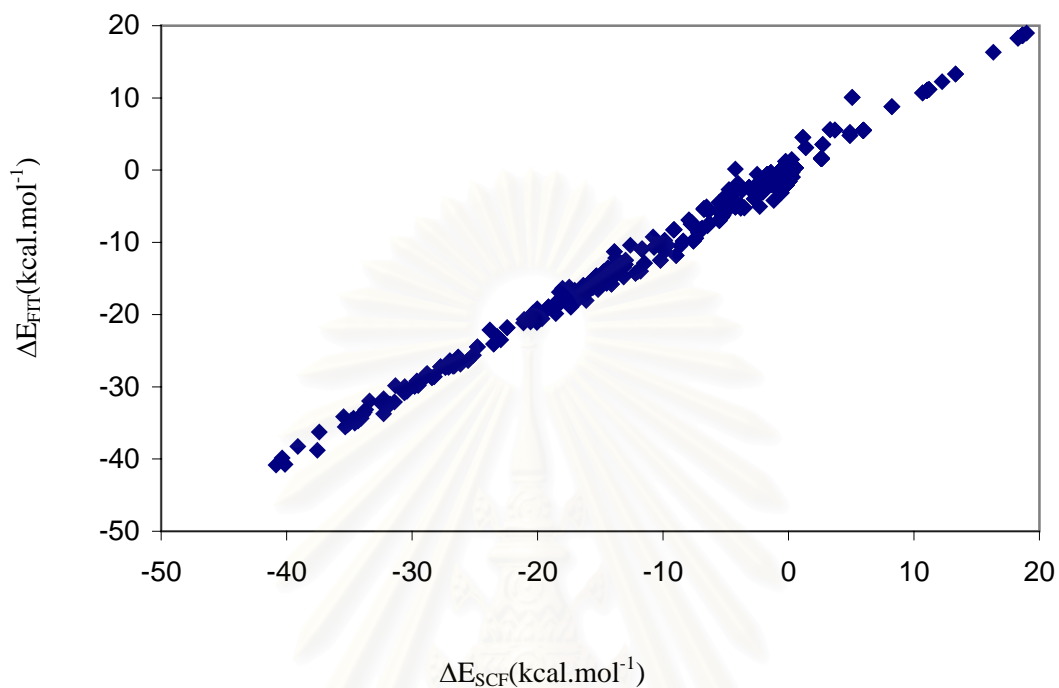


Figure 5.4 Comparison of the stabilization energies from the DZP-ab initio (ΔE_{SCF}) and the potential function (ΔE_{FIT}), which the derivatives calculated using central different technique, with the final values of the fitting parameters given in Table 5.4.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

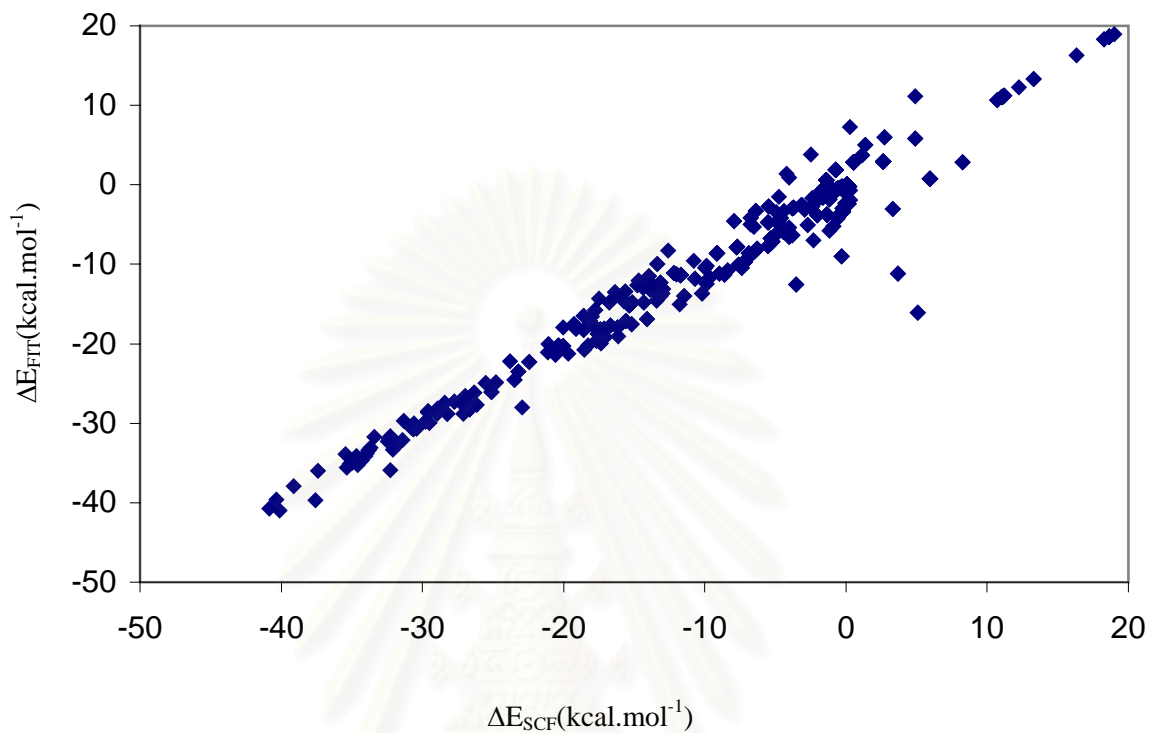


Figure 5.5 Comparison of the stabilization energies from the DZP-ab initio (ΔE_{SCF}) and the potential function (ΔE_{FIT}), which the derivatives calculated analytically, with the final values of the fitting parameters given in Table 5.5.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

5.2 Development of Ammonia - Ammonia Potential Function

5.2.1 Input Elementary Data

In this example, the fitting procedure contains 21 trajectories and 819 SCF energy points in the geometries mentioned in Figure 4.12. Then, one-fifth of the trajectory, taking one every five lines, will be kept for testing.

The selected 5 trajectories as well as the corresponding 195 SCF energies were used for testing. Here, the other 16 trajectories with 624 SCF energy points were used for fitting.

5.2.2 Selection of the Functional Form

Here, the energy points for the ammonia-ammonia dimer, with choice 2 of the functional form, $x = 4$, $y = 7$ and $z = 12$ (see equation 4.3), were displayed, that is

$$\Delta E_{\text{NH}_3-\text{NH}_3} = \sum_i^k \sum_j^l \left(-\frac{A_{ij}}{r_{ij}^x} + \frac{B_{ij}}{r_{ij}^y} + \frac{q_i q_j}{r_{ij}} + \frac{C_{ij}}{r_{ij}^z} \right) \quad (5.2)$$

where $k = 4$ and $l = 4$ (number of atoms of ammonia molecule), A_{ij} , B_{ij} and C_{ij} are fitting parameters, r_{ij} is the distance, $q_i = -0.74207$ and 0.24736 atomic unit for atoms N and H of ammonia.

5.2.3 Set an Initial Guess for the Parameters

In this example, equation (5.2) contains 9 parameters namely A_{NN} , B_{NN} , C_{NN} , A_{NH} , B_{NH} , C_{NH} , A_{HH} , B_{HH} , C_{HH} . The first 3 parameters represents interaction between atom N and N, the second one represents interaction between atom N and H while the last 3 parameters for the H and H pair.

Note that due to symmetry reason, all 3 hydrogen atoms of the two NH_3 are represented by the same set of fitting parameters.

- The initial guess of parameters

A_{NN}	=	99.7486877
B_{NN}	=	269.96283
C_{NN}	=	2017.52527
A_{NH}	=	147.81398
B_{NH}	=	475.154297
C_{NH}	=	2221.89502
A_{HH}	=	68.1932144
B_{HH}	=	251.712234
C_{HH}	=	2970.36646

5.2.4 Fitting Energies to the Functional Form

With 624 ammonia-ammonia interaction energies (see Section 5.2.1), the pair lower than $5.2 \text{ kcal.mol}^{-1}$ were fitted to the analytical function, equation 5.2.

The optimized parameters of fit in which using the first method are given in the Table 5.8.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Table 5.8. The optimized parameters (the derivatives were approximated using the central different technique) representing interaction of atoms between the two ammonia molecules (energy and r are in kcal.mol⁻¹ and angstroms, respectively).

Atom	A (Å ⁴ kcal.mol ⁻¹)	B (Å ⁷ kcal.mol ⁻¹)	C (Å ¹² kcal.mol ⁻¹)
NN	-0.32110724E+04	-0.72573423E+04	-0.25824919E+06
NH	0.85169143E+03	0.51342317E+04	-0.34077085E+05
HH	-0.20315592E+03	-0.99981208E+03	0.57742285E+04

Characteristics of the fit in which the derivatives were approximated using the central different technique:

Minimum energy included in the fit	=	-1.007	kcal.mol ⁻¹ .
Maximum energy included in the fit	=	5.20	kcal.mol ⁻¹ .
Maximum absolute residual	=	0.66	kcal.mol ⁻¹ .
Minimum absolute residual	=	0.001	kcal.mol ⁻¹ .
Average absolute residual	=	0.089	kcal.mol ⁻¹ .
Number of iteration	=	7	cycle.
Standard deviation (σ_{fit1})	=	0.129	kcal.mol ⁻¹ .

The optimized parameters of fit in which using the second method are given in the Table 5.9.

Table 5.9. The optimized parameters (the derivatives were calculated analytically) representing interaction of atoms between the two ammonia molecules (energy and r are in kcal.mol⁻¹ and angstroms, respectively).

Atom	A (Å ⁴ kcal.mol ⁻¹)	B (Å ⁷ kcal.mol ⁻¹)	C (Å ¹² kcal.mol ⁻¹)
NN	-0.11872595E+04	0.10477691E+05	-0.35927283E+05
NH	0.32843524E+03	0.11187137E+04	-0.10420567E+04
HH	-0.60452460E+02	-0.79040317E+02	0.30113348E+04

Characteristics of the fit in which the derivatives were calculated analytically:

Minimum energy included in the fit	=	-1.007	kcal.mol ⁻¹ .
Maximum energy included in the fit	=	5.20	kcal.mol ⁻¹ .
Maximum absolute residual	=	3.64	kcal.mol ⁻¹ .
Minimum absolute residual	=	0.5E-03	kcal.mol ⁻¹ .
Average absolute residual	=	0.17	kcal.mol ⁻¹ .
Number of iteration	=	500	cycle.
Standard deviation (σ_{fit1})	=	0.246	kcal.mol ⁻¹ .

5.2.5 Testing the Quality of the Function

In this part, the 195 ammonia-ammonia configurations (with the corresponding 195 ΔE_{SCF}), which are different from the first 624 configurations, were used to test the quality of the function. The energy points (ΔE_{FIT}) due to the tested configurations were calculated using the pair potential with the optimal parameter for the two fitted methods shown in Table 5.8 and 5.9.

The standard deviations for the tested data set (σ_{test}) have been calculated for both methods, the values of 0.23 and 0.414 kcal.mol⁻¹ have, respectively yielded.

Characteristics of the test in which the energy points (ΔE_{FIT}) were calculated using the pair potential with the optimal parameter shown in Table 5.8:

Minimum energy included in the test =	-1.007	kcal.mol ⁻¹ .
Maximum energy included in the test =	5.20	kcal.mol ⁻¹ .
Maximum absolute residual =	0.66	kcal.mol ⁻¹ .
Minimum absolute residual =	0.001	kcal.mol ⁻¹ .
Average absolute residual =	0.124	kcal.mol ⁻¹ .
Standard deviation (σ_{test}) =	0.23	kcal.mol ⁻¹ .

Characteristics of the test in which the energy points (ΔE_{FIT}) were calculated using the pair potential with the optimal parameter shown in Table 5.9:

Minimum energy included in the test =	-1.007	kcal.mol ⁻¹ .
Maximum energy included in the test =	5.20	kcal.mol ⁻¹ .
Maximum absolute residual =	1.46	kcal.mol ⁻¹ .
Minimum absolute residual =	0.5E-03	kcal.mol ⁻¹ .
Average absolute residual =	0.17	kcal.mol ⁻¹ .
Standard deviation (σ_{test}) =	0.414	kcal.mol ⁻¹ .

The 195 ΔE_{SCF} were, then, added to the fit and the standard deviations ($\sigma_{fit 2}$) of 0.139 and 0.406 kcal.mol⁻¹, respectively observed. The final optimal parameters are given in Table 5.10 and 5.11 for the method 1 and 2, respectively.

Table 5.10. The final optimized parameters (the derivatives were approximated using the central different technique) representing interaction of atoms between the two ammonia molecules (energy and r are in kcal.mol⁻¹ and angstroms, respectively).

Atom	A (Å ⁴ kcal.mol ⁻¹)	B (Å ⁷ kcal.mol ⁻¹)	C (Å ¹² kcal.mol ⁻¹)
NN	-0.31868972E+04	-0.72983431E+04	-0.25665577E+06
NH	0.85014415E+03	0.51776934E+04	-0.349114735E+05
HH	-0.20485087E+03	-0.10305909E+04	0.61156608E+04

Characteristics of the fit in which the derivatives were approximated using the central different technique:

Minimum energy included in the fit	=	-1.007	kcal.mol ⁻¹ .
Maximum energy included in the fit	=	5.20	kcal.mol ⁻¹ .
Maximum absolute residual	=	0.40	kcal.mol ⁻¹ .
Minimum absolute residual	=	0.001	kcal.mol ⁻¹ .
Average absolute residual	=	0.079	kcal.mol ⁻¹ .
Number of iteration	=	7	cycle.
Standard deviation ($\sigma_{fit 2}$)	=	0.139	kcal.mol ⁻¹ .

Table 5.11. The final optimized parameters (the derivatives were calculated analytically) representing interaction of atoms between the two ammonia molecules (energy and r are in kcal.mol⁻¹ and angstroms, respectively).

Atom	A (Å ⁴ kcal.mol ⁻¹)	B (Å ⁷ kcal.mol ⁻¹)	C (Å ¹² kcal.mol ⁻¹)
NN	-0.11229922E+04	0.12181375E+05	-0.33087938E+06
NH	0.31530579E+03	0.10815344E+04	-0.12755814E+04
HH	-0.55114130E+02	-0.401867519E+02	0.30061042E+04

Characteristics of the fit in which the derivatives were calculated analytically:

Minimum energy included in the fit = -1.007 kcal.mol⁻¹.

Maximum energy included in the fit = 5.20 kcal.mol⁻¹.

Maximum absolute residual = 4.06 kcal.mol⁻¹.

Minimum absolute residual = 0.8E-04 kcal.mol⁻¹.

Average absolute residual = 0.18 kcal.mol⁻¹.

Number of iteration = 301 cycle.

Standard deviation ($\sigma_{fit 2}$) = 0.406 kcal.mol⁻¹.

5.2.6 Search for False Minima of the Function

Now come to the last step of the development of the potential function, false minima which lower than the global minimum of the pair was searched. In this step, subroutine **Genco** will be used to generated 3610 ammonia-ammonia configurations, which different from those included in the fitted and tested sets, where $0^\circ \leq \theta \leq 360^\circ$

and $0^\circ \leq \phi \leq 360^\circ$ in steps of 20° and $2.30 \leq r \leq 10.00 \text{ \AA}$. Then, the 3610 energy points have been calculated using the pair potential function.

The results show that among the generated 3610 data points, all of them are higher than the global minimum of the pair ($-1.007 \text{ kcal.mol}^{-1}$). This means that no artificial minima were found in the fitted function.

The minimum interaction energy between ammonia and ammonia is found at the distance of 3.8 \AA , $\theta = 150$, $\phi = 0$ as given in Table 5.12 (ΔE_{FIT} which the derivatives calculate using central different technique (ΔE_{FIT}^1) and calculated analytically (ΔE_{FIT}^2)). Some of the computed energies with different values of θ and ϕ are shown in Table 5.13 (a,b) and plotted in Figure 5.6 and Figure 5.7, respectively.

Table 5.12. The energies from *ab initio* calculations (ΔE_{SCF}) and from the pair potential function which the derivatives calculated using central different technique (ΔE_{FIT}^1) and calculated analytically (ΔE_{FIT}^2) (energy and r are in kcal.mol⁻¹ and angstroms, respectively).

Distance	ΔE_{SCF}	ΔE_{FIT}^1	ΔE_{FIT}^2
3.3000	-0.377	-0.297	-0.28963
3.3500	-0.537	-0.475	-0.44531
3.4000	-0.665	-0.617	-0.57279
3.4500	-0.765	-0.729	-0.67636
3.5000	-0.843	-0.816	-0.75969
3.5500	-0.902	-0.882	-0.82588
3.6000	-0.945	-0.931	-0.87756
3.6500	-0.975	-0.965	-0.91697
3.7000	-0.994	-0.987	-0.94602
3.7500	-1.004	-0.999	-0.96633
<u>3.8000</u>	-1.007	-1.004	-0.97928
3.8500	-1.004	-1.002	-0.98605
3.9000	-0.997	-0.994	-0.98764
3.9500	-0.985	-0.983	-0.98491
4.0000	-0.970	-0.968	-0.97860
4.0500	-0.953	-0.950	-0.96932
6.0500	-0.307	-0.280	-0.35639
8.0500	-0.127	-0.103	-0.13848

Table 5.13. Ammonia-ammonia interaction energies from *ab initio* calculation (ΔE_{SCF}) and from the pair potential function which the derivatives calculated using central different technique (ΔE_{FIT}^1) and calculated analytically (ΔE_{FIT}^2), for (a) $\theta = 150$ and $\phi = 30$, (b) $\theta = 150$ and $\phi = 60$ (energy and r are in kcal.mol⁻¹ and angstroms, respectively).

(a) $\theta = 150$ and $\phi = 30$

Distance	ΔE_{SCF}	ΔE_{FIT}^1	ΔE_{FIT}^2
3.3000	-0.022	-0.054	0.034
3.3500	-0.200	-0.133	-0.144
3.4000	-0.345	-0.287	-0.293
3.4500	-0.463	-0.412	-0.417
3.5000	-0.558	-0.513	-0.519
3.5500	-0.633	-0.593	-0.603
3.6000	-0.692	-0.657	-0.671
3.6500	-0.737	-0.705	-0.725
3.7000	-0.771	-0.742	-0.768
3.7500	-0.795	-0.769	-0.801
3.8000	-0.811	-0.787	-0.826
3.8500	-0.821	-0.798	-0.844
3.9000	-0.825	-0.803	-0.855
3.9500	-0.824	-0.803	-0.862
4.0000	-0.820	-0.799	-0.864
4.0500	-0.813	-0.792	-0.863
6.0500	-0.293	-0.266	-0.346
8.0500	-0.124	-0.101	-0.136

(b) $\theta = 150$ and $\phi = 60$

Distance	ΔE_{SCF}	ΔE_{FIT}^1	ΔE_{FIT}^2
3.3000	0.310	0.393	0.337
3.3500	0.116	0.196	0.137
3.4000	-0.044	0.031	-0.031
3.4500	-0.178	-0.107	-0.173
3.5000	-0.288	-0.222	-0.293
3.5500	-0.378	-0.316	-0.393
3.6000	-0.451	-0.394	-0.476
3.6500	-0.510	-0.457	-0.545
3.7000	-0.558	-0.508	-0.601
3.7500	-0.595	-0.548	-0.646
3.8000	-0.623	-0.579	-0.682
3.8500	-0.645	-0.603	-0.710
3.9000	-0.660	-0.620	-0.731
3.9500	-0.670	-0.631	-0.746
4.0000	-0.676	-0.638	-0.756
4.0500	-0.678	-0.641	-0.762
6.0500	-0.278	-0.252	-0.336
8.0500	-0.121	-0.098	-0.134

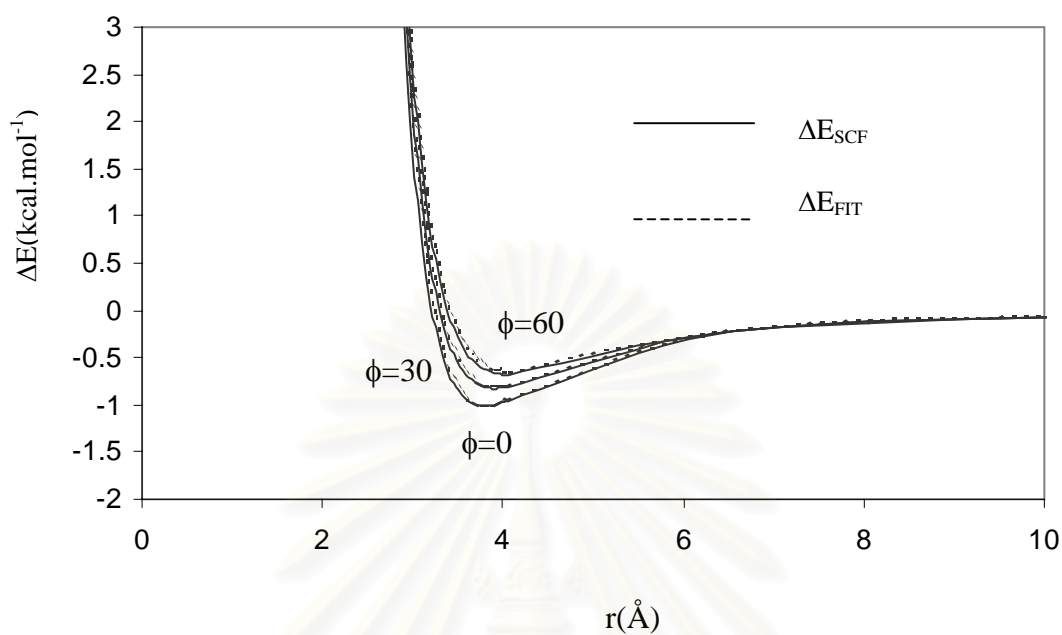


Figure 5.6 The stabilization energies obtained from the DZP-ab initio (ΔE_{SCF}) and from the fitting analytical potential (ΔE_{FIT}), which the derivatives calculated using central different technique, with the parameter given in Table 5.10 ($\theta = 150^\circ$ see Figure 4.12).

It is clearly seen that distance to the energy minimum obtained from the two sources takes place at the same position. Good agreement was yielded, especially in the area around the minimum.

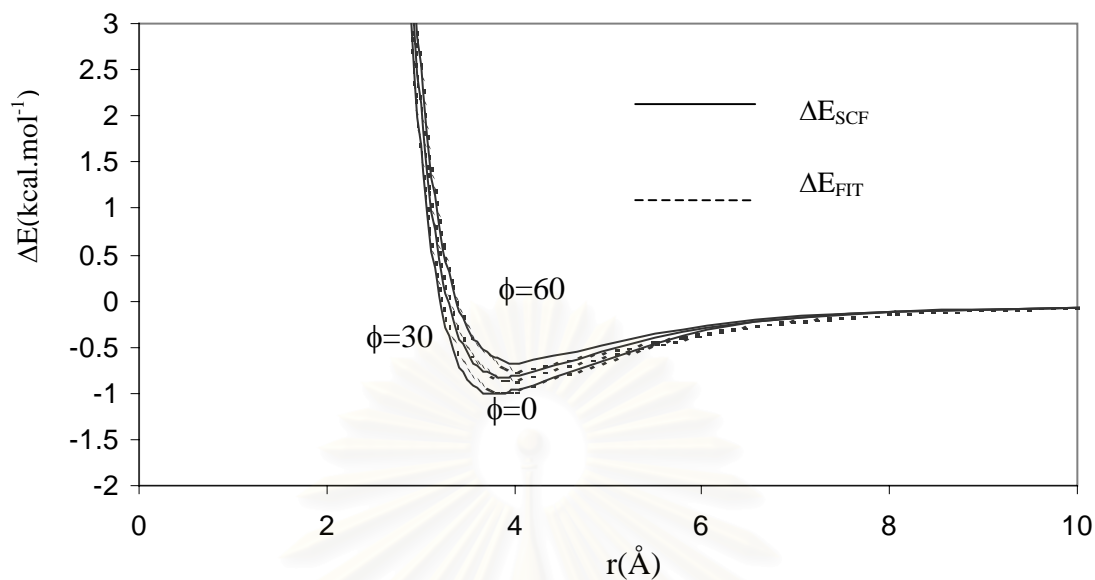


Figure 5.7 The stabilization energies obtained from the DZP-ab initio (ΔE_{SCF}) and from the fitting analytical potential (ΔE_{FIT}), which the derivatives calculated analytically, with the parameter given in Table 5.11 ($\theta = 150^\circ$ see Figure 4.12).

In addition, all data points were, again, compared in Figure 5.8 and Figure 5.9 for first and second method, respectively. In the first method, it is clearly seen that the energies, obtained from the functions are good agreement with the SCF energies, especially for the attractive regions, which is important for prediction of the simulation results.

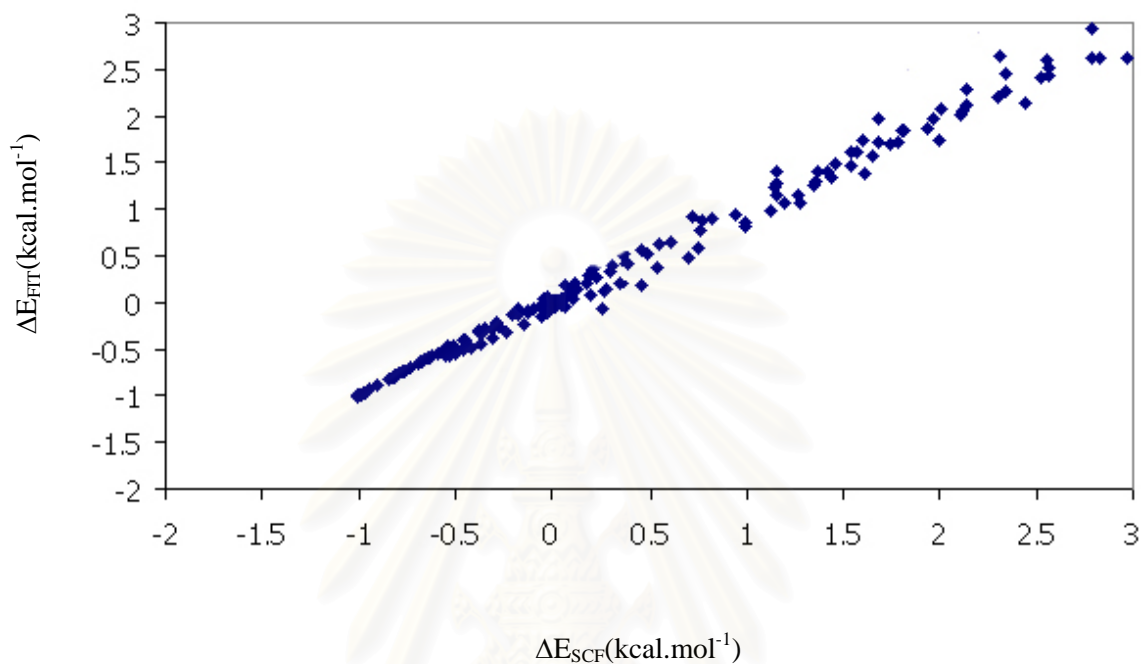


Figure 5.8 Comparison of the stabilization energies from the DZP-ab initio (ΔE_{SCF}) and the potential function (ΔE_{FIT}), which the derivatives calculated using central different technique, with the final values of the fitting parameters given in Table 5.10.

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

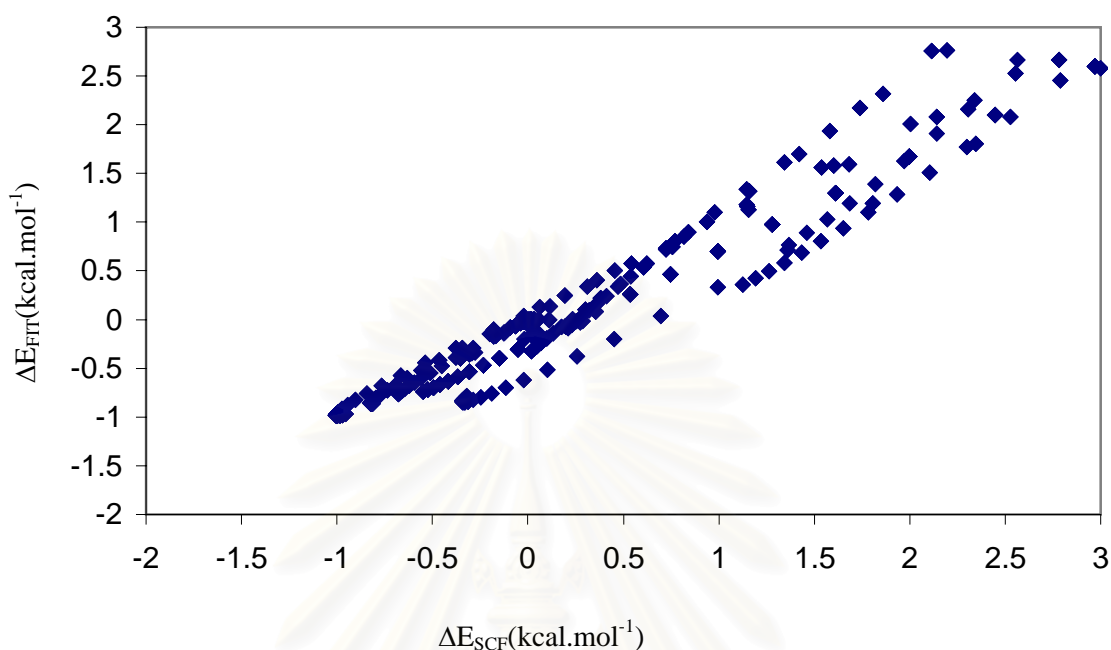


Figure 5.9 Comparison of the stabilization energies from the DZP-ab initio (ΔE_{SCF}) and the potential function (ΔE_{FIT}), which the derivatives calculated analytically, with the final values of the fitting parameters given in Table 5.11.

As mention in Chapter 4 in section 4.2.4 (see equations 4.13 - 4.17) and taking into account all the characteristics data in term of statistical values, we found that the first method where the partial derivative with respect to the parameters (term $\frac{\partial y(x_i; a)}{\partial a_k}$ in equation 3.12) are calculated analytically, is not appropriate especially in the first form because the obtained derivatives are constant for every fitting iterations. Hence, we should select more term and construct the Lagrange's function for the χ^2 merit function that will be used to minimize to get *goodness-of-fit* (see equation 3.11, 3.12 and 4.10) and minimize it. It may give a better result.

Another approach is approximated the partial derivative using the central different technique which is independent of the model we use. This method shows the reliable result so we suggest using this approximation to extract the parameter.



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

REFERENCES

1. Allen, M. P. and Tildesley, D. J. 1987. Computer Simulation of Liquids. New York: Oxford University Press.
2. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. 1953. Equation of State Calculation by Computing Machines. J. Chem. Phys. 21(6): 1087-1092.
3. Alder, B. J., and Wainwright, T. E. 1960. Studied in Molecular Dynamics. II. Behavior of a Small Number of Elastic Spheres. J. Chem. Phys. 33: 1439-1451.
4. Tanabe, Y., and Rode, B. M. 1988. Monte Carlo Simulation of an 18.45 mol % Aqueous Ammonia Solution. J. Chem. Soc. Faraday. Trans. II 84: 679-692.
5. Hannongbua, S., Ishida, T., Spohr, E., and Heinzinger, K. 1988. Molecular Dynamics Study of a Lithium Ion in Ammonia. Z. Naturforsch. 43a: 572-582.
6. Kheawsrikul, S., Hannongbua, S. V., Kokpol, S. U. and Rode, B. M. 1989. A Monte Carlo Study on Preference Solvation of Lithium(I) in Aqueous Ammonia. J. Chem. Soc. Faraday. Trans. II 85: 643-649.
7. Hannongbua, S., Kerdcharoen, T., and Rode, B. M. 1992. Zinc(II) in Liquid Ammonia: Intermolecular Potential Including Three-Body Terms and Monte Carlo Simulation. J. Chem. Phys. 96: 6945-6949.
8. Alagona, G., Ghio, C., and Kollman, P. A. 1985. Monte Carlo Simulations of the Solvation of the Dimethyl Phosphate Anion. J. Am. Chem. Soc. 107: 2229-2238.
9. Lie, G. C., Clementi, E. and Yoshimine, M. 1975. Study of the structural of molecular complexes XIII: Monte Carlo simulation of liquid water with a configuration interaction pair potential. J. Chem. Phys. 64(6), 2314-2323.

10. Swaminathan, S., Harrison, S. W. and Beveridge, D. L. 1977. Monte Carlo studies on the structure of a dilute aqueous solution of methane. J. Am. Chem. Soc. 100(8), 5705-5712.
11. Jorgensen, W. L. and Ibrahim, M. 1980. Structure and properties of liquid ammonia. J. Am. Chem. Soc. 102(10), 3309-3315.
12. Jorgensen, W. L. 1981. Simulation of liquid ethanol including internal rotation. J. Am. Chem. Soc. 103(2), 345-350.
13. Bolis, G., Corongiu, G. and Clementi, E. 1981. Methanol in water solution at 300 K. Chem. Phys. Lett. 6(3), 299-306.
14. Jorgensen, W. L. and Madura, J. D. 1983. Solvation and conformation of methanol in water. J. Am. Chem. Soc. 105(6), 863-869.
15. Linse, P., Karlstrom, G. and Jonsson, B. 1984. Monte Carlo studies of a dilute aqueous solution of benzene. J. Am. Chem. Soc. 106(15), 4096-4102.
16. Byrnes, J. M. and Sandler, S. I. 1984. Monte Carlo simulation of liquid ethane. J. Chem. Phys. 80(2), 881-885.
17. Dietz, W. and Heinzinger, K. 1984. Structure of liquid chloroform: A comparison between computer simulation and neutron scattering results. Ber. Bunsenges. Phys. Chem. 88, 543-546.
18. Okazaki, S., Touhara, H. and Nakanishi, K. 1984. Computer experiments of aqueous solutions V: Monte Carlo calculation on the hydrophobic interaction in 5 mol % methanol solution. J. Chem. Phys. 81(2), 890-894.
19. Thongthep, M. 2001. Fortran programming to generate and calculate complexation energy for the development of intermolecular pair potential function. Master's Thesis, Computational Science, Faculty of Science, Chulalongkorn University.
20. Borse, G. J. 1991. FORTRAN77 and Numerical Methods for Engineers. 2nd ed. Boston: PWS-Kent Publishing.

21. Davis, G. B. and Hoffmann, T. R. 1988. FORTRAN77: A Structured, Disciplined Style. 3rd ed. New York: McGraw-Hill.
22. Leach, A. R. 1996. Molecular Modelling: Principles and Applications. 1st ed. Essex: Addison Wesley Longman Limited.
23. Mulliken, R. S. 1955. J. Chem. Phys. 23, 1833, 1841, 2338, 2343.
24. Press, W. H., and others. 1992. Numerical Recipes in Fortran: The Art of Scientific Computing. New York: Cambridge University Press.
25. Rao, S. S. 1984. Optimization Theory and Applications. New Delhi: Wiley Eastern.
26. Arora, J. S. 1989. Introduction to Optimum Design. New York: McGraw-Hill.
27. Rardin, R. L. 1998. Optimization in Operations Research. N.J.: Prentice-Hall International.
28. Luenberger, D. G. 1984. Linear and Nonlinear Programming. New York: Addison-Wesley.
29. Bazaraa, M. S. 1993. Nonlinear Programming: Theory and Algorithms. New York: Wiley.
30. Taha, H. A. 1997. Operations Research: An Introduction. N.J: Prentice Hall.
31. Evtushenko, Y. G. 1985. Numerical Optimization Techniques. New York: Optimization Software, Inc., Publications Division.
32. Hoffman, J. D. 1993. Numerical Methods for Engineers and Scientists. New York: McGraw-Hill.
33. Murphy, J. 1988. Numerical Analysis, Algorithms and Computation. Chichester: Ellis Horwood.



APPENDICES

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Appendix I

MRQFIT version 1.0

Program Manual

Program MRQFIT version 1.0 has been developed to fit the interaction energies to the functional form (the intermolecular potential function) and improve quality of the function. The source code has been written in Fortran language using FORTRAN-77. This compact manual provided step-by-step instructions, and demonstration with an example in an easy understanding.

Running the MRQFIT Program

1.) If you have the source code of MRQFIT program, you should first compile the MRQFIT program by Fortran-77 (see the compile procedures in the FORTRAN-77 manual). In the other hand, if you already have execute file, i.e. MRQFIT.out, you can run the program by typing MRQFIT.out in the UNIX prompt as:

```
[fon@atc7 ~/MRQFIT]$ MRQFIT.out
```

After pressing the ENTER key, the program start to run as:

```

=====
PROGRAM FOR FITTING
=====
MENU

1)  $-A/R^M + B/R^N + QiQj/R$ 
2)  $-A/R^M + B/R^N + QiQj/R + C/R^Q$ 
3)  $-A/R^M + B*EXP(-C*R) + QiQj/R$ 
4)  $-A/R^M + B*EXP(-C*R) + QiQj/R + D/R^Q$ 
5)  $-A/R^M + B*EXP(-C*R) + QiQj/R + D*EXP(-E*R)$ 
6)  $-A*EXP(-B*R) + C*EXP(-D*R) + QiQj/R$ 
7) Exit

```

2.) Select a choice of functional form from the list by pressing a number such as:

```

SELECT NUMBER OF CHOICE (1,2,3,...,7):
2

```

Note that the program should protect you to select unreasonable choice such as "0", "1.5", "8", etc.

3.) Select a power of denominator of each term such as:

```

PLEASE INPUT VALUE M, N AND Q OF THE FORM
M =
4
N =
7
Q =
12

```

In this step, you must wait for the program optimization. When the program finish, the output which containing three files are obtained (see Appendix II).

```
** NORMAL TERMINATION **  
STOP ** RUN COMPLETED ** statement executed
```



สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Appendix II

Output Files for MRQFIT Program

In this appendix, the example of output files in the NH₃-NH₃ system are shown as follow:

Output I

NUMBER OF ATOMS IN THE MOLECULES : 4 AND 4
NUMBER OF PARAMETERS : 9
NUMBER OF TRAJECTORIES : 21
NUMBER OF CONFIGURATIONS : 819
POINTS ABOVE 5.20 KCAL/MOL WILL BE EXCLUDED
DEGREES OF FREEDOM = 658
ITERATION LIMIT : 500 STEPS
FITTING STOP WHEN PARAMETERS CHANGE LESS THAN : 1.0000%

INITIAL PARAMETER : 99.7486877 269.9628300 2017.5252700
147.8139800 475.1542970 2221.8950200
68.1932144 251.7122340 2970.3664600

STEP	CHISQUARE	STD.	PARAMETERS
0	234.7740522	15.3223383	A 1 = 99.7486877 A 2 = 269.9628300 A 3 = 2017.5252700 A 4 = 147.8139800 A 5 = 475.1542970 A 6 = 2221.8950200 A 7 = 68.1932144 A 8 = 251.7122340 A 9 = 2970.3664600
1	0.0610427	0.2470682	A 1 = -360.4981060 A 2 = 14378.2098057 A 3 = 277997.9301762 A 4 = 99.3577428 A 5 = -455.0639601 A 6 = 11455.3803934 A 7 = 9.9939036 A 8 = 615.5520528 A 9 = -6972.9668208
2	0.0461584	0.2148451	A 1 = -413.0158797 A 2 = 20718.8714343 A 3 = -197955.5697540 A 4 = 167.7351744

			A 5 =	-655.7845477
			A 6 =	23978.3937696
			A 7 =	-31.4595134
			A 8 =	330.3055498
			A 9 =	-4699.7399432
3	0.0274538	0.1656919	A 1 =	-1578.9455580
			A 2 =	10314.9149489
			A 3 =	-365361.8286861
			A 4 =	472.6181664
			A 5 =	1871.0328990
			A 6 =	-176.8536728
			A 7 =	-116.2371518
			A 8 =	-354.3494743
			A 9 =	939.6388911
4	0.0194550	0.1394812	A 1 =	-3005.5489237
			A 2 =	-5298.0599896
			A 3 =	-270375.5392072
			A 4 =	807.7618914
			A 5 =	4805.7010683
			A 6 =	-30990.3440391
			A 7 =	-194.9866118
			A 8 =	-955.5774620
			A 9 =	5544.2037413
5	0.0193528	0.1391142	A 1 =	-3184.6297120
			A 2 =	-7273.3235584
			A 3 =	-256828.2254227
			A 4 =	849.6143362
			A 5 =	5173.0427059
			A 6 =	-34862.4426104
			A 7 =	-204.7276204
			A 8 =	-1029.6538041
			A 9 =	6108.5231306
6	0.0193527	0.1391141	A 1 =	-3186.8943773
			A 2 =	-7298.3114691
			A 3 =	-256655.9947517
			A 4 =	850.1434813
			A 5 =	5177.6875303
			A 6 =	-34911.4115434
			A 7 =	-204.8507240
			A 8 =	-1030.5897867
			A 9 =	6115.6517797

Maximum absolute residual = 0.4020830 kcal/mol
 Minimum absolute residual = 0.0017987 kcal/mol
 Average absolute residual = 0.0796535 kcal/mol

FINAL LAMBDA = 0.0000000010000001

FINAL PARAMETERS AND STATISTICS VALUES:

STEP	CHISQUARE	STD.	PARAMETERS
7	0.0193527	0.1391141	A 1 = -3186.8972441 A 2 = -7298.3430995 A 3 = -256655.7768786 A 4 = 850.1441511 A 5 = 5177.6934100 A 6 = -34911.4735306 A 7 = -204.8508799 A 8 = -1030.5909715 A 9 = 6115.6608036

Output II

NUMBER OF ATOMS IN THE MOLECULES : 4 AND 4
 NUMBER OF PARAMETERS : 9
 NUMBER OF TRAJECTORIES : 21
 NUMBER OF CONFIGURATIONS : 819
 POINTS ABOVE 5.20 KCAL/MOL WILL BE EXCLUDED
 DEGREES OF FREEDOM = 658
 ITERATION LIMIT : 500 STEPS
 FITTING STOP WHEN PARAMETERS CHANGE LESS THAN : 1.0000%

INITIAL PARAMETER :	99.7486877	269.9628300	2017.5252700
	147.8139800	475.1542970	2221.8950200
	68.1932144	251.7122340	2970.3664600

Maximum absolute residual = 0.4020830 kcal/mol
 Minimum absolute residual = 0.0017987 kcal/mol
 Average absolute residual = 0.0796535 kcal/mol

FINAL LAMBDA = 0.0000000010000001

FINAL PARAMETERS AND STATISTICS VALUES:

STEP	CHISQUARE	STD.	PARAMETERS
7	0.0193527	0.1391141	A 1 = -3186.8972441 A 2 = -7298.3430995 A 3 = -256655.7768786 A 4 = 850.1441511 A 5 = 5177.6934100 A 6 = -34911.4735306 A 7 = -204.8508799 A 8 = -1030.5909715 A 9 = 6115.6608036

Output III

NO.	YREAL	YFIT	RESIDUAL	R of CENTERS
1	34.559	34.559	0.0000	2.3000
2	18.728	18.728	0.0000	2.5000
3	16.026	16.026	0.0000	2.5500
4	13.695	13.695	0.0000	2.6000
5	11.684	11.684	0.0000	2.6500
6	9.9510	9.9510	0.0000	2.7000
7	8.4570	8.4570	0.0000	2.7500
8	7.1690	7.1690	0.0000	2.8000
9	6.0610	6.0610	0.0000	2.8500
10	5.1060	4.7897	0.31629	2.9000
11	4.2850	3.9788	0.30620	2.9500
12	3.5780	3.2575	0.32047	3.0000
13	2.9700	2.6159	0.35411	3.0500
14	2.4470	2.1446	0.30236	3.1000
15	1.9980	1.7354	0.26257	3.1500
16	1.6110	1.3807	0.23025	3.2000
17	1.2790	1.0739	0.20512	3.2500
18	0.99300	0.80888	0.18412	3.3000
19	0.74700	0.58048	0.16652	3.3500
20	0.53600	0.38403	0.15197	3.4000
21	0.35400	0.21545	0.13855	3.4500
22	0.19800	0.71163E-01	0.12684	3.5000
23	0.64000E-01	-0.51964E-01	0.11596	3.5500
24	-0.51000E-01	-0.15668	0.10568	3.6000
25	-0.14900	-0.24538	0.96376E-01	3.6500
26	-0.23300	-0.32015	0.87151E-01	3.7000
27	-0.30500	-0.38283	0.77832E-01	3.7500
28	-0.36600	-0.43501	0.69014E-01	3.8000
29	-0.41700	-0.47809	0.61088E-01	3.8500
30	-0.46000	-0.51327	0.53268E-01	3.9000
31	-0.49500	-0.54161	0.46612E-01	3.9500
32	-0.52500	-0.56404	0.39042E-01	4.0000
33	-0.54800	-0.58136	0.33362E-01	4.0500
34	-0.35300	-0.31273	-0.40270E-01	6.0500
35	-0.17000	-0.13340	-0.36595E-01	8.0500
36	-0.93000E-01	-0.67282E-01	-0.25718E-01	10.050
	:			
55	1.6110	1.3807	0.23025	3.2000
56	1.2790	1.0739	0.20512	3.2500
57	0.99300	0.80888	0.18412	3.3000
58	0.74700	0.58048	0.16652	3.3500
59	0.53600	0.38403	0.15197	3.4000
60	0.35400	0.21545	0.13855	3.4500
61	0.19800	0.71163E-01	0.12684	3.5000
62	0.64000E-01	-0.51964E-01	0.11596	3.5500
63	-0.51000E-01	-0.15668	0.10568	3.6000
64	-0.14900	-0.24538	0.96376E-01	3.6500
65	-0.23300	-0.32015	0.87151E-01	3.7000
66	-0.30500	-0.38283	0.77832E-01	3.7500
67	-0.36600	-0.43501	0.69014E-01	3.8000
68	-0.41700	-0.47809	0.61088E-01	3.8500
69	-0.46000	-0.51327	0.53268E-01	3.9000
70	-0.49500	-0.54161	0.46612E-01	3.9500
71	-0.52500	-0.56404	0.39042E-01	4.0000

72	-0.54800	-0.58136	0.33362E-01	4.0500
73	-0.35300	-0.31273	-0.40270E-01	6.0500
74	-0.17000	-0.13340	-0.36595E-01	8.0500
75	-0.93000E-01	-0.67282E-01	-0.25718E-01	10.050
	:			
274	33.729	33.729	0.0000	2.3000
275	16.749	16.749	0.0000	2.5000
276	13.924	13.924	0.0000	2.5500
277	11.515	11.515	0.0000	2.6000
278	9.4660	9.4660	0.0000	2.6500
279	7.7250	7.7250	0.0000	2.7000
280	6.2480	6.2480	0.0000	2.7500
281	4.9990	5.3734	-0.37441	2.8000
282	3.9430	4.3340	-0.39103	2.8500
283	3.0540	3.4235	-0.36948	2.9000
284	2.3070	2.6371	-0.33007	2.9500
285	1.6810	1.9654	-0.28444	3.0000
286	1.1570	1.3971	-0.24009	3.0500
287	0.72200	0.91998	-0.19798	3.1000
288	0.36200	0.52236	-0.16036	3.1500
289	0.64000E-01	0.19336	-0.12936	3.2000
290	-0.17900	-0.76896E-01	-0.10210	3.2500
291	-0.37700	-0.29714	-0.79858E-01	3.3000
292	-0.53700	-0.47504	-0.61963E-01	3.3500
293	-0.66500	-0.61722	-0.47778E-01	3.4000
294	-0.76500	-0.72940	-0.35602E-01	3.4500
295	-0.84300	-0.81645	-0.26553E-01	3.5000
296	-0.90200	-0.88253	-0.19472E-01	3.5500
297	-0.94500	-0.93116	-0.13836E-01	3.6000
298	-0.97500	-0.96534	-0.96596E-02	3.6500
299	-0.99400	-0.98757	-0.64258E-02	3.7000
300	-1.0040	-0.99999	-0.40139E-02	3.7500
301	-1.0070	-1.0044	-0.26391E-02	3.8000
302	-1.0040	-1.0022	-0.18052E-02	3.8500
303	-0.99700	-0.99474	-0.22559E-02	3.9000
304	-0.98500	-0.98306	-0.19412E-02	3.9500
305	-0.97000	-0.96802	-0.19833E-02	4.0000
306	-0.95300	-0.95035	-0.26502E-02	4.0500
307	-0.30700	-0.28070	-0.26301E-01	6.0500
308	-0.12700	-0.10353	-0.23465E-01	8.0500
309	-0.65000E-01	-0.48859E-01	-0.16141E-01	10.050
	:			
781	45.411	45.411	0.0000	2.3000
782	22.513	22.513	0.0000	2.5000
783	18.789	18.789	0.0000	2.5500
790	4.7280	4.4235	0.30449	2.9000
791	3.7850	3.8080	-0.22954E-01	2.9500
792	2.9990	3.1211	-0.12213	3.0000
	:			
811	-0.34000	-0.74065	0.40065	3.8000
812	-0.34400	-0.73506	0.39106	3.8500
813	-0.34300	-0.72368	0.38068	3.9000
814	-0.33800	-0.70795	0.36995	3.9500
815	-0.33000	-0.68901	0.35901	4.0000
816	-0.31900	-0.66778	0.34878	4.0500
817	-0.18000E-01	-0.10425	0.86249E-01	6.0500
818	0.70000E-02	-0.20816E-01	0.27816E-01	8.0500
819	0.70000E-02	-0.50074E-02	0.12007E-01	10.050

CURRICULUM VITAE

Nichayaporn Sangrawee

- 1974 Born November 24th, in Nakhon Sri Thammarat, Thailand
Father : Mr. Sunthorn Sangrawee
Mother : Mrs. Sunisa Sangrawee
- 1980-1986 Elementary School (Vad Pratat Temple School,
Nakhon Sri Thammarat)
- 1986-1989 Primary School (Kallayanee Sri Thammarat School,
Nakhon Sri Thammarat)
- 1989-1992 High School (Kallayanee Sri Thammarat School,
Nakhon Sri Thammarat)
- 1993-1997 Bachelor of Science (Applied Mathematics),
Thammasat University, Bangkok
- 1998-2001 Master degree student at Department of Mathematics,
Faculty of Science, Chulalongkorn University, Bangkok
- 1997-Present Lecturer at Department of Mathematics, Faculty of Science,
Thammasat University, Bangkok