First Principles Simulations: Development of New Density Functionals and Pseudopotentials and Formation Mechanism of Fullerenes

Thesis by

Xinlei Hua

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

California Institute of Technology
Pasadena, California

1996
(Submitted May 3, 1996)

Acknowledgements

First I would like to thank my advisor Bill Goddard for providing me with the opportunity to work in his group. Thanks to Bill for his encouragement and support during my rather long maturation process. Bill's unquenchable passion for science, near super-man's energy and the encyclopedic style is highly contagious. To study under his supervision is a rare fortune I will forever cherish. Bill is my role model of kind heart and an all-round cool guy.

Many thanks to Dr. Xiaojie Chen, under whose guidance I completed my first electronic structure code. Part of the thesis are from our collaboration. Thanks to Xiaojie for constant help and friendship. Thanks to Jean-Marc Langlois for his expert help in PS-GVB and FORTRAN debugging. Thanks to Ching-Hwa Kiang, whose work on buckytubes inspired me to pursue the story on the fullerene formation mechanism. So many people in the Goddard group have given me so much help. Naoki Karasawa was always very patient and clear in teaching me POLYGRAF and lattice dynamics. Thanks to our wonderful system manager Darryl Willick, who always has the right solution for my problem. K. T. Lim and Terry Coley were my ultimate source of computer consulting. Charles Musgrave set a high standard to emulate. Thanks to Charles for being high spirited. Thank Ersan Demiralp, Guanhua Chen and Yuejin Guo for friendship and sharing all the ups-and-downs of our intellectual quest. Siddharth Dasgupta was always available for questions of all sorts. Shaoyi Jiang has been an enthusiastic and inspiring colleague and friend. Every now and then Daqi Lu would come to my office and show me the latest wonders of computer software. Thanks to Guanhua Gao, Cecco Faglioni, Terumasa Yumaguchi, Changmoon Park, Erik Bierwagen, Matt Carlson and Mario Blanco, Sue Melnik and Ruth Ann Bertsch for making working in the group a pleasure. Thanks to our wonderful secretary Debbie Chester. Thanks to all the Goddard group.

My campus life is so rewarding that I developed a tendency of staying here forever.

Thanks to Yixin Liu for sharing all the dreams, from analog VLSI, McKinsey, to China's future. Thanks to my wonderful roommate Shaun Kirby, with whom I ventured into the summit of Mammoth and the white-water of Upper Kerns, and from who I learned so much about American intellectual culture. Thanks to the campus Chinese community, the Caltech C, for providing me with a warm cultural environment in which I work and play. Especially, thanks to Bing Ai, Dawei Dong, Bin Zhao, Yong Guo, Wei Min Lu, Taoling Xie, and the gang of our volley team. Thanks to my violin teacher Chris Baumler and chamber music instructor Delora Bing, who taught me tunes, rhythms and harmonies. Thanks to Newton and Roberta Chu for providing me with a host family here in Pasadena. These are the people that made my graduate life the happiest years of my life and for whom I feel deeply indebted.

Thanks to Prof. Jonnathan Machta and Prof. Robert Guyer for recommending me to study at Caltech. I admired their style of work and cherish their friendship.

Thanks for my parents for their unwavering faith in me. Thanks to them for cultivating the independence in me early on in my life. With my brother Xin-Feng I had so much fun together. Finally I would like to thank my dear wife Yuanyi. Her whole hearted support is everything to me and her love makes my life a joy on earth.

Abstract

This thesis consists of two parts. Part I deals with the development of first principles methodologies. Part II deals with applications of atomistic simulations, i.e. quantum mechanics and molecular dynamics simulations.

Part I includes two topics. One is generalized gradient correction for the density functional theory which constitutes Chapter 2. The other is first principle pseudopotentials which is covered in Chapter 3 and 4. Chapter 3 develops the density functional theory version used mostly in solid state physics, while Chapter 4 develops the Ab initio version used mostly in chemistry. Part II also includes two topics. One is the fullerenes formation mechanism which is covered in Chapter 5 and 6. Chapter 5 deals with their thermodynamical properties and Chapter 6 discusses their formation processes. The other is the lattice properties for the $YBa_2Cu_3O_7$ high-T$_c$ superconductor which is covered by Chapter 7. Also, in Chapter 1 we summarize the hierarchy models for materials simulations and review the state-of-the-art tools at various levels of that hierarchy.

Predicting the band gap from first principles has been hindered by the complexity of the systems and the flaws in the simplified theories. The most successful first principle theory, i.e. DFT, gives a value about the 2/3 of the measured one. [1] This is partly due to the approximate nature of the functionals used in DFT calculations. It has long been known that HOMOs calculated with LDA, the most commonly used DFT, is far from ionization potential. [2] In Chapter 2, after analyzing the nature of gradient corrected functional for DFT, we proposed a new exchange energy functional. The new functional is tested on several atoms and molecules and found to reproduce the Hartree-Fock eigenvalues to a good accuracy. With the incorporation of correlation energy in DFT, we can hope that the new functional would lead to a new and efficient

---

[1] Hartree-Fock, on the other hand, gives energy band gap twice as large as the measured one.

[2] It can be shown that the exact functional would give Koopman's theorem in DFT.

way of predicting energy band gaps for all the materials from first principles.

All chemistries involve mainly valence electrons. The effect of inner shells, shielding the nucleus Coulomb potential and providing Pauli repulsions to support the valence electrons, can be implemented by a potential called effective core potential(ECP). Since valence electrons of different angular momentum would experience different Pauli repulsion of the core electrons, this ECP is angular momentum dependent. This leads to a nonlocal potential and requires three-center integrals, which scales quadratically with the size of the system. On the other hand, the wave nature of electrons allows us to lower the resolution in describing the potential. By using a set of Gaussian functions to replace real-space grid in representing the ECP, we factorized the three-center integral into a sum of products of two-center integrals. We have found a set of Gaussian functions that gives accuracy of better than $0.01eV$ for all the elements in the periodic table, sufficient for all ECP calculations. The new method scales linearly with the system size. At 128 atoms, the cost is 1/15 of the old ECP method. This cleared a bottleneck for first principle programs that use ECP to study heavy elements.

Why C_{60} fullerenes are so stable and how this highly symmetric molecule is formed in the super-heated vapor is the two most fundamental questions in fullerene research. Prompted by the recent observations of the monocyclic ring and bycyclic rings as intermediates, we performed DFT calculations on the ring isomers and fullerene isomers of carbon clusters of various sizes. From it we extracted a force field (FF) for molecular dynamics simulations. This FF is used to calculate the free-energies at different temperatures. Based on our analysis of the physical forces that drive the carbon clustering and isomerization, we describe an evolution process for fullerene formation that is consistent with all the observations so far. With a combined DFT/MD method we are able to provide the energetics for a complete path of fullerene formation. Hint for synthesis improvement are suggested.

High-Tc superconductors are ceramics. To improve their mechanical and electrical properties, we conducted a molecular dynamics simulation for the $YBa_2Cu_3O_7$ superconductor. We derived a ionic-covalent force field from fitting the experimental data.

Our FF is able to reproduce the structure and Raman modes accurately. The FF predictions of isotope shifts of Raman frequencies, phonon dispersion spectra, phonon density of states, elastic stiffness constants, and volume thermal expansion are all in fairly good agreement with experiments.

Contents

I Sharpening the Simulation Tool 1

List of Figures

List of Tables

Part I

Sharpening the Simulation Tool

Chapter 1 Advances in Simulation Methods

1.1 Hierarchy of Models in Materials Simulation

Materials are used for their varieties of properties. Some are used for their strength on a macroscopic scale to build bridges. Other are used for their microscopic shapes and affinity for making drugs. The goal of material simulation is to map out the *property-structure relationship* of materials. For the diversity of properties we always use models, which correlate with the length scale of the properties involved, see Figure 1.1. With the advances in quantum chemistry and molecular dynamics, for the first time it

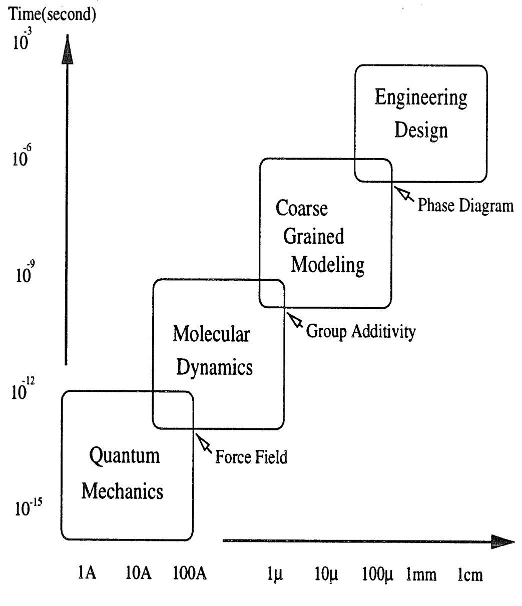


Figure 1.1: The hierarchy models for materials simulation.

is now possible to calculate from first principles properties at the various length and time scales.

In the hierarchy of models one level of theory provides the "reasons" and calculates parameters for the coarser grained theory. In turn itself needs a more microscopic theory to provide the underlying mechanism for its axiomatical parameters. For two theories in this ladder, we sometimes refer to the coarser one as the *phenomenological* theory and to the finer grain one as the *microscopic* one. The grain sizes correlate with the time, space and energy. At one end one has the finest grain for the small distance, short time and high energy, while at the other end, one has large distance, long time and low energy for the coarse grain.

The "fundamental" parameters in a theory of a certain grain size are determined either through measurements of that grain size or through the predictions from the finer grained theory. The former, top down approach, is inductive and involves statistical analysis; the later, bottom up approach, is deductive and involves dynamics.

At each level there are *correlation functions* that connect both sides. They are intrinsic properties for the upper theory but averaged properties in the lower one. The most prominent example is the equilibrium probability distribution of microscopic states. [1]

Both quantum physics and the classical physics have the property of scale invariance. Hierarchy models are not about the simple scale transformations. The hierarchy models deal with the exploding complexity that comes with the increase of length scales and time scales. It studies the emerging coarse grained variables and the laws governing them, e.g., how the behavior of the four valence electrons of carbon atom is related to the yield of C_{60} fullerenes in an arc.

In this section we will review some of the commonly used methods for moving up and down this ladder.

---

[1] Just like curvature is an intrinsic quantity characterizing the curved space without having to embed them in the space of higher dimension.

1.1.1 Coarse Graining in Quantum Mechanics

In quantum mechanics the coarse graining is usually done by taking the partial trace. The remaining ones would be the Hilbert space we are interested in. For example, Born-Oppenheimer approximation trace out the electronic degree of freedom and study the nucleus motion (by virtue of the disparate scales, visible light versus infra-red in energy and sub-femtosecond versus sub-picosecond in time). The emerging coarse grained variables is the potential energy surface (PES) for the nuclei. Another example is the commonly used one-particle Hamiltonian,

$$h[\{\phi_i\}] = Tr_{jkl\ldots}\langle\Psi_{ijkl\ldots}|\mathbf{H}|\Psi_{ijkl\ldots}\rangle, \tag{1.1}$$

with the Hamiltonian

$$\mathbf{H} = \frac{-\hbar^2}{2m}\sum_{i=1}^{n}\nabla_i^2 + \sum_{A=1}^{M}\sum_{i=1}^{n}\frac{Z_A e^2}{r_{iA}} + \sum_{i=1}^{n}\sum_{j>i}^{n}\frac{e^2}{r_{ij}}. \tag{1.2}$$

All but one of the degrees of freedom have been summed out, leading to a single-particle theory. The coarse grained variables $|\phi_i\rangle$ emerge from the fine grained many-particle wavefunction $|\Psi\rangle$.

In most cases partial trace is impractical, because we don't have $|\Psi\rangle$ to start with. Instead of performing the sum in (1.1) from $\langle\Psi_{ijkl\ldots}|\mathbf{H}|\Psi_{ijkl\ldots}\rangle$ one can model the $h[\{\phi\}]$ based on the known behaviors of ϕ. Examples are the BCS theory of superconductivity, [1] soliton theory of polyacetylyne, [2] etc..

Most of the coarse graining is done with modeling. Partial trace is the overall formalism but in practice it is usually performed within a given model.

1.1.2 Coarse Graining in Classical Physics

Standard formalism of coarse graining in classical mechanics is by virtue of center of mass theorem and in classical electrodynamics by multipole expansion.

- Center of mass theorem says that the motion of a large molecule obeys the

$\ddot{\mathbf{R}}_c = \mathbf{F}_c/M$ where $\mathbf{R}_c$ is the coordinate of the center of mass. The kinetic energies of the random motion of the atomic constituents is incorporated into M via E_k/c^2.

- For a point afar the electromagnetic field of a charge distribution $\rho(\mathbf{r})$ could be derived from the superposition of fields from a point charge, a dipole, a quadropole, an octopole, etc.. In so doing one can enlarge the length scale, from microscopic to mesoscopic.

Most of the coarse graining, however, are done with either mapping or modeling or averaging.

- In continuum elastics, by expand the displacement in the long wave limit, one can map the elastic constants, $C_{11}, C_{33}, C_{13}, ...$, etc. to the interatomic potential. [17]

$$\tilde{C}_{\alpha\beta,\gamma\lambda} = -\frac{1}{2V_a}\sum_k \phi_{\alpha\beta}(0|k) r_{k\gamma} r_{k\lambda}. \tag{1.3}$$

- Modeling the dielectric constant ϵ from atomic polarization α [4]

$$\begin{aligned} \epsilon &= 1 + 4\pi\chi & (1.4) \\ &= 1 + 4\pi\frac{N\alpha}{1-(4\pi/3)N\alpha}. & (1.5) \end{aligned}$$

1.1.3 Statistical Physics — *from Jaynes, Landau to Wilson*

A microscopically random state corresponds to the macroscopically homogeneous state. Low symmetry at one resolution leads to high symmetry phase of a coarser resolution.

Jaynes' Information Theory of Statistical Mechanics

Assuming that the thermodynamic entropy and information-theory entropy to be the same concept, Jaynes in 1957 showed that statistical mechanics can be derived from the maximum entropy principle. [5]

Suppose all we know is the expectation value of the function $f(x)$:

$$\langle f(x)\rangle = \sum_{i=1}^{n} p_i f(x_i), \tag{1.6}$$

where p_i is the unknown probability of having value x_i subjecting to the constraint

$$\sum p_i = 1. \tag{1.7}$$

On the basis of this information, what is the expectation value of the function $g(x)$, i.e. $\langle g(x)\rangle$? The most unbiased, nonsubjective estimate would be to maximize our "ignorance". Shannon showed [6] that there is unambiguous criterion for representing the amount of uncertainty. He proofed that the quantity which is positive, which increases with increasing uncertainty, and is additive for independent sources of uncertainty, is

$$S(p_1...p_n) = -K\sum_i p_i ln p_i, \tag{1.8}$$

where K is a positive constant. This is just the expression for entropy as found in statistical mechanics. Jaynes thus proposed the following:

> *In making inferences on the basis of partial information we must use that probability distribution which has maximum entropy subject to whatever is known.*

Using Lagrangian multipliers for the constraint maximization, one gets

$$p_i = e^{-\lambda-\mu f(x_i)}. \tag{1.9}$$

The constants λ and μ are determined by substituting into the constraints (8.57) and (8.77),

$$\langle f(x)\rangle = -\frac{\partial}{\partial\mu} lnZ(\mu), \tag{1.10}$$

$$\lambda = lnZ(\mu), \tag{1.11}$$

where

$$Z(\mu) = \sum_i e^{-\mu f(x_i)}. \tag{1.12}$$

Given information about the state of the system, if

$$p_i = e^{-\lambda - \mu f(x_i)} \tag{1.13}$$

becomes exceedingly small, then one can view $\langle f(x_i) \rangle$ as a constraint and build models based on such condition. This is another formal way of saying how you can focus on a smaller block of Hilbert space, i.e., a formal foundation of model building.

The Hierarchy of Coarse Graining

Suppose a system is described by the Hamiltonian H. Eigenstates are characterized by $\{\Psi_n\}$, and eigenvalues E_n. The partition function is

$$Z = \sum_n e^{-\beta E_n}. \tag{1.14}$$

If at one level there appears a $\phi(\mathbf{r})$ that is discernible and thus could be used as a state variable then we say that we need a new parameter to describe the system (we don't need it if there is no discernible pattern on that resolution). Landau called this kind of parameter the order parameter, since it is when the order appears that the symmetry breaks requiring a new variable for describing it.

From the grain level below it, when order parameter appears, we need to do coarse graining with the constraint $\{\phi(\mathbf{r})\}$. Thus

$$Z = \sum_{\{\phi(\mathbf{r})\}} \sum_{n \in \{\phi(\mathbf{r})\}} e^{-\beta E_n}, \tag{1.15}$$

where n is the most microscopic state index. $\phi(\mathbf{r})$ is the macroscopic state description, since **r** is macroscopic coordinate (macroscopically small but microscopically large). $n \in \{\phi(\mathbf{r})\}$ denotes the constraint in summing over the microscopic states n.

One can apply this constrained summation to obtain a hierarchy of summations, from the most microscopic states to the macroscopic state:

$$Z(\gamma) = \sum_{\beta\in\gamma}\sum_{\alpha\in\beta}\sum_{n\in\alpha} e^{-\beta E_n} \tag{1.16}$$

$$Z(\gamma) = \sum_{\beta\in\gamma} Z(\beta;\gamma) \equiv \sum_{\beta} e^{-ln(Z(\beta;\gamma))} \tag{1.17}$$

$$Z(\beta) = \sum_{\alpha\in\beta} Z(\alpha;\beta) \equiv \sum_{\alpha} e^{-ln(Z(\alpha;\beta))} \tag{1.18}$$

....

$Z(\gamma), Z(\beta)$, etc. are sometimes called the partial trace. Defining free energy $F(\beta)$ by:

$$\lambda(\gamma)F(\beta) = ln(Z(\beta;\gamma)). \tag{1.19}$$

We have:

$$Z(\gamma) = \sum_{\beta\in\gamma} e^{-\lambda(\gamma)F(\beta)}, \tag{1.20}$$

$$Z(\beta) = \sum_{\alpha\in\beta} e^{-\lambda(\beta)F(\alpha)}. \tag{1.21}$$

The expressions of the free energy depend on the level of coarse graining. For instance, $F(\beta) = (1/\lambda)ln(Z(\beta))$, $F(\alpha) = H(\alpha)$, etc.. If $\gamma = (T, V, N)$ $\beta = \phi(\mathbf{q})$ and $\alpha =$

$\{p_i, q_i\}$, then we have:

$$Z(T,V,N) = \int_{\{\phi(\mathbf{q})\}\in(T,V,N)} D\phi(\mathbf{q})e^{-\lambda(T,V,N)F[\phi(\mathbf{q})]} \tag{1.22}$$

with

$$F[\phi(\mathbf{q})] = ln \int_{\{\mathbf{p}_i,\mathbf{q}_i\}\in\phi(\mathbf{q})} \Pi_i d\mathbf{p}_i d\mathbf{q}_i e^{-\lambda(\phi(\mathbf{q}))H(\{\mathbf{p}_i,\mathbf{q}_i\})}. \tag{1.23}$$

Landau-Ginzburg Theory

Landau proposed a phenomenological theory. Instead of calculating the $F[\phi(\mathbf{q})]$ in (1.23) from a finer grain (e.g., from $H(\{\mathbf{p}_i, \mathbf{q}_i\})$), he modeled, with some symmetry consideration, the $F[\phi(\mathbf{q})]$ with polynomials,

$$F[\phi] = a\phi^2 + b\phi^4 + c|\nabla\phi|^2 + ..., \tag{1.24}$$

Equations (1.22) and (1.24) constitute an empirical method for studying the phenomenon on the length scale of $\{\mathbf{q}\}$.

One can get the equilibrium distribution $\phi_0(\mathbf{q})$ by the *most probable* method. The largest contribution in the integral in (1.22) comes from the $\phi(\mathbf{q})$ that minimize the $F[\phi(\mathbf{q})]$. This $\phi_0(\mathbf{q})$ is the equilibrium state and it can be derived via solving the Euler-Lagrangian equation,

$$\frac{\delta F[\phi(\mathbf{q})]}{\delta\phi(\mathbf{q})}|_{\phi_0} = 0. \tag{1.25}$$

This is the famous Ginzburg-Landau theory.

Wilson's Renormalization Group Theory

Yet another way of doing coarse graining is by scale transformation. This method is appropriate for situations where scaling law is everything, e.g., near the critical point of a continuum phase transition.

Consider a spin system described by the Hamiltonian

$$H = K_0 \sum_{i<j} \mathbf{S}_i \mathbf{S}_j. \tag{1.26}$$

One way of doing coarse graining is through the decimation of the variables $\{\mathbf{S}_i\}$ upon scaling transformation. After one iteration of decimation(i.e., partial trace within block and cast the resulting Hamiltonian $H^{(1)}$ into the same form as H), the spin variables becomes $\{\mathbf{S}_i^{(1)}\}$ and the coupling constant becomes $K^{(1)}$. [2] The relationships among the coupling constants $\{K^{(i)}\}$ of different grain levels,

$$K^{(i+1)} = f^{(i)}(K^{(i)}) \tag{1.27}$$

describes the coarse graining.

Under the cutoff approximation(see footnote), the same f holds between any adjacient level of coarse graining, i.e.,

$$f^{(i)} = f^{(i-1)} = ... = f^{(0)}. \tag{1.28}$$

Thus the coarse graining process is reduced to a recursion relation. The sequence of $f^{(i)}$ is called renormalization group and is denoted by $\mathbf{R}_L$ [3]

At critical points, the *value* of coupling constant K should be independent of the level of coarse graining. One can thus identify the K with the fixed point of renormalization transformation $\mathbf{R}_L$, and find it by solving

$$K^* = \mathbf{R}_L K^*. \tag{1.29}$$

Equation (1.27) constitutes the method of doing coarse graining near the critical

---

[2] During decimation, extra interaction can arise. For example, the next nearest neighbor interaction, which is absent in the original Hamiltonian, can occur. Thus one need to do cutoff in casting the coarse grained Hamiltonian into the same form as the physical Hamiltonian.

[3] $\{\mathbf{R}_L\}$ is actually a semigroup since there is no inverse operation.

points. It is very powerful to treat a class of problems that have the property of self-similarity, e.g., calculation of the critical exponents. It was invented by Kadanoff, Fisher and Wilson in the 70s to study the phase transition and critical phenomenon. [27] However, equation (1.27) is not valid when the system is away from the critical point. When the system is away from the critical point, there is no self-similarity. Then the characteristic length scale is the guide line to model the coarse graining.

1.2 Review of the First Principles Methods

In this chapter we review the quantum chemistry method.

Quantum mechanics is the fundamental law governing all the material properties. Solving quantum mechanics equations for the system containing a large number of interacting particles constitute the subject of the first principle method. Over the years, the algorithms, both physically based and pure numerical, coupled with the development in computer hardware technology, has advanced to a stage that it can study the property of close to a hundred atom system. Advances in speed allows us to do larger systems, while advances in accuracy allow us to raise the predictive power (higher resolution). The current quantum chemistry (QC) methods grow into two branches which have different formal roots. One is Hartree-Fock based method (HF, RHF and UHF, MP2, MP4, CI, MCSCF, GVB, etc.. [20] The other is the density function theory (DFT), an economic tool that recently acquired significant improvement in accuracy.

The Schödinger equation

$$H|\Psi\rangle = E|\Psi\rangle \tag{1.30}$$

with the omnipotent Hamiltonian

$$H = \frac{-\hbar^2}{2m}\sum_{i=1}^{n}\nabla_i^2 + \sum_{A=1}^{M}\sum_{i=1}^{n}\frac{Z_A e^2}{r_{iA}} + \sum_{i=1}^{n}\sum_{j>i}^{n}\frac{e^2}{r_{ij}} \tag{1.31}$$

and the antisymmetric property of $|\Psi\rangle$ for the indistinguishable fermions, encode all the chemistry and physical property of materials.

$N-$electron wave function $|\Psi\rangle$ can be expanded in the space spanned by basis as,

$$|\Psi\rangle = \sum_{\{\phi_1,\phi_2,...,\phi_n\}} C_{12,...,N}|\phi_1\phi_2\phi_3....\phi_N\rangle. \tag{1.32}$$

Antisymmetric property of $|\Psi\rangle$ is implemented as

$$|\phi_1\phi_2\phi_3....\phi_N\rangle = \sum_{P\in S_N} sgn(P)|\phi_1\rangle \otimes |\phi_2\rangle \otimes ... \otimes |\phi_N\rangle, \tag{1.33}$$

where S_n is the symmetric group of order N, $|\phi_i\rangle$ are single particle states. Using its eigenvalue to denote the single-particle state, $|\phi_i\rangle$ could be $|\mathbf{r}\rangle$, $|\mathbf{k}\rangle$, or any $|n\rangle$, etc.. For example, if $|\phi_i\rangle = |\mathbf{r}\rangle$ we have $C_{12,...,N} = \Psi(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_n)$ in (1.32).

1.2.1 Hartree-Fock Method

When the electron correlation is not strong, we can form a single-particle picture. Namely, we can find a set of optimal single-particle $\{|\phi_i\rangle\}$, such that

$$|\Psi\rangle \approx |\phi_1\phi_2....\phi_N\rangle. \tag{1.34}$$

Hartree-Fock wavefunction $|\Psi^{HF}\rangle$ is defined as the best such wavefunction in the sense that it gives the lowest ground state energy $E_0 = \langle\Psi|\mathbf{H}|\Psi\rangle$.

Putting

$$\Phi(\mathbf{r}_1\mathbf{r}_2....\mathbf{r}_N) = \langle\mathbf{r}_1\mathbf{r}_2....\mathbf{r}_N|\phi_1\phi_2\phi_3....\phi_N\rangle = \frac{1}{\sqrt{N!}}\begin{vmatrix} \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) & .... & \phi_N(\mathbf{x}_1) \\ \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) & .... & \phi_N(\mathbf{x}_2) \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ \phi_1(\mathbf{x}_N) & \phi_2(\mathbf{x}_N) & .... & \phi_N(\mathbf{x}_N) \end{vmatrix} \tag{1.35}$$

into the above expectation value of $\mathbf{H}$, expressing each orbital with a set of L functions (called basis set) $\{\chi_\mu\}$,

$$\phi_i = \sum_{\mu}^{L} C_{i\mu}\chi_\mu, \tag{1.36}$$

and doing variation with respect to $\{C_{i\mu}\}$ under the constraint,

$$\langle \Psi^{HF} | \Psi^{HF} \rangle = 1, \tag{1.37}$$

leads to an Euler-type of equation:

$$\sum_{\nu} H_{\mu\nu} C_{\nu n} = E_n \sum_{\nu} S_{\mu\nu} C_{\nu n}. \tag{1.38}$$

Or

$$\mathbf{FC} = \mathbf{ESC} \tag{1.39}$$

for short. $\mathbf{E}$ is a diagonal matrix of eigenvalues. The equation 1.39 is called the Roothaan equation. If we choose $|\chi_\mu\rangle$ to be continuous basis $|\mathbf{r}\rangle$, the sum $\sum_\mu$ becomes integral $\int d\mathbf{r}$, the variation changes from $\partial/\partial C_{n\mu}$ to $\delta/\delta\phi_n(\mathbf{r})$, and we have the differential-integral equation formalism,

$$h(\mathbf{x}_1)\phi_n(\mathbf{x}_1) + \sum_{m\neq n} \Big[\int d\mathbf{x}_2 \frac{|\phi_m(\mathbf{x}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \Big] \phi_n(\mathbf{x}_1) - \sum_{m\neq n} \Big[\int d\mathbf{x}_2 \frac{\phi_m(\mathbf{x}_2)^* \phi_n(\mathbf{x}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \Big] \phi_m(\mathbf{x}_1) = \epsilon_n \phi_n(\mathbf{x}_1) \tag{1.40}$$

Fock operator

$$F(1) = h(1) + \sum_{m} J_m(1) - K_m(1) \tag{1.41}$$

is the sum of a one-particle operator $h(1)$ and an effective one-electron potential operator called Hartree-Fock potential $v^{HF}(1) = \sum_m J_m(1) - K_m(1)$

$$\mathbf{F}|\phi_n\rangle = \epsilon_n|\phi_n\rangle. \tag{1.42}$$

This is the famous Hartree-Fock equation. It is the workhorse of first principle method.

Of course QM is independent of the representation used to describe it. In molecular study Roothaan equation is usually solved.

To form the Fock matrix $\mathbf{F}$ We need to evaluate these kind of matrix elements:

$$t_{ij} = \int d\mathbf{r}\chi_i(\mathbf{r} - \mathbf{R}_A)\nabla^2\chi_j(\mathbf{r} - \mathbf{R}_B) \tag{1.43}$$

and

$$V_{ijkl} = \int\int d\mathbf{r}d\mathbf{r}'\chi_i(\mathbf{r} - \mathbf{R}_A)\chi_j(\mathbf{r} - \mathbf{R}_B)\frac{1}{|\mathbf{r} - \mathbf{r}'|}\chi_k(\mathbf{r} - \mathbf{R}_C)\chi_l(\mathbf{r} - \mathbf{R}_D). \tag{1.44}$$

Since $|\chi_i\rangle$ are not orthogonal to each other we have the overlapping matrix

$$S_{ij} = \int d\mathbf{r}\chi_i(\mathbf{r} - \mathbf{R}_A)\chi_j(\mathbf{r} - \mathbf{R}_B). \tag{1.45}$$

Roothaan equation can be solved by symmetrization and diagonalization.

$$(\mathbf{S}^{\frac{1}{2}}\mathbf{F}\mathbf{S}^{-\frac{1}{2}})(\mathbf{S}^{\frac{1}{2}}\mathbf{C}_{\mathrm{i}}) = \epsilon_i(\mathbf{S}^{\frac{1}{2}}\mathbf{C}_{\mathrm{i}}). \tag{1.46}$$

For the purpose of developing the perturbation methods we ask the following question: Is there an approximate Hamiltonian for which $|\Psi\rangle$ is an exact eigenvector? The answer is yes:

$$H_0 = \sum_i^N f(i). \tag{1.47}$$

Then

$$H = H_0 + V \tag{1.48}$$

with

$$V = \sum_{i>j}^{N} \frac{1}{r_{ij}} - \sum_{i=1}^{N} v^{HF}(i). \tag{1.49}$$

The Hartree-Fock energy

$$\begin{aligned} E_0 &= \langle \Psi^{HF} | H_0 | \Psi^{HF} \rangle + \langle \Psi^{HF} | V | \Psi^{HF} \rangle & (1.50) \\ &= \sum_n \epsilon_n + \langle \Psi^{HF} | V | \Psi^{HF} \rangle & (1.51) \\ &= E_0^{(0)} + E_0^{(1)}. & (1.52) \end{aligned}$$

So the zeroth order is the eigenvalue sum and the first order is the normal elimination of the double counting.

1.2.2 Beyond Hartree-Fock —Perturbation Schemes

Rayleigh-Schrödinger scheme

$$\langle \Psi_0^{(0)} + \lambda_1 \Psi_0^{(1)} + ...| H_0 + V | \Psi_0^{(0)} + \lambda_1 \Psi_0^{(1)} .... \rangle. \tag{1.53}$$

First split Hamiltonian into H_0 and V with H_0 having the properties of (a) It covers majority of the H and (b) Its eigenfunction is known or easy to solve. Then use the eigenvectors of H_0 to form a space in which to correct for the residue V. For example, if H_0 is taken as Hartree-Fock Hamiltonian (equation refeq:hfh), then ϕ_0 would be the Hartree-Fock wavefunction.

$$\psi_0^{(1)} = \sum_n{}' \frac{\langle \psi_n^{(0)} | V | \psi_0^{(0)} \rangle}{E_n^{(0)} - E_0^{(0)}} \psi_n^{(0)}, \tag{1.54}$$

where $\psi_n^{(0)}$ are Slater determinants made from combination of orbitals, occupied as well as virtual, and V is from equation 1.49. The E_{tot} can be expressed in terms of matrix elements of Fock operator, an expression first derived by Moset-Ploset and are now called MP2 theory for the second order formula.

Feynman-Dyson Scheme

$$\langle\Psi_0|H|\Psi_0\rangle = \langle\Phi_0|U(-\infty,0)HU(0,-\infty)|\Phi_0\rangle \tag{1.55}$$

start from asymptotically turning on the interaction term, analogous to scattering problem. At $t \to -\infty$ the eigenfunction is $\Phi_0(\mathbf{r})$. At $t = 0$ it is $|\Psi_0\rangle$. Using Wick's theorem and the property that the $[\hat{\phi}_\alpha, \hat{\phi}_\beta] = iG_{\alpha\beta}$ is a $C-$number, this leads to the Feynman-Dyson perturbation expansion, which is relativistic invariant.

The Schrödinger scheme leads to MP2, MP3, MP4, etc. and is normally used to calculate the total energy of ground state. The Feynman scheme leads to the Dyson equation and is normally used for studying the excitation spectrum.

1.2.3 Density Functional Theory

Hohenberg and Kohn noticed that at the nondegenerate ground state there is one energy and one electron density. [7]

$$E_0[\rho_0] = \int \rho_0(\mathbf{r})v(\mathbf{r})d\mathbf{r} + F_0[\rho_0], \tag{1.56}$$

where the functional $F_0[\rho_0]$ is a universal functional of the *electron system*, independent of external field v. Since $F_0[\rho_0]$ is universal, people had hoped of finding THE $F_0[\rho_0]$, that, upon entering the ρ_0, gives the correct ground state energy. But the real work is how to find this ρ_0, the ground state density. Unlike the Lagrangian $L(\phi, \dot{\phi}, t)$, whose Euler-Lagrangian equation is used to find the behavior of the dynamical variable $\phi_i(t)$, the uniqueness of $F_0[\rho_0]$ by definition is a property of the ground-state itself.

Levy [8] realized this and constructed an operational definition for $F[\rho]$ using an operator which involves only electron system, $\mathbf{T} + \mathbf{V}_{ee}$,

$$F[\rho] \equiv Min_{\Psi\to\rho}\langle\Psi|\mathbf{T} + \mathbf{V}_{ee}|\Psi\rangle. \tag{1.57}$$

From now on we will understand the functionals $F[\rho]$, $T[\rho]$, $V[\rho]$ as the constraint search definition of Levy.

Kohn and Sham [11] proposed mapping the unknown $F[\rho]$ to a fictitious independent electron system having the same density ρ. Then one can take the ride with this independent electron system to the ground state ρ_0 and $F_0[\rho_0]$. Since at the ground state there is only one ρ_0 and one E_0, we reach our destination. For this independent electron system(IES),

$$T^{ind}[\rho] = \sum_i^N \langle \phi_i | -\frac{1}{2}\nabla^2 | \phi_i \rangle \tag{1.58}$$

with

$$\rho(\mathbf{r}) = \sum_i^N f_i |\phi_i(\mathbf{r})|^2, \tag{1.59}$$

where f_i is the occupation number. Coulomb potential energy is the normal one,

$$J[\rho] = \frac{1}{2}\int d\mathbf{r}d\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{1.60}$$

They then put the residue kinetic energy due to correlation, exchange energy and correlation potential energy into one term called exchange-correlation energy functional:

$$E_{xc}[\rho] = T[\rho] - T^{ind}[\rho] + V_{ee}[\rho] - J[\rho]. \tag{1.61}$$

Kohn-Sham then model this $E_{xc}[\rho]$ by the expression of homogeneous electron gas having the same density ρ. So the final functional is

$$E[\rho] = T^{ind}[\rho] + J[\rho] + E_{xc}[\rho] + E_{ext}[\rho]. \tag{1.62}$$

This functional leads to an Euler-Lagrangian equation

$$\left[-\frac{1}{2}\nabla^2 + v_{ext} + \int d\mathbf{r}\frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{r}'|} + v_{xc}(\mathbf{r})\right]\psi_n = \epsilon_n\psi_n \tag{1.63}$$

with

$$v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})} \tag{1.64}$$

and

$$\rho(\mathbf{r}) = \sum_{i}^{N} f_i |\phi_i(\mathbf{r})|^2, \tag{1.65}$$

which need to be solved self-consistently. This is the famous Kohn-Sham scheme. The total energy E_{tot} is

$$E_{tot}[\rho] = \sum_n \epsilon_n - \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[\rho] - \int d\mathbf{r} v_{xc}(\mathbf{r}) \rho(\mathbf{r}). \tag{1.66}$$

Levy's operational definition of $F[\rho]$ has no practical meaning. To get the exact $E_{xc}[\rho]$ we resort to perturbation theory. From the last section on perturbation we have,

$$\langle \Psi_0 | H | \Psi_0 \rangle = \langle \Phi_0 | U(-\infty, 0) H U(0, -\infty) | \Phi_0 \rangle. \tag{1.67}$$

One can view DFT as an empirical way of modeling

$$H^{eff} = U(-\infty, 0) H U(0, -\infty) \tag{1.68}$$

which involves the self-consistently incorporation of $\rho(\mathbf{r})$ from $|\Phi_0\rangle$

In turn, perturbation provides the formal link between $E_{xc}[\rho]$ and H. The insight of the origin of E_{xc} can lead to improvement over the current functionals. Improvement of E_{xc} will be the subject of Chapter **2**.

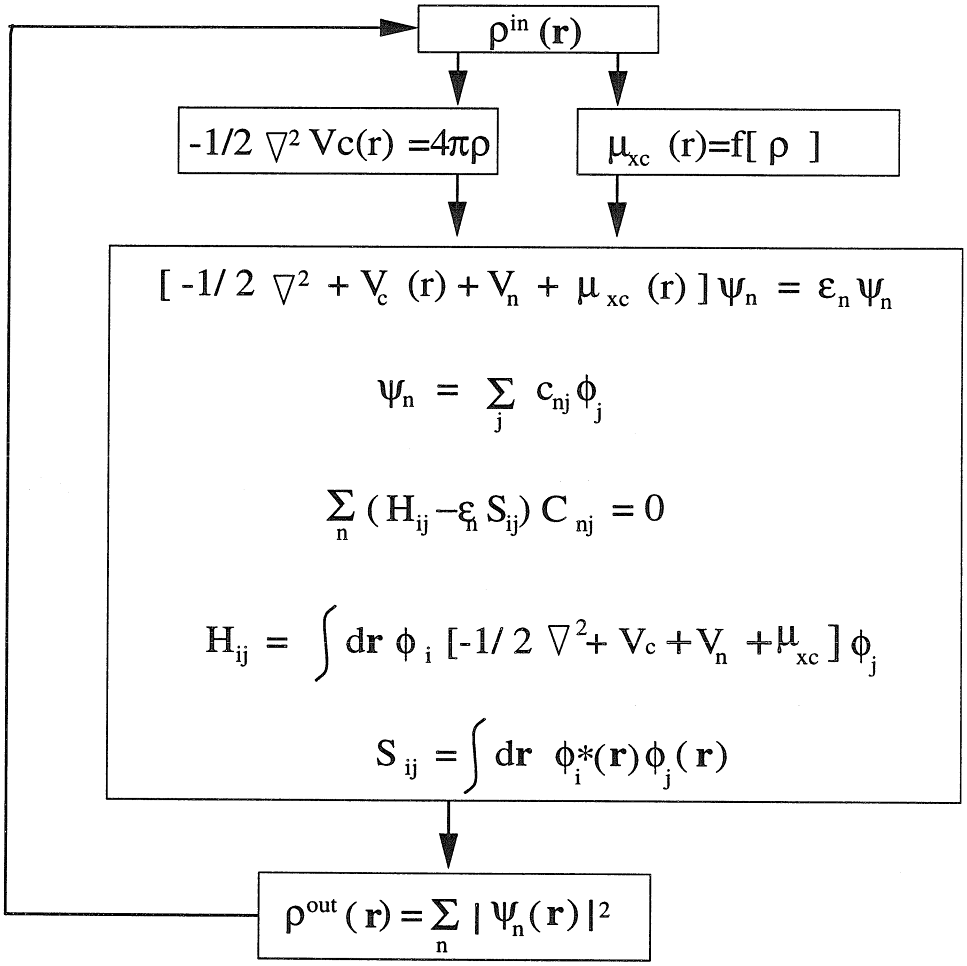


Figure 1.2: The self-consistent solution of the Kohn-Sham equation

Potentail	Basis Sets	Core Electrons	$E_{xc}[\rho]$
periodic	plane-waves	all electron	LDA
non-periodic	numerical grid	pseudopotential	X_α
	APW[a]		VWN
	Slater-type orbitals		BLYP
	Gaussian-type orb.		B3LYP
	Gaussian-plane waves		GGA

[a] Augmented Plane-waves.

Wavefunction update	Geometry update
Conjugate Gradient	Analytical derivative
DIIS	numerical derivative
Linear Response(C-P)	Car Parrinello
Harris Functional	

Table 1.1: The implementation choices of the DFT calculation by each part.

1.2.4 Implementations

SCF Procedure

Implementation Choices

The advances in implementations of each part of the above procedure have enlarged the power of the first principles machinery. Some of these are based on physics of the electron systems, some on numerical algorithms.

In table 1.1 we listed several choices for each part of a DFT computation program.

For the content of Table 1.1, $E_{xc}[\rho]$ will be the subject of Chapter **2**; Pseudopotential will be discussed in Chapter **3** and Chapter **4**. Wavefunction updating and geometry updating will be covered in Appendix. So here we just briefly discuss the boundary condition and the basis sets, i.e., the optimal representation.

Representation choices

First, symmetry allows us to block diagonalize the Fock operator, thus significantly cuts down the size of the matrix.

system	symmetry group	representation
Atom	K_h	s,p,d,f,....,
Molecule	Point group	A_1,B_2,E_1, ...,
Crystal	Space group	Γ,L,Δ,....,

Take the crystal for example. To a good approximation a single electron in crystal experiences an averaged periodic potential. We can impose the translation symmetry to simplify calculations. [14]

Plane waves $|\mathbf{k}\rangle$ has the periodic boundary condition built in. $|\mathbf{k}\rangle$ is the index of the irreducible representation of space group. Using the *Bloch* function as basis,

$$|\chi_\alpha(\mathbf{k})\rangle = \frac{1}{\sqrt{N}} \sum_i e^{i\mathbf{k}\cdot\mathbf{R}_i} |\chi_\alpha(\mathbf{R}_i)\rangle, \tag{1.69}$$

the orbitals in crystal is

$$|\phi_{n\mathbf{k}}\rangle = \sum_\alpha^L u_{n\alpha} |\chi_\alpha(\mathbf{k})\rangle. \tag{1.70}$$

Minimizing $\langle\phi_\mathbf{k}|H|\phi_\mathbf{k}\rangle/\langle\phi_\mathbf{k}|\phi_\mathbf{k}\rangle$ leads to secular equation,

$$\sum_\beta^L H_{\alpha\beta}(\mathbf{k}) u_{n\beta} = E_{n\mathbf{k}} u_{n\alpha} \tag{1.71}$$

with the matrix element

$$H_{\alpha\beta}(\mathbf{k}) = \frac{1}{N} \sum_{ij} e^{i\mathbf{k}\cdot(\mathbf{R}_i - \mathbf{R}_j)} \langle\chi_\alpha(\mathbf{R}_i)|H|\chi_\beta(\mathbf{R}_j)\rangle, \tag{1.72}$$

which is of the size of the basis set employed for the unit cell. Thus, knowledge of the translational symmetry reduces the size of the Hamiltonian matrix from $10^{23} \times 10^{23}$ to $L \times L$.

Within the irreducible representation of the system's symmetry group, one still has choices over which basis set to use. Gaussian functions $x^m y^n z^q e^{-\alpha r^2}$ have been the choices for molecules, plane-waves $e^{-i\mathbf{G}\cdot\mathbf{r}}$ and some augementation have been the choices for crystals, while numerical grid $\mathbf{r}$ has been used only for atoms. Recent pseudo-spectra technology and dual-space Gaussian technology for crystals employed the optimal combinations of these three representations. [15, 15] These optimizations are based on the physical pictures of atoms in crystals or atoms in molecules. From a numerical algorithm point of view, different types of basis functions correspond to different numerical grids, just like the one used in finite element for fluid mechanics. The optimal grid is always problem dependent, requiring the foresight of what the systems should be like. [4]

Car-Parrinello's *Ab initio* Molecular Dynamics

Electrons move at a speed three orders of magnitude faster than ions, they can follow ionic movement almost instantaneously. Electron *orbitals*, therefore, change at the same rate as ions except at some transition regions. When ions are standing still orbitals do not change since *orbital indices* are good quantum numbers of the *electronic* Hamiltonian under BO approximation. When Car and Parrinello noticed that electronic *orbitals* are dynamic variables of the time scale of ions, [5] they designed a dynamics whereby one propagates electronic orbitals alongside the ions, instead of doing self-consistent calculation to get them. [15] The elimination of self-consistent loop significantly cut down the cost, allowing the molecular dynamics to be performed with the ab initio PES, which now has a finite thickness. However, the stability of such dynamics depends on the step size $\delta\mathbf{R}$. In a section of appendix we will discuss the nature of such dynamics. We will show that the time step in the Car-Parrinello scheme is limited by the extent of validity of linear response.

---

[4] In a broader sense, the grid optimization is similar to choosing the sample space in Monte Carlo, designing the wiring of a neural net, finding eigenvectors, etc..(among these symmetry is the obvious first step.) They constitute the very procedure of modeling.

[5] electron can respond to external field with high frequency, now this external field is the field produced by $\{\delta\mathbf{R}_i\}$

1.3 Review of the Molecular Dynamics Method

All molecular dynamics (MD) and Monte Carlo (MC) methods take PES as its input. How to get this PES constitutes the subject of the next section. In this section we shall focus on the MD procedure itself.

1.3.1 NTP Dynamics

All MD assumes ergodicity. To simulate finite temperature properties one must generate configurations according to the (N,V,T) ensemble. This can be achieved by incorporating a larger system as heat bath.

By introducing variables to represent the degree of freedom of the heat bath and other external constraints, one can build an extended system. The Hamiltonian corresponding to this extended system still generate a microcanonical ensemble. One can build the extended system in such a way that trajectory average of the extended system is formally equivalent to the average over the trajectories of the original Hamiltonian H_0 *weighted by the corresponding Boltzmann factor.*

(N, P, H) dynamics

[23, 17] In

$$L_0(\mathbf{r}, \dot{\mathbf{r}}) = \frac{1}{2} m \sum_{i=1}^{N} \dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_i - \sum_{i<j=1}^{N} u(r_{ij}), \tag{1.73}$$

replacing the coordinates $\mathbf{r}_i$ by the scaled coordinates $\rho_i = \mathbf{r}_i / V^{1/3}$, and considering the following Lagrangian,

$$L(\rho, \dot{\rho}, Q, \dot{Q}) = \frac{1}{2} m Q^{2/3} \sum_{i=1}^{N} \dot{\rho}_i \cdot \dot{\rho}_i - \sum_{i<j=1}^{N} u(Q^{1/3} \rho_{ij}) + \frac{1}{2} M \dot{Q}^2 - \alpha Q \tag{1.74}$$

leads to the Hamiltonian form:

$$\pi_i = \frac{\partial L}{\partial \dot{\rho}_i}, \tag{1.75}$$

$$\Pi_i = \frac{\partial L}{\partial \dot{Q}}, \tag{1.76}$$

$$H(\rho^N, \pi^N, Q, \Pi) = (2mQ^{2/3})^{-1}\sum_{i=1}^{N}\pi_i \cdot \pi_i + \sum_{i<j=1}^{N} u(Q^{1/3}\rho_{ij}) + (2M)^{-1}\Pi^2 + \alpha Q \tag{1.77}$$

The dynamics of L generates the trajectories $\{\rho(t), \pi(t), Q(t), \Pi(t)\}$. Anderson showed [6] that average along *this* trajectory, which is the (N,V,E) average of the scaled system, for any function $F(\mathbf{r}^N, \mathbf{p}^N, V)$, is formally equivalent to (N, P, H) average of F along the trajectory of the original L_0.

(N, T, P) dynamics

Nosé defined the virtual variables $(\mathbf{q}_i, \mathbf{p}_i, s, V, t)$ as follows, [5]

$$\mathbf{q}' = V^{1/3}\mathbf{q}_i, \tag{1.78}$$

$$\mathbf{p}' = \frac{\mathbf{p}_i}{V^{1/3}s}, \tag{1.79}$$

$$\mathbf{t}' = \int^t dt/s, \tag{1.80}$$

where $(\mathbf{q'}_i, \mathbf{p'}_i, s, V, t')$ are real variables. Write the Hamiltonian in terms of virtual variables as,

$$H = \sum_i \frac{\mathbf{p}_i^2}{2m_i V^{2/3}s^2} + \phi(V^{1/3}\mathbf{q}) + \frac{p_s^2}{2Q} + gkTlns + \frac{p_V^2}{2W} + P_{ex}V \tag{1.81}$$

Take the *virtual* variable as canonical variables, the Hamiltonian equations

$$\frac{d\mathbf{q}_i}{dt} = \frac{\partial H}{\partial \mathbf{p}_i} = \frac{\mathbf{p}_i}{m_i V^{2/3}s^2}, \tag{1.82}$$

---

[6]

$$\bar{F} = C\int d\rho^N d\pi^N \int dQ d\Pi \;\; \delta[H(\rho^N, \pi^N, Q, \Pi) - E]F(Q^{1/3}\rho^N, \pi^N/Q^{1/3}; Q)$$

$$F_{NPH} = C\int dV \int d\mathbf{r}^N d\mathbf{p}^N \;\; \delta[H(\mathbf{r}^N, \mathbf{p}^N; V) + PV - H]F(\mathbf{r}^N, \mathbf{p}^N; V)$$

$$\bar{F} = F_{NPH}.$$

$$\ldots . \tag{1.83}$$

leads to a set of dynamical equations in terms of virtual variables. Replacing the virtuals back by the real variable leads to the dynamics equation for the real variables, which produce trajectories in real variables. Again, using the standard coarse graining method of integrating out the variables in the partition function we have:

$$\begin{aligned} Z &= \int dp_V dV dp_s ds \int d\mathbf{p} d\mathbf{q} \delta(H-E) \\ &= \int dp_V dV dp_s ds \int d\mathbf{p} d\mathbf{q} \delta[H_0(\mathbf{p'}_i, \mathbf{q'}_i) + p_s^2/(2Q) + gkTlns + p_V^2/(2W) + P_{ex}V - E)] \\ &= C \int dV \int d\mathbf{p'} d\mathbf{q'} exp\left[-\left(\frac{3N+1}{g}\right)(H_0(\mathbf{p'}_i, \mathbf{q'}_i) + P_{ex}V - E)\right]. \end{aligned} \tag{1.84}$$

So, for virtual time sampling, with g=3N+1, the equilibrium distribution function is

$$\rho(\mathbf{p'}, \mathbf{q'}, V) = exp\{-[H_0(\mathbf{p'}, \mathbf{q'}) + P_{ex}V]/kT\}. \tag{1.85}$$

Because of the momentum and angular momentum conservation, the ensembles produced by the above MD method are slightly different from the usual statistical mechanical ensembles.

Whether Nosé dynamics is just a way of generating canonical ensembles or whether it also represents a physically meaningful dynamics of the system was investigated by K. Cho and J. D. Joannopoulos, [20] who showed that the thermostat parameters Q in Equation(1.74) can be chosen to obtain the real physical dynamics.

In Chapter **6** we will use the (N,T,H) dynamics to estimate the attempting frequency for carbon cluster reactions. In Chapter **7** we will use it to calculate the thermal expansions.

Correlation Functions

Correlation functions tell us *a postari* how the microscopic variable should be averaged to give macroscopic transport properties. They are the bridges between macro and micro world. On the one hand, they represent the intrinsic properties of the *macroscopic* system, and on the other hand they can be determined from averaging the *microscopic* configurations. For quantum mechanical many-electron system, correlation function of various observables are directly related to two-particle Green's functions, which boil down to the expectation value of creation and annihilation operators and which can be calculated through Feynman-Dyson perturbation. For classical mechanics the average is done over the trajectories. They are not determined directly from FF, rather they are from the trajectories generated by FF.[7] Now that we have an efficient way of generating ensembles, we can calculate the correlation functions at any temperature. Furthermore, as contrast to the Monte Carlo method, MD allows one to calculate the temporal correlation function. The relationship between correlation and the linear response properties was derived by Kubo in 1959. [19] For example, the electrical conductivity is related to the current-current correlation.

$$\mathbf{j}(\mathbf{r}) = \frac{ie\hbar}{2m}[\psi^+(\mathbf{r})\nabla\psi(\mathbf{r}) - \nabla\psi^+(\mathbf{r})\psi(\mathbf{r})] - \frac{e^2}{m}\mathbf{A}\psi^+(\mathbf{r})\psi(\mathbf{r}) \tag{1.86}$$

$$\sigma_{\alpha\beta}(\mathbf{q},\omega) = \frac{iNe^2}{m\omega}\delta_{\alpha\beta} + \frac{1}{\hbar\omega}\int_{-\infty}^{0} dt\langle\Psi|[j_\alpha(\mathbf{q},0), j_\beta(-\mathbf{q},t)]|\Psi\rangle e^{-i\omega t}. \tag{1.87}$$

Examples in classical physics: diffusion constant is related to autocorrelation of velocities $\{\mathbf{v}_i\}$.

$$D = \int_0^\infty dt < v(0)v(t) > . \tag{1.88}$$

Therefore one can calculate any linear response property by running MD and calculating correlation functions, which amounts to averaging over trajectories of dynamics variables.

---

[7] Just like Green's function are not from Hamiltonian itself but from the expectation value of $\psi^+\psi$ on many-body ground state wavefunction determined by Hamiltonian.

1.3.2 Implementation

The advances in implementing the molecular dynamics allows the simulation to reach a million atoms. The new capacity makes simulations on nanotechnology and large biological(proteins, viruses) systems possible.

The most computation costly terms is evaluating Coulomb energy summation. It scaled as N^2. Recently advances in implementation has made Coulomb sum a N process.

Ewald

The slow decaying of $[1/R_{ij}]^n$ requires very large cut offs, making direct summation slow to converge. The local **R** description hit the local details but misses the long wave-length phenomenon. On the other hand, Fourier **G**-space representation hits the long wave phenomenon, but misses the short wave phenomenon by the finite cut off G_{cut}. As a way to calculate the Coulomb interaction in lattice, Ewald [21] in 1921 found a way to combine the **R**-space for the local and **G**-space for the long wave. He separated the sum into two parts, one of which converged rapidly in real space, and the other converged rapidly in reciprocal space (Fourier transforms).

$$\sum_{i,j,L} \frac{Q_i Q_j}{R_{ijL}} = \frac{1}{2} \sum_{i,j,L} \frac{Q_i Q_j}{R_{ijL}} erfc\left(\frac{1}{\eta} R_{ij}\right) + \frac{1}{2} \sum_{i,j,L} \frac{Q_i Q_j}{R_{ijL}} \left[1 - erfc\left(\frac{1}{\eta} R_{ij}\right)\right]. \tag{1.89}$$

The first term converged very fast in real space, the second term converged very slowly in real space. However Fourier transforming the second term leads to

$$E_k = \frac{2\pi}{\Omega_0} \sum_h \frac{1}{h^2} e^{-\eta^2 h^2/4} \left\{ \sum_{i,j} Q_i Q_j cos[\mathbf{h} \cdot (\mathbf{R}_i - \mathbf{R}_j)] \right\} - \frac{1}{\eta\sqrt{\pi}} \sum_i Q_i^2 \tag{1.90}$$

which can converge fast using appropriate η.

Extending the idea to quantum mechanics calculations leads to the dual-space technology, GDS/DFT. [15]

CMM

Cell multipole method (CMM) employs the idea of multipole expansion to calculate the Coulomb energy of a large cluster of charge. The field which an atom experiences, $V(\mathbf{R}_i)$, can be divided into two parts, contribution from nearby atoms and the contributions from the distant atoms. For $V_{far}(\mathbf{R}_i)$ we don't need to evaluate in terms of the individual atoms. We can group them together and count their contribution by their multipole expansion. The atoms further away can be grouped into larger groups.

$$V(\mathbf{R}_i) = V_{far}(\mathbf{R}_i) + V_{near}(\mathbf{R}_i) \tag{1.91}$$

$$V_{near}(\mathbf{R}_i) = \sum_{j,near} Q_j R_{ij}^{-1} \tag{1.92}$$

$$V_{far}(\mathbf{R}_i) = \sum_{A,far} V_A^{pole}(\mathbf{R}_i - \mathbf{R}_A). \tag{1.93}$$

This algorithm scale as N compared with the $(1/2)N^2$ scaling of the exact method. On starburst dendrimers (β-alanine) it is 2377 faster than the exact method. One can increase the accuracy of CMM by including more atoms into the near field and including higher multipoles in the far field expansion. Typically including multipoles up to quadrupole (CMM/2) leads to rms force errors of 0.4 to 0.2 (kcal/mol)/Å [22]

The extension of this idea to PBC leads to reduced cell multipole method, or RCMM. The computation cost for Ewald scales as

$$T = C_{real} R_{cut}^3 N + C_{recip} k_{cut}^3 N^2 \tag{1.94}$$

where N is the number of atoms in the unit cell, R_{cut} and k_{cut} are cutoffs for the sums in real and reciprocal spaces, and C_{real} and C_{recip} are constants. [23] For large systems, e.g., million atoms per unit cell, this is not practical. RCMMM, the crystal version of CMM, cuts the cost to scale N.

In molecular dynamics simulation of $YBa_2Cu_3O_7$ systems(Chapter 7), we will use the Ewald method.

There are also major advances in solving Newton equations. In studying large systems like macro-molecules, there are very fast mode and very slow mode. Averaging out the fast mode when studying the slow mode, or freezing the slow mode when studying the fast mode lead to the NEIMO dynamics. [8] Since we won't be using it in this thesis we will just give the reference. [24]

1.4 Linking QM with MD–Force Field

All MD requires the potential energy surface (PES) as input, with it MD outputs structure properties related to ions (crystal structures, heat of formation, melting temperature, defect energetic, surface reconstructions, phonon dispersion curves, elastic constant, bulk and share modulus; phonon density of state, and phonon contribution of specific heat, thermal-conductivity, thermal expansions, etc.).

Since PESs differ greatly from system to system, having many scales of roughness and many local minima, it is very difficult to come up with a universal functional that applies to all the properties of all the systems.

For bindings mainly due to the electrostatic forces, there are successful force field[9] for pure ionic crystals, e.g., NaCl. For bindings due to covalent bond there are valence force field for the pure covalent crystals, e.g., diamond, C_{60}. For other materials the force field could be complicated. In general we need to find the balance between two extremes, accurate but specific targeting, and general but approximate.

The FF is normally modeled by some functional of atomic position and by fitting the corresponding parameters to the experimental data or ab initio calculations.

Smart choice of the functional form of the FF, which may not be unique, is crucial to its accuracy and transferability, and thus the predictive power. Physically the question is "to what type of approximate solution of the electronic Schrödinger equation

---

[8] Classical analogous of Born-Oppenheimer approximation.

[9] Here we use force field, potential energy surface, interatomic potential, interchangeably

does the energy functional correspond?" [26] Often, the empirical fitting schemes tend to reproduce well the properties they fit with and do poorly for other properties. For instance, FFs fitted to Hessians would not give accurate prediction of the thermal-conductivity, since the latter is due to anharmonicity, a realm beyond the harmonic approximation. Another example is the defect energetics, where a drastic change in local electron density makes the FF around the defect quite distinctive from the bulk ones. So in doing MD simulations one needs to constantly be aware of the power and limitation of the force field one uses. We will give detailed examples in Chapter **7**.

1.4.1 Covalent and Ionic FF

The binding between atoms can be categorized into reducing kinetic energy and reducing the electrostatic energy(including exchange). The former makes covalent bond, which is directional, short ranged, and has a fixed number of coordination, while the later makes ionic bond, which is pure attraction, is long ranged, isotropic, no saturation. Correspondingly one can divide force field terms into bonding and nonbonding terms.

$$E_{ff} = E_{val} + E_{nb} \tag{1.95}$$

$$E_{val} = E_{bond} + E_{angle} + E_{torsion} + E_{inv} \tag{1.96}$$

$$E_{nb} = E_{vdw} + E_Q \tag{1.97}$$

Nonbonding Force Field

Potential Derived Charges and Charge Equilibrium [25]

Using point charge to represent electron cloud is only well defined for the far field. Far field criterion gives the definition of the potential derived charge(PDC).

For large systems, calculating PDC at each geometries is too expensive and does not serve the purpose for cases like calculating polarizations and running dynamics. Charge equilibrium method(QEq) is an efficient, physics based method. The central idea is to treat charge as a gas that flows between spherical atoms of radius $\{R_A\}$

until the chemical potentials are equal(like the grand canonical ensemble). The charge free energy E_Q is

$$E_Q = \sum_A E_A(Q_A) + \sum_{A>B} Q_A Q_B J_{AB}(R_{AB}) \tag{1.98}$$

where

$$E_A(Q_A) = E_0 + \chi_A^0 Q_A + \frac{1}{2} J_{AA}^0 Q_A^2, \tag{1.99}$$

with $\chi^0 = (IP_A + EA_A)/2$ and and $J_{AA}^0 = (IP_A - EA_A)$ determined by the atomic ionization potential(IP) and electron affinity(EA). The function $J_{AB}(R_{AB})$ is the shielded Coulomb interaction between the two atoms.

$$\frac{\delta E_Q}{\delta Q} = 0 \tag{1.100}$$

leads to a set of linear equation for the unique equilibrium charges $\{Q_A\}$ at a given geometry. QEq is successful in organics systems and semiconductor systems. But for ceramics like $YBa_2Cu_3O_7$ system we found that Qeq gives charges too small(about half for Y and Ba).

Bonding Terms

Functional form

Bond stretch,

$$E_{stretch} = \frac{1}{2} k_r (R - R_0)^2 \tag{1.101}$$

Bond angle,

$$E_{angle} = \frac{1}{2} k_\theta \left(cos(\theta) - cos(\theta_0)\right)^2 \tag{1.102}$$

Torsion, Inversion, etc, are also standard and available in POLYGRAF manual.

Force field parameters

The parameters, e.g., $k_r, k_\theta, R_0, \theta_0$, etc., are tuned against either *ab initio* Hessians or experimental vibration frequencies. Once determined they are used for predicting new properties.

1.4.2 Metallic Force Field

The delocalized electrons in metals make them quite different from either ionic or covalent bonded crystals. Upon bonding, the change of electron densities are centered on the interstitial regions, [31] rather than on atoms as in ionic crystals or at the bond mid point as in covalent bonded crystals.

Based on this picture Li and WAG proposed interstitial electron model (IEM). [30] In IEM electrons are represented by classical particles. Pair wise potential between these interstices and ions are constructed. The parameters in the potentials are fitted to experiments. Explicit inclusion of the valence electrons in the interstitial region achieves the anisotropic description without using angular force.

Another force field commonly used for metal are based on the tight binding theory of electrons. For example, the following form is used with success in studying the transition metals Zr(zirconium) [29]

$$E_c = A \sum_{r_{ij}<r_c} exp\left[-p\left(\frac{r_{ij}}{r_0} - 1\right)\right] - \sqrt{\lambda^2 \sum_{r_{ij}<r_c} exp\left[-2q\left(\frac{r_{ij}}{r_0} - 1\right)\right]}. \tag{1.103}$$

The first term is the Born-Mayer type repulsion term. The second term is the second-moment approximation of the tight-binding band energy.

Other forms of force field include embedded-atom-method [10], etc.

$$E = \sum_i F_i\left(\sum_{j\neq i} f_j(r_{ij})\right) + \frac{1}{2}\sum_{ij}\phi_{ij}(r_{ij}). \tag{1.104}$$

The first terms takes up the change in kinetic energy which is determined by the elec-

---

[10] Effective medium theory, embedded atom method, pari-functional methods all refer to similar method.

tron density. The function $f(R)$ is the electron density at a site due to the neighboring atom. F is the energy associated with placing an atom in that electron environment. The second term takes up the potential energy. $\{\phi_{ij}\}$ represent the Coulomb or van der Waals interaction.

We see the force field varies greatly from system to system. This is the case because of the phenomenological nature of the force field. One need first to identify the system and then choose the force field accordingly. Deriving an accurate FF is very important, for it underlies the atomistic simulations. As a tool it allows us to study large systems efficiently. This tool is getting sharper and sharper. We will illustrate the functionality of force field in the Part **II** of this thesis.

Semi-empirical Theories

Yet another class of approaches in trading accuracy for speed are the semi-empirical (SE) methods. The semi-empirical methods use the first principle functional and make universal approximations. By neglecting some integrals(INDO,MNDO), setting up some cutoffs, reinterpreting certain terms, of the *Ab initio* method etc., the SE method can save cost substantially. Examples are Harris functional, [27] tight-binding total energy calculation, [28] etc. But the justification for the SE methods lies *a postari* on their performance, and they can't be tailored individually, the predictive power of SE methods, thus, is limited. On the other hand, one can envision a hierarchy of SE methods to bridge the *Ab initio* and MD method, to manage the different level of accuracy of an PES efficiently.

1.5 Summary

There has long been the dream that theory(quantum mechanics, molecular dynamics and statistical mechanics) properly incorporated into computer software could be used to design new drugs, new chemicals, and new materials. The advances outlined in this chapter foretell the enormous potential opportunities for applications to problems previously unattainable. With the advances in each level of the hierarchy models, the range of each grain increases, and the linking between them becomes more accurate and more efficient. We have reason to hope that, with the advances in hardware technologies, the long held dream will be realized.

Bibliography

[1] J. Bardeen, L. N. Cooper, J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957)

[2] W. P. Su, J. R. Schrieffer, A. J. Heeger, Phys. Rev. Lett. **42**, 1698 (1979); Phys. Rev. B **22**, 2099 (1980)

[3] P. Brüesch and W. Bührer, Z. für Physik B-Condensed Matter 70, 1-8 (1988)

[4] E. M. Purcell, Electricity and Magnetism *Berkeley Physics Course Vol. 2* (1965)

[5] E. T. Jaynes "Information Theory and Statistical Mechanics". Phys. Rev. **106** 620,(1957)

[6] C. E. Shannon, Bell System Tech. J. **27**, 379, 623 (1948) reprinted in C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (University of Illinois Press, Urbana, 1949)

[7] See References in L. E. Reichl *A Modern Course in Statistical Physics* (1980) Chapter 10.

[8] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry*, McGraw-Hill, (1989)

[9] P. Hohenberg and W. Kohn, Phys. Rev. **136**:B864-B871 (1964)

[10] M. Levy, Proc. Natl. Acad. Sci. USA **76**: 6062-6065. (1979)

[11] W. Kohn, L. J. Sham, Phys. Rev. **140**, A1133 (1965)

[12] X. J. Chen, J-M. Langlois, and W. A. Goddard III, Phys. Rev. B, **52**, 2348, (1995)

[13] R. A. Friesner, Chem. Phys. Lett. **116**, 39 (1985); and reference in PS-GVB USER's Guide, Schrödinger, Inc. (1996)

[14] W. Harrison, *Electronic Structure and the Properties of Solids* Freeman and Company (1980)

[15] R. Car and M. Parrinello, Phy. Rev. Lett. **55** 2471, (1985)

[16] Hans C. Anderson, J. Chem. Phys. **72** 2384 (1980)

[17] M. Parrinello and A. Rahman J. Appl. Phys.**52** 7182 (1981)

[18] S. Nosé, J. Chem. Phys. **81** 511 (1984)

[19] R. Kubo, *Lectures in Theoretical Physics*, Vol. I(Boulder) (Wiley-Interscience, New York, 1959), pp. 120-203; J. Phys. Soc. Japan **12** 570 (1957)

[20] K. Cho and J. D. Joannopoulos, Phys. Rev. A **45** 7089 (1992)

[21] P. P. Ewald, Ann. Phys. (Lpz.) (4) **64**, 253 (1921)

[22] W. A. Goddard, MSC Bi-annual Report March, 1992 — March, 1994

[23] H. Q. Ding, N. Karasawa, and W. A. Goddard, J. Chem. Phys. **97**, 4309 (1992)

[24] A. Jain, N. Vaidehi, and G. Rodriquez, J. Comp. Phys. **106**, 258 (1993)

[25] A. K. Rappé and W. A. Goddard III, J. Phys. Chem. **95**, 3358 (1991).

[26] A. E. Carlsson, *Solid State Physics* Vol. 43, p9,1990 Academic Press.

[27] J. Harris, Phys. Rev. B **31** 1770 (1985); W. M. C. Foulkes and R. Haydock, Phys. Rev. B **39** 12520, (1989)

[28] G. B. Adams, O. F. Sankey, J. B. Page and M. O'Keeffe, Chem. Phys. **176**, 61, (1993)

[29] F. Willaime and C. Massobrio, Physical Review letter, **63** 2244 (1989)

[30] M. Li and W. A. Goddard III, J. Chem. Phys. **98**, 7995 (1993); Physical Review B, **40** 12155 (1989);

[31] M. H. McAdon and W. A. Goddard III, J. Chem. Phys. **91**, 2607 (1987); Phys. Rev. Lett., **55** 2563 (1985).

Chapter 2 Generalized Gradient Approximation of DFT

2.1 Introduction

2.1.1 Overview of DFT

Thomas and Fermi [1] independently proposed a density theory in 1927 to treat the the many-electron systems. Thomas-Fermi(TF) theory picturizes the many-electron system as being made up of many regions of locally homogeneous electrons gas. The total energy, E^{TF}, is the integral of the local energy densities $\epsilon(\mathbf{r})$. TF theory used the uniform electron gas expression for this $\epsilon(\mathbf{r})$. Then variational principle is used to find the ground-state density ρ_0. After Hartree-Fock equation is proposed (Hartree in 1928 and Fock in 1930 [2]) the orbitals have become the dominant description of the many-electron systems. Because of the nonlocal nature of exchange operator, HF equation is a differential-integral equation. Slater(1951) [3] proposed replacing the exchange operator by a local one. He *modeled* the exchange potential after the homogeneous electron gas and introduced an adjustable parameter α, giving the name X_α method. The local potential leads to a differential equation. In 1965 Hohenberg, Kohn and Sham developed the density functional theory(DFT) that not only put Slater's method on a formal basis but also allows one to incorporate the correlation energy term $E_c[\rho]$. Kohn-Sham scheme is very successful and has been very popular in solid-state physics.

With the advances in computational methods, [15] the non-locality of exchange-energy matrix is no longer the most important issue. On the other hand, the developments of gradient corrected methods [5] elevate the accuracy of DFT to that of MP2 [1]. Now DFT is used mainly as an efficient method to incorporate correlation.

For crystal calculations DFT is the only cost-effective choice. For molecules calculation, the HF based *Ab initio* method with the incorporation of correlations are in many places replaced by DFT, due to the laters economic way of incorporating the correlations.

DFT is now a practical tool to study both the structure of matter and electronic structure, both in crystal and in molecules.

2.1.2 Hohenberg-Kohn-Sham Theory

Before going into the details of the standard Kohn-Sham DFT theory and the subsequent gradient corrected form we would like to point out some subtle flaws of its formulation. Our new understandings lead us to the new functionals which for the first time restored the Koopman's theorem in DFT. This constitutes the major step toward improving the accuracy of DFT in predicting such properties like band gap and other excitation spectrum.

Hoherberg and Kohn showed that at the ground state there is a unique density ρ_0 and that ground state energy E_0 is a unique functional of this ρ_0. [7] They "proved" it by comparing such ground state energy E_0 with that due to a different external potential v'_{ext}. They had assumed that, except for some pathological distributions, all the distribution $\rho(\mathbf{r})$ can be realized by some external potential $v_{ext}(\mathbf{r})$. Levy [8] pointed out this is not true and the condition under which this holds constitutes the subject of $N-representability$. [1]

We will show that what HK have proved is something trivial, just by definition it has to hold. While what they used implicitly is not proved. For an N-electron system, v_{ext} defines the system. Obviously, given a v_{ext} there is only one $\rho(\mathbf{r})$ at ground state, since Ψ_0 determines ρ_0. And, being an observable, ρ_0 has to be unique. [1] Therefore *at* the ground state the mapping $E_0[\rho_0]$ from the density distribution $\rho_0(\mathbf{r})$ to a real number E_0 has to be one-to-one.

---

[1] since there is no unitary transformations to play around, as in the case of orbitals.

A Lesson From Equilibrium Thermodynamics

In thermodynamical equilibrium, (μ,T,P) determines the state of a system. At equilibrium, the variation w.r.t. N leads to another thermodynamical equilibrium quantity chemical potential μ: $\partial G/\partial N = \mu(N,T,P)$. On the other hand the search for the equilibrium, $\delta G = 0$, requires variation w.r.t. the finer grained variable which has meaning away from equilibrium state, i.e. the probability distribution of microscopic quantities, $\rho(H(\{\mathbf{q},\mathbf{p}\}))$. Thus $\delta G = 0$, while assumed conceptually in thermodynamics, has to be carried out with a finer grained theory, in which the thermal equilibrium "states" is just one kind of averaging for the microscopic states. In the fine grained theory one can assign the probability distribution to each fine grained states, and from which construct the entropy, $s = -k_B \sum_i \rho_i ln \rho_i$. Find the probability distribution that maximize the entropy and identify this probability with thermodynamical equilibrium "states". This way one can get the *equilibrium probability distribution*, $\rho_0(\{\mathbf{q}_i,\mathbf{p}_i\}) = e^{-(E(\{\mathbf{q}_i,\mathbf{p}_i\})-\mu N)/k_B T}$, and the equilibrium free-energy $\overline{G}(\mu,N,T) = -k_B T ln(\sum_{\{\mathbf{q}_i,\mathbf{p}_i\}} e^{-(E(\{\mathbf{q}_i,\mathbf{p}_i\})-\mu N)/k_B T})$. We noticed that for the equilibrium states only energy $E(\{\mathbf{q}_i,\mathbf{p}_i\}) - \mu N$ and the density of states matter, detailed microscopic description $\{\mathbf{q}_i,\mathbf{p}_i\}$ doesn't.

In steady state one can extend the thermodynamical equilibrium to have spatial dependent $\mu(\mathbf{r})$, where $\mathbf{r}$ is a macroscopic position vector representing a microscopically large block of space.

In deriving Hartree-Fock equation(see Chapter 1) the variation is w.r.t. $\{\phi_i\}$. In DFT, on the other hand, the variation is w.r.t. ρ, which is a massive coarse graining of the orbitals, $\rho = \sum_i |\phi_i|^2$.

A Lesson from Variational Calculus

Near the ground state, one can expand the unknown energy functional $E[\rho]$ around the ground state density ρ_0.

$$E_x[\rho] = E_x^0[\rho_0] + \int \frac{\delta E}{\delta \rho} |_{\rho_0} \delta\rho d\mathbf{r} + \frac{1}{2} \int \int \frac{\delta^2 E}{\delta\rho(\mathbf{r})\delta\rho(\mathbf{r}')} |_{\rho_0} \delta\rho(\mathbf{r})\delta\rho(\mathbf{r}') d\mathbf{r} d\mathbf{r}' + O((\delta\rho)^3)$$

(2.1)

2

Note that the mapping from the ground state electron density ρ_0 to ground state energy E_0 could be totally different from the assumed mapping between the variable ρ and energy functional E. Clearly, using

$$\frac{\delta E_0[\rho_0]}{\delta\rho_0(\mathbf{r})} = 0 \tag{2.2}$$

to find ρ_0 is not legal. So unlike the Lagrangian or Hamiltonian, which is used to derive the ground state, the ground state energy functional $E_0[\rho_0]$ should not be used for deriving the Euler-Lagrangian type of dynamical equations.

Hohenberg and Kohn [7] assumed implicitly that $E[\rho]$ and $E_0[\rho_0]$ are of the same functional, at least near ρ_0. In their argument they are assumed that all density distributions can be realized by an external potential v_{ext}, a property called v-representable. Levy provided a more rigorous arguments, [8] enlarging the domain of the possible $\rho(\mathbf{r})$ from v-representable to N-representable[3]. Here we shall illustrate that ρ_0 as constants of density-space should be incorporated in $E[\rho]$ as parameters, i.e. $E[\rho;\rho_0]$.

For instance, in homogeneous electron systems, $\rho_0 = N/V$, with particle number N and volume V being the global constraints. Then the exchange-energy functional should be $E_x[\rho; N, V] = E_x[\rho, \rho_0]$. Thus for different Vs there are a bundle of $E_x[\rho]$s. The ground state *value* $E_{x0}[\rho_0; N, V]$ for each V fall onto a curve. It is this curve that LDA has assumed to be $E_x[\rho]$, see Figure 2.1.

In this chapter, we will first follow Kohn-Sham scheme and show how it works for various systems. Then we shall introduce a new functional, based on our new understanding of the DFT. We shall show how the new functional improves the calculations of the eigenvalues(orbital energies), a problem that has muffled the power of DFT for

---

[2] As discussed in the appendix, the word functional, variation, and the square bracket $[\rho]$, are nothing but the extension of many-component function $F(\phi_1, \phi_2, \phi_3, ...)$ to the continuum. (A manifold) This continuity allows one to define derivatives of any order, thus enlarge the space from $\{\phi(\mathbf{r})\}$ to $\{\phi(\mathbf{r}), \nabla\phi(\mathbf{r}), \nabla^2\phi(\mathbf{r}), ...\}$.

[3] $\rho(\mathbf{r})$s that correspond to all possible many-particle wavefunction Ψs.

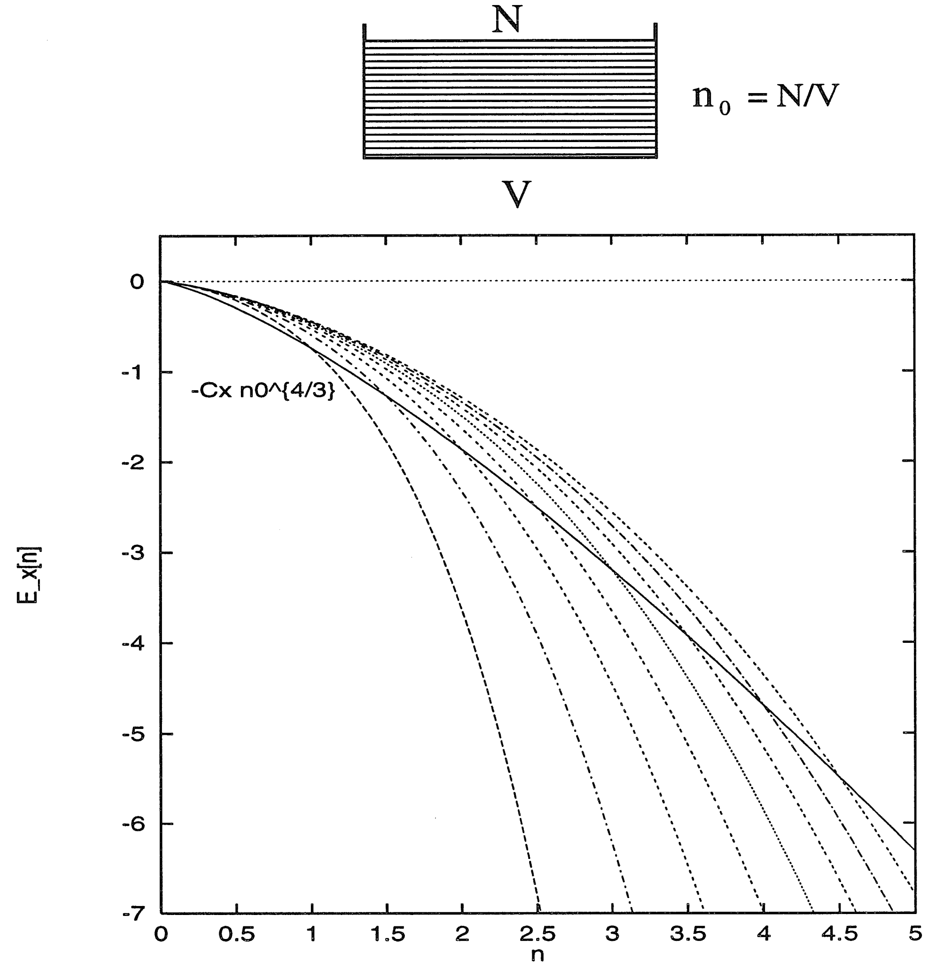


Figure 2.1: The bundle of $E_x[\rho]$s corresponding to different Vs of a box of homogeneous electron gas. Imagine V being the third dimension perpendicular to the paper. The ground state density $\rho_0 = N/V$ is just a combination of constraints N and V.

30 years.

Kohn-Sham Equation

From the energy functional $E[\rho]$ we get an Euler-Lagrangian equation for ground-state density ρ_0 by $\delta E[\rho]/\delta\rho = 0$. Introducing orbitals $\phi_i(\mathbf{x})$ that satisfies $\rho(\mathbf{r}) = 2\sum_i^{occ.} |\phi_i(\mathbf{r})|^2$, we can do variations with respect to $\phi_i(\mathbf{r})$: [10]

$$\frac{\delta E[\rho]}{\delta\phi_i(\mathbf{r})} = 0 \tag{2.3}$$

which leads to

$$\{-\frac{1}{2}\nabla^2 + v_{ext}(\mathbf{r}) + v_{coul}(\mathbf{r}) + v_{xc}(\mathbf{r})\}\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r}) \tag{2.4}$$

where

$$\nabla^2 v_{coul}(\mathbf{r}) = -4\pi\rho(\mathbf{r}) \tag{2.5}$$

and

$$v_{xc}(\mathbf{r}) = \frac{\delta E_{xc}[\rho]}{\delta\rho(\mathbf{r})}. \tag{2.6}$$

The fundamental permutation symmetry of Fermions are best encoded in the formulation in terms of *orbitals*, though these orbitals are themselves not observables. As a first approximation, one assumes that $E_x[\rho] = -C_x \int \rho^{4/3} d\mathbf{r}$, with $C_x = \frac{3}{4}\left(\frac{3}{\pi}\right)^{1/3}$ an expression for the homogeneous electron gas of density ρ. This is the famous local-density-approximation (LDA).

2.2 Generalized Gradient Approximation

The LDA neglected the gradient($|\nabla\rho|$) effect. From the very beginning people have tried to incorporate the $|\nabla\rho|$ effect into the energy functional $E_{xc}[\rho]$. The direct expansion of $E_{xc}[\rho]$ [7] in terms of $|\nabla\rho|$,$|\nabla^2\rho|$, etc. could lead to divergence. Further more, as Perdew and Wang pointed out [11], the exchange energy functional $E_x[\rho]$ have to satisfy certain constraints arising from the exact expression:

$$E_{xc} = \frac{1}{2}\int\int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\rho(\mathbf{r}_1)\rho_{xc}(\mathbf{r}_1,\mathbf{r}_2)}{r_{12}}. \tag{2.7}$$

For instance, the hole normalization constraint

$$\int d\mathbf{r}_2\rho_{xc}(\mathbf{r}_1,\mathbf{r}_2) = -1, \tag{2.8}$$

and the asymptotic behavior for exchange energy *density* $\epsilon_x(\mathbf{r})$:

$$E_{xc}(r \to \infty) = -\frac{1}{2}\int d\mathbf{r}\frac{\rho(\mathbf{r})}{r}. \tag{2.9}$$

One must formulate $E_x[\rho]$ that satisfies those constraints.

The degree of inhomogeneity can be measured by the de Brogie wavelength λ, which is measured by the Fermi wavevector k_F. Define the reduced density gradient

$$s = \frac{|\nabla\rho|}{2k_F\rho}, \tag{2.10}$$

where $k_F = (3\pi^2\rho)^{1/3}$ is the Fermi wavevector of homogeneous electron gas of density ρ. $s = \lambda/l$ is a measure of the inhomogeneity, λ is the de Brogie wavelength and l is the extent of smoothness. Thus the inhomogeneity could either be due to the large λ or due to the small l. At the bond midpoint, $|\nabla\rho| = 0$, s is zero, as expected. But near the nucleus where the density is high, thus the λ is small, the reduced gradient s could also be small. Figure 2.2 shows the spatial distribution of the reduced gradient s from Ar atom. Notice the shell structure reflected in $s(r)$. s becomes very small each time $|\nabla\rho|$ becomes zero. Also we learned that the important regions are $s < 2.5$.

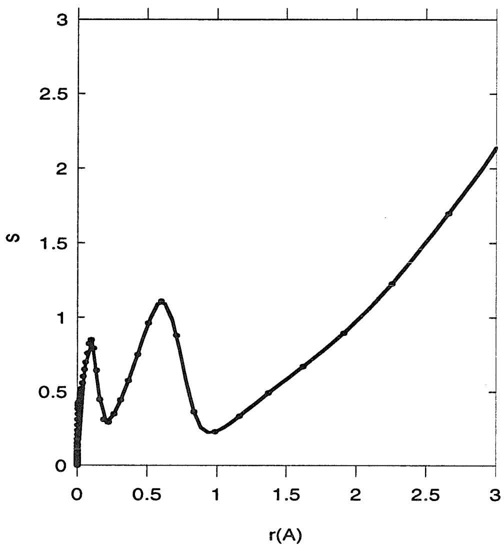


Figure 2.2: The spatial distribution of reduced gradient calculated for Ar atom. $s = |\nabla\rho|/2k_F\rho$. r is in unit of $\mathring{A}$.

2.2.1 Gradient Corrected $E_x[\rho]$

A general energy functional of ρ and $|\nabla\rho|$ can be written as

$$E_x[\rho] = \int f(\rho, s) d\mathbf{r}. \tag{2.11}$$

Expanding the exchange-hole correlation, Perdew [11] proposed the generalized gradient approximation, GGA, which has the following form:

$$E_x^{GGA}[\rho] = -C_x \int \rho^{4/3} F(s) d\mathbf{r}. \tag{2.12}$$

(Namely, he factorized the $f(\rho, s)$ in equation 2.11.)

For homogeneous electron $s \to 0$, $F(s) \to 1$; for electrons in molecules and atoms, $F(s)$ need to be determined. Here we performed HF calculation on a set of atoms and molecules. From our calculation we can get the exchange energy density $e_x(\mathbf{r})$. This $e_x(\mathbf{r})$ divided by $e_x^{LDA}(\mathbf{r})$ gives the factor $F(s) = F(s(\mathbf{r}))$ for each case, see equation 2.12. Figure 2.3 2.4 shows such a $F(s(\mathbf{r}))$.

We observed that $F(s(\mathbf{r}))$ fall more or less onto one curve. This allows us to model the e_x by a universal function $F^{GGA}(s)$. We noticed that $F(s)$ is not single valued at $s = 0$. It has two major branches. However, the lower branch arises from those regions of s which correspond to the high k_F at the nucleus. As one can see in the figure 2.2, this region of s is in the extreme vicinity of nucleus. So this branch contributes very little to the total energy. We can safely ignore this branch altogether. We adopted an analytical expression due to Becke [12]

$$F(s) = 1 + \frac{b}{2C_x} \frac{x^2}{1 + 6bx ln(x + \sqrt{1 + x^2})} \tag{2.13}$$

with

$$x \equiv \frac{|\nabla\rho_\sigma|}{\rho_\sigma^{4/3}} = (48\pi^2)^{1/3} s. \tag{2.14}$$

$\rho_\sigma = 0.5\rho$ is the density of spin component σ. b=0.0042 a.u. comes from the fitting.

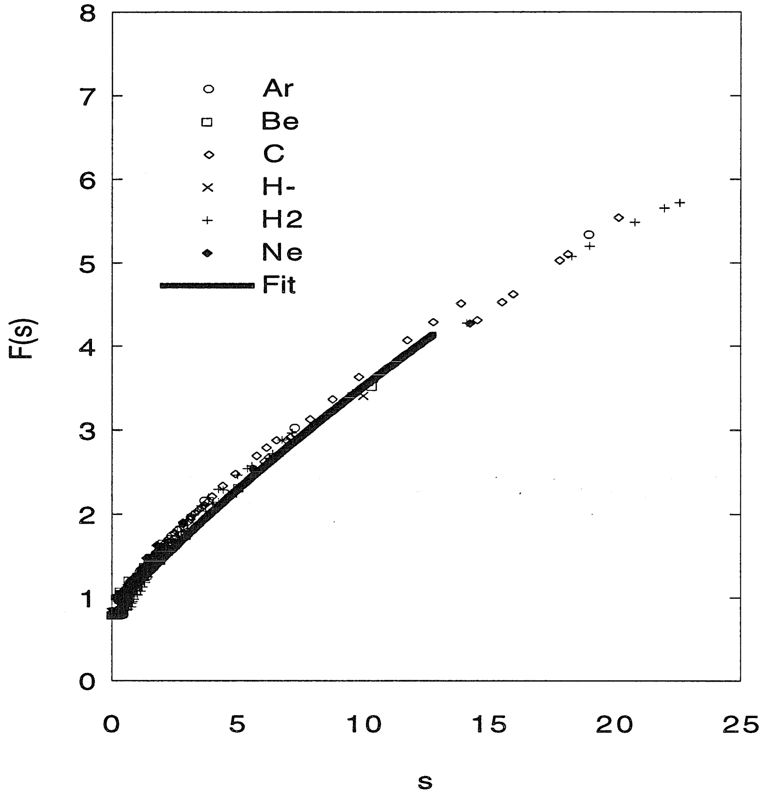


Figure 2.3: $E_x^{GGA}[\rho]$ scale factor $F(s) = e^{HF}(\mathbf{r})/e^{LDA}(\mathbf{r})$ versus s($\mathbf{r}$). $F(s)$ of all the cases fall onto one curve. HF is done with 6-31G* basis set.

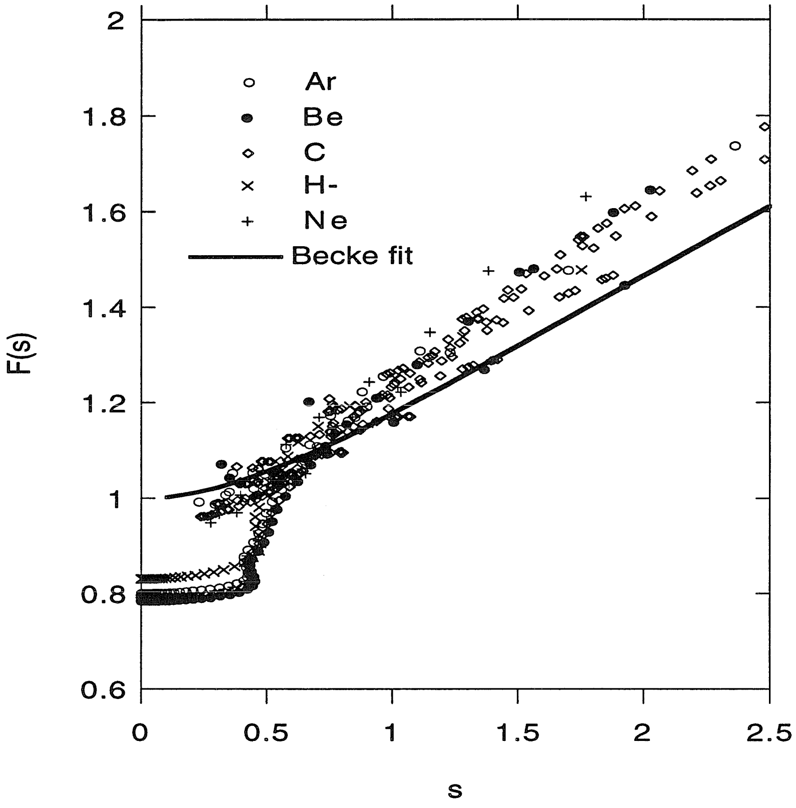


Figure 2.4: $E_x^{GGA}[\rho]$ scale factor $F(s)$. A closer look shows that there is a major branch at the very small s. This region corresponds to the extreme vicinity of nucleus and has a negligible contribution due to small volume element.

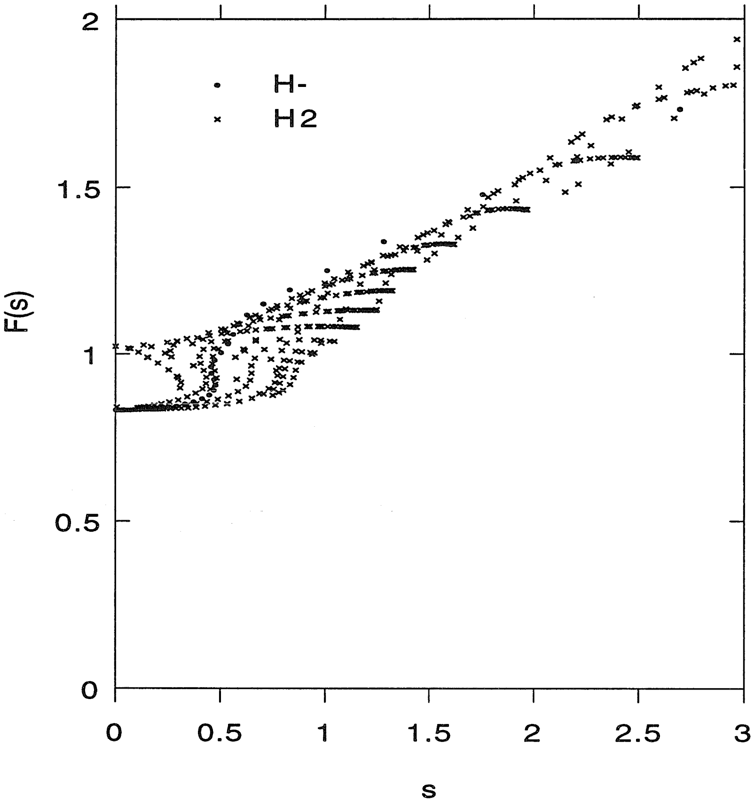


Figure 2.5: Scale factor $F(s)$ for $E_x^{GGA}[\rho]$, a closer look on H_2 to show the multivalueness of the $F(s(\mathbf{r}))$. One branch at $s \to 0$ corresponds to the bond mid-point, the other nucleus.

	E_x^{LDA}	E_x^{GGA}	E_x^{HF}	$E_x^{GGA} - E_x^{HF}(eV)$
He	-0.86221	-1.023625	-1.026732	0.08
Be	-2.317598	-2.670203	-2.682562	0.34
C	-4.400052	-5.013522	-5.029892	0.44
O	-7.260082	-8.130775	-8.135856	0.14
Ne	-10.998326	-12.131625	-12.126511	-0.14
Ar	-27.792172	-30.111732	-30.181348	1.89
H_2	-0.552381	-0.655275	-0.653440	-0.05
H_2O	-8.099892	-8.996603	-8.978400	-0.50
C_2H_6	-11.139656	-12.481290	-12.523264	1.14

Table 2.1: Exchange energy by GGA compared with LDA(exchange only) and HF. Energy unit is Hartree.

Becke designed this formula to reproduce the asymptotic behavior(see equation 2.9).

In Figure 2.3 2.4 we compare our fitted $F(s)$ with that from HF. The overall fit is excellent. A closer look of $F(s)$ shows some minor scatterings. How does the quality of the overall fit translate into the accuracy in reproducing the HF exchange energy? Table 2.1 listed the E_x^{GGA} calculated with equation 2.12 with our fitted $F(s)$. We can see that GGA with Becke's formula [12] made significant improvement over LDA.

2.2.2 Gradient Corrected $E_c[\rho]$

There are two commonly used approximations to E_c in DFT calculations. One is from the Jastrow type of wavefunction [13] due to Colle and Salvetti [14] and its density formulation(including gradient) by Lee, Yang and Parr [15]. The other is based on the accurate Monte Carlo calculation on uniform electron gas by Ceperley and Alder [16] and its Padé analytical fit by Vosko, Wilk and Nusair(VWN), [17] and later by Perdew and Wang. [18] For calculating the electronic excitation spectrum ϵ_i, on the other hand, GW method based on the Feynman-Dyson perturbation is usually used. [19, 20]

In this work we add the gradient correction to the $E_c[\rho]$ of VWN due to Perdew and Wang. [18] [5] Our result is shown in Table 2.2. More than 90% of the experimental

	E_{xc} (Hartree)	E_x^{HF} (Hartree)	$-E_c$ (ev)	$-E_c(Expt.)$ (ev)
He	-1.071565	-1.026732	1.2200	1.14
Be	-2.768872	-2.682562	2.3486	2.57
C	-5.189122	-5.029892	4.3321	4.26
O	-8.406445	-8.135856	7.3632	7.02
Ne	-12.526312	-12.126511	10.8792	10.61
Ar	-30.907508	-30.181348	19.7600	21.42

Table 2.2: Correlation energy calculated with GGA compared with experiment. 6-31G* basis set with S=0.

	HF	LDA	GGA	$expt.$
C_2	0.73	6.07	7.01	6.36
C_2H_4	18.71	27.51	24.71	24.65
C_2H_6	24.61	34.48	31.16	31.22

Table 2.3: Atomization energy calculated with GGA compared experiments, together with HF and LDA. Energy unit is in eV

correlation energy are recovered in these cases.

Combining the gradient corrected $E_x[\rho]$ with $E_c[\rho]$, we get a much improved atomization energy. Table 2.3 shows our calculation on some systems. The equilibrium geometries are not very sensitive to whether one uses HF or LDA or GGA. All of the theories give bond lengths to within $0.02\AA$ and bond angles to about 2^o. Pople's group had documented the performance of a family of density functional methods [1] and reached the similar conclusion. We conclude that with GGA E_x and GGA E_c DFT becomes a practical method of getting the structure and atomization energy (cohesive energy).

2.3 Band Gap and Optimal v_x

Photoemission spectra, electric conductivity, photoconductivity, superconductivity, and many others are determined by the electronic structures. These properties depend sensitively on the detailed shape of one particle potentials. DFT, by providing orbital energies through eigenvalues of Kohn-Sham equation, is a first principle tool to predict electronic structures. DFT can give the correct band-structure of a crystal, $\epsilon(\mathbf{k})$, but the band gap E_g is always off by a factor of about 2/3. In this section we first analyze the nature of orbital energy and show the condition for the correct one-particle potential $v_x(\mathbf{r})$, which directly effects $\{\epsilon_i\}$. Then we propose a new functional whose $v_x(\mathbf{r})$ satisfies these conditions. We will show that our functional leads to the much improved eigenvalues, $\{\epsilon_i\}$.

2.3.1 Band Gap and (IP-EA)

The electron affinity(EA) is defined by

$$EA = E(N+1) - E(N). \tag{2.15}$$

The ionization potential(IP) is defined by

$$IP = E(N) - E(N-1). \tag{2.16}$$

The band gap E_g is defined by $E_1 - E_0$, the energy difference between the first excited states and the ground state energy. As a commonly used approximation, in both HF and DFT, the band gap is calculated as $\epsilon_{LUMO} - \epsilon_{HOMO}$. This scheme found its base on Koopman's theorem in HF and VonBarth's derivation for DFT [21]. But, as is well known, such a scheme in HF gives a band gap almost twice as large as the measured one, while in DFT gives only about half to 2/3 of the measured one. In *crystal*, since adding or removing one electron wouldn't change the whole band structure, $-\epsilon_{HOMO} = IP$ and $-\epsilon_{LUMO} = EA$ could be used as the zeroth order approximation upon which to add perturbation terms. So

$$E_g = \epsilon_{LUMO} - \epsilon_{HOMO} = IP - EA. \tag{2.17}$$

	ϵ^{sahni}	ϵ^{HF}	$\delta\epsilon^{sahni}$	ϵ^{Talman}	ϵ^{HF}	$\delta\epsilon^{talman}$
C	-0.409	-0.434	-0.025	-0.40	-0.43	-0.03
Ne	-0.8565	-0.8505	-0.006	-0.86	-0.80	+0.06

Table 2.4: Highest occupied orbitals calculated within work formalism of Sahni [22] and calculated with Talman's scheme. Both are calculated on a radial grid and used spherical average. Energy unit is in Hartrees.

In atoms, however, the addition or removal of an electron would change the orbital so much that $\epsilon_{HOMO} - \epsilon_{LUMO}$ is not an accurate description of the $IP - EA$ at all. So we use the equation 2.15 and 2.16 to calculate IP and EA. For molecules, one needs to find the optimal trade off between the costly but accurate $E_0(N) - E_0(N-1)$ and cheap but approximate $\epsilon_{LUMO} - \epsilon_{HOMO}$.

2.3.2 Eigenvalue Problem

For HF the problems are (i) Neglecting the *correlation energy*; (ii) the virtual orbitals are not accurate enough, since all the resources are used to make the optimal set of occupied orbitals that enter the E_{tot} calculations. For LDA the problems are the wrong asymptotic behavior and self-interaction (SI).

One-Particle Potentials

Eigenvalues are very sensitive to the one-particle potential. In HF the SCF potential an electron feels is orbital dependent, or nonlocal potentials. Following Slater's X_α method of 1951, there are many developments to replace the nonlocal potential with a local one, most noticeably the LDA potential $v_x = C\rho(\mathbf{r})^{1/3}$. In 1976, Talman and Shadwick [23] devised a method to extract such local potentials. Talman's scheme involves minimizing the expectation value of the Hartree-Fock Hamiltonian and leads to a complex linear integral equation for the local exchange potential. It can reproduce eigenvalues of HF to within 5%. [23] See table 2.4.

In 1989, Sahni's group proposed the work formalism. Imagine an electron located at the $\mathbf{r}'$. Its Fermi hole charge distribution $\rho_{xc}(\mathbf{r}', \mathbf{r}'')$ created an electric field strength

$$E_{xc}(\mathbf{r}') = \int \rho_{xc}(\mathbf{r}', \mathbf{r}'') \frac{(\mathbf{r}' - \mathbf{r}'')}{|\mathbf{r}' - \mathbf{r}''|^3} d\mathbf{r}'' \tag{2.18}$$

The work done in bringing an electron from infinity to $\mathbf{r}$ against this electric field,

$$w_{xc}(\mathbf{r}) = -\int_{\infty}^{\mathbf{r}} E_{xc} \cdot d\mathbf{l}, \tag{2.19}$$

is the local effective exchange-correlation potential seen by the electrons. This $w_{xc}(\mathbf{r})$ has to be derived from SCF solution with the orbital representation of the exchange hole $\rho_{xc}(\mathbf{r}, \mathbf{r}')$. The work formalism provided the physically appealing interpretation of the exchange potential. For most atoms, it leads to eigenvalues and total energies that are within $10mH$ of that by HF method. (Table 2.4.) This established that the nonlocal exchange potential, at least in atoms, can be accurately represented by an *effective local potential.* However, the cost of work formalism is larger than that of HF, since on top of the normal HF SCF, it requires the integration of equation 2.18 and 2.19.

GGA Eigenvalues

Talman [23] showed that the effective local potential $v_x[\rho]$ has to have the asymptotic behavior [23]

$$v_x \rightarrow -1/r. \tag{2.20}$$

In LDA both e_{xc} and $v_x \rightarrow e^{-\lambda r}$ as $r \rightarrow \infty$. This is because the LDA e_{xc} and v_x are expressed in terms of $\rho^{1/3}$ which decays exponentially at infinity.

By design, Becke's analytical fit(equation 2.13) can give the correct asymptotic behavior for e_x, the exchange energy density. But the variational relationship $v_x = \delta E_x / \delta \rho(\mathbf{r})$ leads to quite different asymptotic behavior for v_x. This is why that even though E_x^{GGA} gives good total energy it does not improve the eigenvalues, see Table 2.5. Clearly, we need to improve the GGA exchange functional 2.12.

	E_{tot}^{GGA}	E_{tot}^{HF}	$homo^{GGA}$	$homo^{HF}$
He	-2.853999	-2.855160	-.539726	-0.9149
Be	-14.560105	-14.56676	-.170738	-0.3013
C	-37.588045	-37.585673	-.138214	-0.3420
O	-74.672631	-74.656607	-.251667	-0.5758
Ne	-128.496452	-128.474402	-.396733	-0.8306
Ar	-526.745126	-526.773735	-.330767	-0.5900

Table 2.5: HOMOs calculated with GGA-exchange-only compared with HF. 6-31G* basis set with S=0. Energy unit is in Hartrees

The New Functional

We propose a new functional that has the correct asymptotic behavior for both the energy density $\epsilon(\mathbf{r})$ and one-particle potential $v_x(\mathbf{r})$.

One can assume that the constraint search [8] gives, close to the true ground-state density ρ_0:

$$E_x[\rho] = -C_x \int \rho^{4/3} F(s_0) e^{\beta(\rho-\rho_0)/\rho_0} d\mathbf{r}. \tag{2.21}$$

Then we have

$$E_x[\rho_0] = -C_x \int \rho_0^{4/3} F(s_0) d\mathbf{r}, \tag{2.22}$$

and

$$v_x\Big|_{\rho_0} = -C_x \rho_0^{1/3} F(s_0) \left(\frac{4}{3} + \beta\right) = \left(\frac{4}{3} + \beta\right) \epsilon_x, \tag{2.23}$$

where

$$\epsilon_x = -C_x \rho_0^{1/3} F(s_0) \tag{2.24}$$

and β may have spatial dependence.

Note that equation 2.21 cannot be considered as a known functional since we don't know $\rho_0(r)$ other than that it exists and is unique.

According to their asymptotic behavior (see equation 2.9,2.20) we have $v_x = 2\epsilon_x$ at $r \rightarrow \infty$. Thus we have $\beta = 2/3$ for $r \rightarrow \infty$. For simplicity, we assume β to be a constant, independent of positions. The proposed exchange energy functional is then

$$E_x[\rho] = -C_x \int \rho^{4/3} F(s_0) e^{2(\rho-\rho_0)/3\rho_0} d\mathbf{r}. \tag{2.25}$$

Variation on this $E_x[\rho]$ leads to the following relationship:

$$v_x(\mathbf{r}) = 2\epsilon_x(\mathbf{r}). \tag{2.26}$$

The key feature of functional 2.25 is that *the gradient term s does not participate in the variation $\delta E_x/\delta\rho$.*

In the case of homogeneous electron gas, $F(s_0) = 1$, $\epsilon_x = -C_x\rho^{1/3}$.

$$v_x = -2C_x\rho^{1/3} = -\frac{3}{2}\left(\frac{3}{\pi}\right)^{1/3} \rho^{1/3}. \tag{2.27}$$

This is exactly the exchange potential of a homogeneous electron gas averaged over $\mathbf{k}$, the orbital quantum number for the uniform gas.

Proof:

$$\omega_{\mathbf{k}\sigma} = \frac{k^2}{2m} - 1/\Omega \sum_{\mathbf{q}} \frac{4\pi}{q^2} n_{\mathbf{k}+\mathbf{q},\sigma}. \tag{2.28}$$

The sum over $\mathbf{q}$ can be performed and gives

$$\begin{aligned} \omega_{\mathbf{k}\sigma} &= \frac{k^2}{2m} - k_F/2\pi\left(2 + \frac{k_F^2 - k^2}{kk_F} ln\left|\frac{k+k_F}{k-k_F}\right|\right) && (2.29)\\ &= \frac{k^2}{2m} - k_F/2\pi F(k/k_F). && (2.30)\\ &= \frac{k^2}{2m} - v_x(\mathbf{k}). && (2.31) \end{aligned}$$

	E_{tot}^{GGGA}	E_{tot}^{HF}	$homo^{GGGA}$	$homo^{HF}$
He	-2.8491	-2.85516	-0.8897	-0.9149
Be	-14.5570	-14.56676	-0.3238	-0.3013
C	-37.5594	-37.58567	-0.3589	-0.3420
Ne	-128.4406	-128.47440	-0.8221	-0.8306
H_2O	-76.011392	-76.04940	-.489160	-0.497470
C_2H_6	-79.194130	-79.20862	-.520626	-0.483115
glysine	-282.819449	-282.844442	-.414989	-0.397651

Table 2.6: HOMO calculated by the GGGA method. 6-31G* basis set, S=0. Energy unit is in Hartrees

Averaging over all occupied states, i.e., the Fermi sphere, leads to

$$\begin{aligned} v_x &= -\frac{k_F}{2\pi}\overline{F(k/k_F)} && (2.32) \\ &= -\frac{3}{2}\left(\frac{3}{\pi}\right)^{1/3}\rho^{1/3} && (2.33) \end{aligned}$$

where we used

$$\overline{F(k/k_F)} = \frac{3}{4\pi k_F^3}\int_{k<k_F} d\mathbf{k}F(k/k_F) = 3. \tag{2.34}$$

Q.E.D.

Slater put a factor α in front of the above expression for v_x and LDA gives $\alpha = 2/3$. Our functional shows that α should be 1. We shall refer our new functionals as the generalized generalized gradient approximation, GGGA. For a nonhomogeneous electron system, the gradient factor, $F(s_0)$ plays the role of setting the scale. We have tested our new functional on several atoms and molecules. The results are listed in table 2.6. Our functional 2.25 reproduces the HF eigenvalue to a good accuracy.

Comparing with Talman and Sahni in Table 2.4, both of them use orbitals and are very costly, our GGGA gives the same accuracy in eigenvalues at a much less cost. The new functional is still a density functional. It does not evoke orbitals.

Since our functional requires as constants the ground-state density ρ_0, which is

	ϵ^{LDA}	ϵ^{GGA}	ϵ^{GGGA}	ϵ^{HF}	$\epsilon^{GGGA} - \epsilon^{HF}$
C	-9.8685	-10.0190	-11.0990	-11.3461	0.247
	-0.4298	-0.4415	-0.6969	-0.7099	0.013
	-0.1415	-0.1382	-0.3590	-0.3422	0.017
Ne	-30.1933	-30.4461	-32.1177	-32.7628	0.645
	-1.2105	-1.2319	-1.6708	-1.9120	0.241
	-.38266	-0.3891	-0.8221	-0.8306	0.008

Table 2.7: Eigenvalues calculated with the LDA, GGA, GGGA method and HF. Our calculation is on 6-31G* basis set and uses spherical average. Energy unit is in Hartrees

unknown, the evaluation of the $E_x[\rho]$ has to be done in an iterative way. (Similar to the case of *natural orbitals*, [20] which uses the eigenfunctions of the density-matrix as the fastest converging basis for Hamiltonian without first knowing the density-matrix.) Here the bootstrapping process is just the normal SCF procedure. Like in ordinary calculus, one is always safe in plugging the ρ_0 *after* the variational

$$v_x(\mathbf{r}) = \frac{\delta Ex[\rho]}{\delta\rho(\mathbf{r})} \tag{2.35}$$

is done, and since the Kohn-Sham equation itself is solved self-consistently, we can safely *switch the order between finding* ρ *in KS equation and setting* ρ *equal to* ρ_0. After all when we find the SCF ρ, it *is* assumed to be our ρ_0.

This scheme gives E_{tot} close to the exact one(HF) but leads to slower convergence than the GGA. Also, we noticed that some of the eigenvalues are larger than HF while some are smaller. As one can see from Table 2.6, our GGGA improves eigenvalues $\{\epsilon_i\}$ from within 50% to within 7% of HF while only worsening the E_{tot} by less than 0.07%. (see Table 2.6)

We also tested a scheme whereby one performs the normal GGA and uses v_x^{GGGA} to get the eigenvalues *after* the SCF. Our tests show that such a scheme leads to results closer to that of normal GGA, i.e. better E_{tot} but bad ϵ_i. We found that v_x^{GGA} and v_x^{GGGA} lead to two different densities, which lead to different $\{\epsilon_i\}$ and E_{tot}.

The E_{tot}^{GGGA} are off by about $20mH$. Though it is only about 0.07%, it is too large for calculating atomization energy. However, in calculating the atomization energy the important quantity is not the absolute energy but the energy difference. The error in atomization energy should be much smaller. The tests are underway and results will be presented later.

The successful improvement of eigenvalues by our GGGA indicates the importance of having the correct boundary condition (asymptotic behavior) for the eigenvalues problem.

Self-Interaction

E_{coul} include self-interaction energy(SI), as is evidenced by hydrogen atom.

$$E_{coul} = \frac{1}{2}\int d\mathbf{r}_1 d\mathbf{r}_2 \frac{|\phi_0(\mathbf{r}_1)|^2|\phi_0(\mathbf{r}_2)|^2}{r_{12}} \tag{2.36}$$

$E_x[\rho]$ should include the SI to cancel exactly the SI from Coulomb term. This requirement, while satisfied by HF, is not met by most DFT functionals. Compare the total energy expression for HF and DFT:

$$E_{tot}^{HF} = 2\sum_{i}^{occ.}\epsilon_i - \frac{1}{2}\sum_{ij}^{occ.}\left(\langle\phi_i\phi_j|\frac{1}{r_{12}}|\phi_i\phi_j\rangle - \langle\phi_i\phi_j|\frac{1}{r_{12}}|\phi_j\phi_i\rangle\right), \tag{2.37}$$

$$E_{tot}^{DFT} = 2\sum_{i}^{occ.}\epsilon_i - \frac{1}{2}\int\frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}}d\mathbf{r}_1 d\mathbf{r}_2 + E_{xc}[\rho] - \int v_{xc}(\mathbf{r})\rho(\mathbf{r})d\mathbf{r}. \tag{2.38}$$

Take He atom for example where all $E_x[\rho]$ should be self-interaction. We should have,

$$E_x^{DFT}[\rho_0] - \int d\mathbf{r} v_{xc}(\mathbf{r})\rho(\mathbf{r}) = -E_x^{HF}[\rho_0] \tag{2.39}$$

if we want to have $E_x^{DFT} = E_x^{HF}$ we must have:

$$\int v_x(\mathbf{r})\rho_0(\mathbf{r}) = 2E_x^{DFT}[\rho]. \tag{2.40}$$

But for LDA we always have

$$\int v_x(\mathbf{r})\rho_0(\mathbf{r}) = \frac{4}{3}E_x^{LDA}[\rho]. \tag{2.41}$$

Thus the requirement is not met. On the other hand, our new functional 2.26 satisfies this requirement. So at least in this case our new functional is self-interaction free.

Fermi and Amaldi [1] realized this in 1934 and proposed the simple self-interaction-corrected formula

$$E_{coul}[\rho] = \frac{N-1}{N}\int\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{r_{12}} d\mathbf{r}_1 d\mathbf{r}_2. \tag{2.42}$$

Perdew and Zunger proposed the self-interaction-correction (SIC) version for a given approximation of E_{xc} [24]. SIC improves the total energy calculation E_{tot} and the E_x, E_c separately; It improves the eigenvalues, giving better band gaps; It also gives the correct long range behavior of the potential and density. However, SIC posed considerable cost increase. In fact, it requires the potential v_{xc} to be orbital dependent, already deviate from the spirit of an economic density functional, moving toward *ab initio* orbital representation. Self-interaction, long range behavior, eigenvalue problems all seem to be related. Since the complicated SIC usually can give the correct band gap, and since our new functional can also be self-interaction free, at least for some cases, we can hope that the new functional 2.21, with more sophisticated choice of β, e.g., $\beta(\rho_0, s_0)$, will give the right energy band gap. For its computation efficiency, this will be the first practical theoretical method for the accurate prediction of the energy band gaps. Further test need to be done to see whether this is actually the case.

2.4 Summary

The orbital nature of the QM has deep roots. Using observables alone to represent QM has a fundamental problem due to Bell's inequality. However the lure of possible savings and the appeal of building theories with observables attracted many attempts to build density theory. All of such theories involve modeling the various terms of total energy with that of known systems.

$$\begin{aligned} E_{tot}[\rho] &= T[\rho] + V_{ee}[\rho] + E_{ext}[\rho] \\ &= T_o[\rho] + E_{coul}[\rho] + E_x[\rho] + E_c[\rho] + E_{ext}[\rho]. \end{aligned} \tag{2.43}$$

The last term from electron-ion interaction depends strictly on density only. Thomas-Fermi theory model all energy terms after a homogeneous electron gas, kinetic energy as well as potential energy. The kinetic energy functional was too crude and Fermi-Dirac distribution is hard to implement. Parr and Yang [1] documented several attempts to improve the kinetic energy functional using only density and its gradient. Kohn-Sham's DFT treats the kinetic energy with orbitals, leaving potential energy functional to density. With gradient correction such approach enjoys great success due to the error cancelation and relative insensitiveness of potential energy to density error. It has been developed into a practical tool of structure calculations. Further extensions like Hartree-Fock-Kohn-Sham treat the exchange part, the major non-classical energy term, exactly, leaving only the correlation energy to density representation $E_c[\rho]$. But the improvements are not satisfactory (10% errors in atomic correlation energy [25]) due to the lack of error cancelations. Becke recently argued that DFT needs to have a mixture of exact exchange and density functional exchange,

$$E_{xc} = E_{xc}^{DFT} + a_0(E_x^{Exact} - E_x^{DFT}) \tag{2.44}$$

due to the "adiabatic connection". [26] Finally in self-interaction-corrected (SIC) scheme, v_{xc} bears the mark of individual orbitals. [24] The results are better but the cost also goes up.

Nowadays, gradient corrected DFT, with density representation for potential only, has become a practical tool. It delivers, on average, an MP2 accuracy for structural calculation at a significantly smaller cost. Comparing with the local MP2 [?] DFT requires much smaller memory. However, KS eigenvalues are typically too shallow by

50%. In this work we provide a direct and an easy way to improve the eigenvalues within the density approach. Our approach rests on the fundamental assumptions of DFT and takes advantage of the asymptotic behavior to determine the relevant parameters. The exchange-only version gives eigenvalues very close to those given by HF, demonstrating the importance of having the right boundary condition on the eigenvalues. Further incorporation of correlation energy functionals would leads to the long sought after improvement for predicting the energy band gap.

Here we can see the interplay between the bottom up(*Ab initio*) and empirical, top down (E_{xc}) approaches at the level of fundamental tools. Calibrating $F(s)$ against HF result, extracting $E_c[\rho]$ from quantum monte carlo, etc. This kind of interplay between computing and modeling is the optimal approach in material simulation.

Bibliography

[1] L. H. Thomas, Proc. Cambridge Philos. Soc. **23**, 542 (1927) E. Fermi, Z. Phys. **48**, 73 (1928)

[2] D. R. Hartree, Proc. Cambridge Philos, Soc. **24**, 89 (1928); V. Fock, Z. Phys. **44**, 455 (1927)

[3] J. C. Slater, Phys. Rev. **81**, 385 (1951)

[4] R. A. Friesner, Chem. Phys. Lett. **116**, 39 (1985); and reference in PS-GVB USER's Guide, Schrödinger, Inc. (1996)

[5] John P. Perdew, J. A. Chevary, S. H. Vosko, Koblar A. Jackson, mark R. Pederson, D. J. Singh and Carlos Fiolhais, Phys. Rev. B **46** 6671 (1992)

[6] B. G. Johnson, P. M. W. Gill and J. A. Pople, J. Chem. Phys. **98** 5612, (1993)

[7] P. Hohenberg and W. Kohn, Phys. Rev. **136**:B864-B871 (1964)

[8] M. Levy, Proc. Natl. Acad. Sci. USA **76**: 6062-6065. (1979)

[9] R. G. Parr and W. Yang *Density-Functional Theory of Atoms and Molecules* Oxford. (1989)

[10] Note: variation w.r.t. $\phi_i(\mathbf{r})$ covers the domain of definition of variation w.r.t.$\rho(\mathbf{r})$, but not vice versa.

[11] J. P. Perdew and W. Yue, Phys. Rev. B **33** 8800, (1986)

[12] A. D. Becke, J. Chem. Phys. **96** 2155, (1992) A. D. Becke, J. Chem. Phys. **97** 9173, (1992)

[13] R. Jastrow, Phys. Rev. **98**, 1479 (1955)

[14] R. Colle and O. Salvetti, Theoret. Chim. Acta (Berl.) **37** 329-334 (1975)

[15] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B, **37** 785 (1988)

[16] D. M. Ceperley, B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980)

[17] S. H. Vosko, L. Wilk and M. Nusair, Can. J. Phys. **58**, 1201 (1980)

[18] J. P. Perdew and Y. Wang, Phys. Rev. B **45** 13244 (1992)

[19] P. Fulde, *Electron Correlations in Molecules and Solids*, Springer-Verlag 1991

[20] Attila Szabo and Neil S. Ostlund, *Modern Quantum Chemistry*, McGraw-Hill 1989

[21] C.-O. Almbladh and U. von Barth, Phys. Rev. B **31** 3231, (1985)

[22] M. K. Harbola and V. Sahni, Phys. Rev. Lett., **62** 489 (1989) V. Sahni, Y. Li, and M. Harbola, Phys. Rev. A, **45** 1434, (1992)

[23] J. D. Talman and W. F. Shadwick, Phys. Rev. A, **14** 36 (1976)

[24] J. P. Perdew and A. Zunger, Phys. Rev. B **23** 5048 (1981)

[25] S. Baroni and E. Tuncel, J. Chem. Phys. **79**, 6140 (1983)

[26] A. D. Becke, J. Chem. Phys. **104**, 1040 (1996)

Chapter 3 Separable First Principles Pseudopotentials For Density Functional Calculations

Abstract

We describe a general method to generate transferable separable pseudopotentials (PP/S) for any element of the periodic table. With PP/S, the effort in calculating the matrix elements for localized basis sets (e.g., Gaussians) or for plane-waves scales linearly with the total number of basis functions. We illustrate the approach by extracting separable pseudopotential from the Bachelet-Hamman-Schlüter (BHS) pseudopotentials for several very different atoms (elements from groups 1, 8, 12, 14, and 17). This BHS/S pseudopotential is applied to crystals of diamond, silicon, germanium, and CdTe, with results in excellent agreement with those obtained using the non-separable potentials.

3.1 Introduction

The chemistry of molecules and materials can be well understood in terms of valence electrons (e.g., eight for Fe, four for Si, Ge, and Pb). However, because of the Pauli principle the quantum mechanical description requires a proper treatment of the core electrons. Usually the core electrons are treated using doubly-occupied orbitals in which case the major effect of the Pauli principle is accommodated by requiring the valence orbitals to be orthogonal to the core orbitals. Such explicit description of the core orbitals requires very localized orbitals which can dominate the cost of quantum mechanical calculations for large molecules and solids. Consequently, various approaches have been developed to replace the core electrons by pseudopotentials (PP) constructed to accurately represent the effects of the core electrons and the Pauli principle in the valence electrons. These methods started with the work of Phillips and Kleinman. [1] Goddard and coworkers [2, 3, 4, 5] developed approaches to extract first principles PP [called effective core potentials (ECP)] directly from *ab initio* calculations, culminating the work of Hay and Wadt [6] (which included relativistic effects) who constructed PP for all atoms Na through Bi. These methods for first principles PP were extended to density functional theory (DFT) by Hamann, et al, [7] leading to the BHS [8] PP commonly used in DFT calculations. Such PP's greatly simplify calculations; however, these first principles PP involve angular momentum projection operators (to account for the Pauli principle) that become computationally intensive for large systems. We present here an approach to modify the first principles PP in such a way as to eliminate angular momentum projection operators while obtaining equal quality results.

The organization of the paper is as follows. Section II described the method, PP/S, and derives working formulas. Section III discusses aspects of accuracy, robustness, efficiency, generality, and systematic improvement in the PP/S. Results on several systems, atoms and solids are reported in Section IV. Section V discusses relationships with the previous approaches.

3.2 The Separable First Principles PP Method, PP/S

3.2.1 Background

The one-particle equation to be solved in a rigorous all-electron calculation of an atom has the form

$$F\phi_i = \epsilon_i \phi_i, \tag{3.1}$$

where ϕ_i must be orthogonal to the core orbitals

$$\langle \phi_i | \phi_c \rangle = 0. \tag{3.2}$$

The Fock operator, F, is derived from the variational principle and has the form

$$F = t + V_{nuc} + V_{val} + V_{core}. \tag{3.3}$$

Here t is the kinetic energy operator,

$$V_{nuc} = -\frac{Z_{atom}}{R} \tag{3.4}$$

is the nuclear attraction potential, V_{val} involves the Coulomb, exchange, and correlation potentials for interactions with other valence electrons, and V_{core} contains all Coulomb, exchange, and correlation potentials involving the core electrons. In a first principles PP:

i. the Z_{core} core electrons are deleted leaving a total of

$$Z_{val} = Z_{atom} - Z_{core} \tag{3.5}$$

valence electrons to be calculated (for a neutral atom)

ii. the orthogonality condition (Eq. 8.77) is dropped and the wiggly valence orbital, ϕ_i, is replaced by a smooth pseudo-orbital, ψ_i, *iii.* the $V_{nuc} + V_{core}$ in (Eq. 3.3) is replaced by

$$V_{net} + V_{pp} \tag{3.6}$$

where

$$V_{net} = -\frac{Z_{val}}{R} \tag{3.7}$$

iv. the V_{pp} in (Eq. 3.6) is chosen so that

$$F_{pp}\psi_i = \epsilon_i\psi_i \tag{3.8}$$

where

$$F_{pp} = t + V_{net} + V_{pp} + V_{val}. \tag{3.9}$$

Finding the V_{pp} satisfying (Eq. 3.8) is referred to as inverting the orbital. The V_{pp} in (Eq. 3.9) depends on the angular momentum, l, of the pseudo-orbital, ψ_i. Thus V_{pp} can be written as [2]

$$V_{pp} = \sum_{l=0}^{\infty} V_l P_{lA} \tag{3.10}$$

where V_l is a function only of distance from nucleus A and P_{lA} is the projection operator onto states of angular momentum l about center A. For a valence orbital having the same angular momentum as a core orbital ($l \leq l_{core}$), the V_l will be very repulsive in the core region (leading to a smooth orbital that is weakly bound instead of a strongly bound core-like orbital). However, for a valence orbital having an angular momentum not corresponding to a core orbital ($l > l_{core}$), the V_l will be very smooth in the core region, corresponding to a shielded potential. It was found [2] that

$$U_l \approx V_{loc} \; if \; l > l_{core} \tag{3.11}$$

where

$$V_{loc} = U_{l_{core}+1}. \tag{3.12}$$

Thus using the identity

$$V_{loc} = \sum_{l=0}^{\infty} V_{loc} P_{lA} \tag{3.13}$$

we can write (eq. 3.14)

$$V_{pp} = V_{loc} + \sum_{l=0}^{l_{core}} U_l P_{lA}. \tag{3.14}$$

where

$$U_l = V_l - V_{loc} \tag{3.15}$$

Thus the first principles PP for an atom A becomes [2]- [8] (Eq. 3.14) where V_{loc} (a shielded Coulomb potential) is a function only of distance from the nucleus of A, and

$$V_{pp}^{nl} = \sum_{l=0,lcore} U_l P_{lA} \tag{3.16}$$

contains all non-local terms (with integral operators).

Given the smooth pseudo-orbital, ψ_i, the conditions (Eq. 3.8), (Eq. 3.9), and (Eq. 3.14) define a unique PP. However, one can choose the shape of the smooth pseudo-orbital ψ_i using several different criteria. The original ECP of Goddard [2] was based on a GVB-like wavefunction which included electron correlation in the core and valence electrons and led to a unique smooth pseudo-orbital. However, Melius and Goddard [3] showed that simple Hartree-Fock (HF) wavefunctions could be used (leading to doubly occupied core orbitals) with the pseudo-orbital ψ_i chosen as a smooth combination of the valence orbital ϕ_i and the all-electron core orbitals. Redondo, McGill, and Goddard [4] showed that because of normalization conditions the

Melius-Goddard formulation leads to a pseudo-orbital that is slightly too small far from the atom, and they proposed the coreless valence orbital (CVO) approach in which the pseudo-orbitals is taken as exactly the HF orbitals far from the nucleus and then smoothed to zero in the core region while retaining normalization (a procedure now called norm conservation [7]). This CVO PP for Si, Ga, Ge, and As was used in several applications. [9] Smedly, Rappè, and Goddard [5] (SRG) pointed out that smoothing the valence orbitals modifies V_{val} and suggested optimizing the smoothed pseudo-orbital to obtain the least change in V_{val} while retaining norm conservation. This approach (denoted as SHC for shape and Hamiltonian consistent) was applied to the second row atoms (Na-Cl). [5] However, SRG showed that even this most optimum PP leads to significant errors for first row atoms and recommended that PP not be used for Li-Ne. Although SHC was applied in several calculations, most *ab initio* ECP calculations use the Hayt-Wadt (HW) potentials [6] which are based on the original Melius-Goddard formulation [3] (without norm conservation). The HW PP have been successfully used in numerous *ab initio* calculations.

Hamman and coworkers [7] suggested that first principles PP be derived by imposing norm conservation and directly inverting the radial scalar Dirac equation solved using the LDA exchange-correlation potentials. Using this approach Bachelet-Hamman-Schlüter (BHS) developed PP for all atoms H-Pr. This BHS PP reproduces accurately the relativistic all-electron results on atoms and leads to accurate descriptions (within LDA) of molecules and solids.

Despite the general success of the first principles PP in (Eq. 3.14) the presence of the angular momentum projection operator, P_{lA} (where A is the center of the atom containing the PP), lead to problems for *ab initio* calculations using Gaussian basis functions. This requires evaluating integrals having the form

$$\langle \chi_B | U_{lA} P_{lA} | \chi_C \rangle \tag{3.17}$$

where A, B, and C may be on three different centers. Assuming that all atoms have PP, this leads to a computational effort scaling as NM^2, where N is the number of basis functions (or atoms), and M is the number of basis functions in the neighbor within the range of U_{lA}.

For the plane-waves methods most common for crystals, (Eq. 3.17) amount to N^2 number of integral like

$$\int j_l(kr)U_l j_l(k'r)r^2 dr P_l(cos\theta_{kk'}), \tag{3.18}$$

leading to an N^2 scaling. For first principles plane-waves calculations, such as Car-Parrinello,[3] the PP part becomes the computational bottleneck.

Kleiman and Bylander [11] showed a way to replace the nonlocal part of (Eq. 3.14) by a form that leads to the factorization when evaluating the matrix element (Eq. 3.17). Their form reduced the cost to N for each k-point in the Brillouin zone. But Kleiman and Bylander form can cause artifacts(so called ghost-states), where the excited orbitals fall into the valence levels, for some systems and is not generally valid. In this section we will develop a general approach for modifying the nonlocal part of the PP, V_{pp}^{nl} in (Eq. 3.14), to reduce the computational cost to N(for plane-waves) or M^2(for Gaussian), while avoiding other problems (ghost states).

3.2.2 The Method

Expanding the angular momentum projections operator in (Eq. 3.14) leads to

$$V_{cor}^{nl} = \sum_{l=0}^{l_{core}} \sum_{m=-l}^{l} U_l(r)Y_{lm}(\theta,\phi) \int d\bar{\Omega} Y_{lm}(\bar{\theta},\bar{\phi}) \tag{3.19}$$

where l_{core} in (Eq. 3.19) is generally the highest angular momentum contained in the core. The superscript nl denotes a nonlocal potential since it is an integral operator.

The radial function U_l is repulsive and short-ranged, decreasing exponentially with distance. This suggests expanding V_{cor}^{nl} using a basis set of Gaussian functions $\{\chi_{\mu}^{pp}\}$

$$\chi_{\mu}^{pp}(r,\theta,\phi) = N_{\mu} r^{l_{\mu}} e^{-\alpha_{\mu} r^2} Y_{l_{\mu}m}(\theta,\phi) \tag{3.20}$$

where N_μ is a normalization constant, l_μ is the angular momentum quantum number, and the Y_{lm} are standard normalized spherical or cubic harmonics.

We consider the eigenfunctions $\{\theta_i\}$ of V_{cor}^{nl} in this basis,

$$\theta_i = \sum_\mu \eta_{\mu i} \chi_\mu^{pp}, \tag{3.21}$$

such that

$$\langle \theta_i | V_{cor}^{nl} | \theta_j \rangle = \lambda_i \delta_{ij}. \tag{3.22}$$

Here the columns of η are the eigenfunctions and λ_i the eigenvalues. The θ_i have the form

$$\theta_i = \theta_{nlm} = R_{nl}(r) Y_{lm}(\theta, \phi). \tag{3.23}$$

A. The Direct Separable PP, V^{dir}

Using (Eq. 3.21) we define the direct separable pseudopotential as

$$V_{cor}^{dir} = \sum_{i=1}^{N_a} |\theta_i\rangle \lambda_i \langle \theta_i| \tag{3.24}$$

$$= \sum_{l=1}^{l_{core}} \sum_{m=-l}^{l} \sum_{n=1}^{N_{la}} |\theta_{nlm}\rangle \lambda_{nl} \langle \theta_{nlm}|, \tag{3.25}$$

where N_{la} is the number of functions with angular momentum l on center a, and (Eq. 3.25) emphasizes the angular momentum dependence in (Eq. 3.24).

Using (Eq. ??) the general three center integral (Eq. 3.17) becomes

$$\langle \chi_B | V_{cor,A}^{nl} | \chi_C \rangle \approx \langle \chi_B | V_{cor,A}^{dir} | \chi_C \rangle = \sum_{l=0}^{l_{core}} \sum_{m=-l}^{l} \sum_{n=1}^{N_{la}} \langle \chi_B | \theta_{nlmA} \rangle \lambda_{nl} \langle \theta_{nlmA} | \chi_C \rangle. \tag{3.26}$$

Thus it is only necessary to calculate the two-center overlaps

$$\langle \theta_{nlmA} | \chi_B \rangle \tag{3.27}$$

in order to obtain the general three-center integrals. Thus we refer to (Eq. 3.25) as a separable PP. In the limit that the basis sets $\{R_{nlA}\}$ are complete, the right-hand side of (Eq. 3.26) approaches the left-hand side.

B. The Potential Weighted Separable PP, V^{vw}

The optimal radial functions $\{R_{nl}\}$ for (Eq. 3.25) are dictated by the specific form of U_l and this may differ dramatically from atom to atom. In order to avoid reoptimizing the PP basis set for every atom, we propose an alternative approach in which $|V^{nl}_{cor}\chi^{pp}_{\mu}\rangle$ is used as the basis set for describing the potential. This potential weighted (VW) separable pseudopotential leads to the following form:

$$V^{vw}_{cor} = \sum_{i=1}^{N_a} |V^{nl}_{cor}\theta_i\rangle \frac{1}{\lambda_i} \langle \theta_i V^{nl}_{cor}| \tag{3.28}$$

$$= \sum_{l=0}^{l_{core}} \sum_{m=-l}^{+l} \sum_{n=1}^{N_{la}} |U_l R_{nl} Y_{lm}\rangle \frac{1}{\lambda_{ln}} \langle U_l R_{nl} Y_{lm}|. \tag{3.29}$$

Using (Eq. 3.29) the general three-center integral (Eq. 3.17) becomes

$$\langle \chi_B | V^{nl}_{cor,A} | \chi_C \rangle \approx \langle \chi_B | V^{vw}_{cor,A} | \chi_C \rangle = \sum_{l=0}^{l_{cor}} \sum_{m=-l}^{l} \sum_{n=1}^{N_{lA}} \langle \chi_B | V^{nl}_{cor,A} | \theta_{nlm,A} \rangle \frac{1}{\lambda_{lm}} \langle \theta_{nlm,A} | V^{nl}_{cor,A} | \chi_C \rangle \tag{3.30}$$

As with (Eq. 3.26), (Eq. 3.30) requires calculation only of two-center functions

$$\langle \theta_{nlm,A} | V^{nl}_{cor,A} | \chi_B \rangle \tag{3.31}$$

in order to obtain the three-center integrals.

C. Comparison of V^{dir}_{sep} and V^{vw}_{sep}

The difference between V^{vw}_{sep} and V^{dir}_{sep} are *i.* V_{sep}^{dir} requires calculation of two-center overlaps

$$\langle \theta_{nlm,A} | \chi_B \rangle \tag{3.32}$$

whereas V_{sep}^{vw} requires calculation of two-center PP integrals

$$\langle \theta_{nlm,A} | V_{cor,A}^{nl} | \chi_B \rangle. \tag{3.33}$$

Since programs have already been written for doing the latter,[11–13] this difference is mainly one of computational cost. And since the two-center integrals are a small part of the cost and scale linearly with size, this is not an issue.

ii. The expansion function $\{R_{nl,A}\}$ required for the right-side of (Eq. 3.29) to converge to the left-side depend mainly on the spatial extent of the core electrons on atoms A rather than on the detailed functional form of the potentials $\{U_l\}$. Consequently the criteria for choosing the basis functions are similar to those used for expanding the wavefunction rather than for expanding the potential. This has two advantages:

a. the standard approach of choosing basis sets can be used for $\{R_{nl,A}\}$,

b. since the sizes of the core electrons of all atoms of the periodic table differ only by about a factor of 2, one can imagine finding a universal set of basis functions $\{R_{nl}^{pp}\}$ to be used for all atoms.

For the reason given in (*ii*), we focus in this paper on the use of V_{sep}^{vw}. The procedures outlined above can be used for any first principles PP and we will denote the resulting separable potential as PP/S. Thus starting with the BHS PP, [8] the separable version is denoted as BHS/S; starting HW PP[6] leads to HW/S. In this chapter we will focus on BHS/PP.

3.2.3 Tests of the BHS/S

To be useful the separable pseudopotential $V^{BHS/S}$ must reproduce the electronic properties obtained by the non-separable pseudopotential V^{BHS}. Namely, PP/S should describe the effect of the PP on the energies and wavefunctions for molecules and solids. On the other hand, since the PP is an intrinsic property of the element, characterizing the core of that element, we should be able to define a testing procedure that involves only the atomic calculation, such that once passed the PP/S should be valid for other environments, molecules as well as solids.

In the following, we will test the separable potentials BHS/S by solving for the wavefunctions of atoms. compare the results with those obtained via BHS. We will examine the full spectrum of calculations and pseudo-orbitals.

With a pseudopotential V_{pp}, the Kohn-Sham equations for obtaining the optimum self-consistent orbitals, (Eq. 3.8) and (Eq. 3.9), become

$$F^{KS}\psi_i = \epsilon_i \psi_i \tag{3.34}$$

where

$$F^{KS} = t + V_{pp} + V_{ee} + V_{xc}. \tag{3.35}$$

This can be solved using a Gaussian basis set $\{\chi_\mu\}$,

$$\psi_i = \sum_\mu C_{\mu i} \chi_\mu, \tag{3.36}$$

where the basis funtions are the standard basis functions for describing the valence orbitals.

The Gaussian basis set $\{\chi_\mu\}$ used in (Eq. 3.36) to describe the pseudo-orbitals $\{\psi_i\}$ need not be related to the Gaussian basis set $\{\chi_\mu^{pp}\}$, used in (Eq. 3.21) for describing the potential V_{cor}^{nl}. This is demonstrated in Figure 3.1, where for Fe atom we plot $U_s(r)$ (the radial part of the s-component of V_{cor}^{nl}) and $\psi_{4s}(r)$ (the $4s$ pseudo-orbital) (both are on logarithmic scale). Here we see that $U_s(r)$ is concentrated within 2 a_0 of the nucleus while $\psi_{4s}(r)$ extends well beyond 9 a_0. Therefore different basis sets $\{\chi_\mu^{pp}\}$ and $\{\chi_\mu\}$ are needed to cover the respective regions of the real space.

We will be particularly concerned about the descriptions of core-like states with the PP/S. In an all-electron calculation, the orthogonality condition (Eq. 8.77), essentially removes all core orbital from the Hilbert sub-space of valence electrons. However, (Eq. 3.8) no longer has any orthogonality conditions so that the Hilbert spaces are not reduced. This means that somewhere in the spectra of states for (Eq. 3.8), there may occur core-like states that are artifacts of the procedure. These extra states are called ghost states and they should remain extremely high in energy so that they cannot play a role in any physical phenomenon. Hence it should be useful to test for the possibility of ghost states contamination of the results.

The basis sets $\{\chi_\mu\}$ in (Eq. 3.36) for describing the eignenfunctions of (Eq. 3.34) will be the primitive basis functions optimized for the MS4 all-electron basis, [9] which have been found to be accurate for all-electron calculations. (Table 1 lists the MS4 basis for Fe atom which contains 10 s-type, 8 p-type, and 4 d-type Gaussian functions.) This contains basis functions for describing the 1s, 2s, 2p, 3s, and 3p core orbitals of Fe.

For the basis $\{\chi_\mu^{pp}\}$ used to expand the PP of (Eq. 3.21), we will use even tempered Gaussian basis sets with N_l basis functions and two parameters. Thus for $N_l = 2p+1$, we choose the exponents in (Eq. 3.20) as

$$\{\alpha_n\} = \alpha_0\beta^{-p}..., \alpha_0\beta^{-1}, \alpha_0, \alpha_0\beta^{1}, ...\alpha_0\beta^{p}; \tag{3.37}$$

while for $N_l = 2p$ we choose them as

$$\{\alpha_n\} = \alpha_0\beta^{-(p-\frac{1}{2})}..., \alpha_0\beta^{-1/2}, \alpha_0\beta^{1/2}, ...\alpha_0\beta^{-(p-\frac{1}{2})}. \tag{3.38}$$

For simplicity we use the same exponents $\{\alpha_n\}$ for all angular momenta l. For the even tempered Gaussian basis, the whole basis set is characterized by two adjustable parameters: α_0 which specifies the mean size of the functions, and β which specified the spacing of the functions. Given (α_0, β), $\{\alpha_n\}$ can be written in terms of the powers in (Eq. 3.38) as $\{n\}$=$(..., -3, -2, -1, 0, 1, 2, 3, ...,)$ or $(...-2.5, -1.5, -0.5, 0.5, 1.5, 2.5, ..)$. The resolution of this n-grid can be modified by inserting sub-divisions, either homogeneously, like $(..., -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, ...)$, or inhomogeneously like $(..., -2, -1.5, -1, -0.5, 0, 1, 2, 3, ...)$, etc.

I. Fe Atom

The ground state of Fe atoms has the valence electron configuration $(3d)^6(4s)^2$ with 1s, 2s, 3s, 2p, and 3p core electrons. Low lying excited states of Fe have valence configuration of $(3d)^7(4s)^1$, $(3d)^8$, $(3d)^6(4s)(4p)$, etc. We first examine the ground state configuration, i.e. $(3d)^6(4s)^2$. Since in this case the MS4 basis contains ten s functions, eight p functions, and six d functions (after decontracting to get the primitives), we will obtain 10 orbitals of the type $(3d)^6(4s)^1(ns)^1$, 8 of the type $(3d)^6(4s)^1(np)^1$, and 6 of the type $(3d)^6(4s)^1(nd)^1$. For simplicity we denote the orbitals of $(3d)^6(4s)^1(ns)^1$ as $n = 4, 5, ..13$. Because the basis set was optimized for valence states $4s$, $4p$ and $3d$, the calculated unoccupied orbitals (virtual orbitals) will *not* describe the true single-particle excited state (which has continious spectrum). But our purpose is to obtain a sensitive measure of where the PP/S deviates from PP in the the description of core-like virtual orbitals. Comparing the virtual, as well as occupied orbitals, between PP and PP/S, we gain insights of the nature of ghost states.

In Figure 3.2 we show the full spectrum of orbital energies with various choices of $\{n\}$. Here 3, 4, and 5 denote $\{n\} = (-1, 0, 1), (-1, 0, 1, 2)$, and $(-1, 0, 1, 2, 3)$. $3'$ denotes $\{n\} = (-3, -2, -1)$. We observed that for 5, the states are accurate up to $9s$ at $\sim$ 60 h; for 4 they are accurate up to $7s$ at 7h, for 3 they are accurate up to $6s$ at 2h. The deviations at those high unoccupied states can be viewed as incipient ghost states. For $3'$ the accuracy collapsed and we have $5s$ falling below the occupied $4s$ state, a ghost state appears! In this case the eigenvalues are calculated from diagonalization *without* doing self-consistency, using the exact electron density. The eigenvectors are assigned to the corresponding exact ones by their overlaps.

In Figure 3.3 we show the errors as a function of N. δE_{tot} is the total energy of the

$(3d)^6(4s)^2$ ground state. $\delta\epsilon_{1h}$ is the the largest absolute error in orbital eigenvalues below 1 Hartree (i.e.ϵ_{4s} and ϵ_{5s}), and $\delta\epsilon_{2h}$ is similar ones below 2h (i.e. ϵ_{6s}). For $N = 5$ the error [$\delta E_{tot} = 0.002mh = 0.000054$ eV and $\delta\epsilon_{1h} = 0.09mh = 0.0024$ eV] are quite acceptable. Even for $N = 3$ the error of $\delta E_{tot} = 0.09mh = 0.024$ eV and $\delta\epsilon_{1h} = 0.7mh = 0.015$eV are well within the cases due to using PP in the first place.

Given $N = 5$ we show in Figure 3.4a the effect of optimizing α_0 with a fixed $\beta = 3.0$. Here we see that δE_{tot} approaches a minimum at $\alpha_0 \sim 0.75$. $\delta\epsilon_{1h}$ decrease as α_0 decrease until $\alpha_0 \sim 0.25$ when it shoots up due to the occurrence of ghost-states. Since at $\alpha_0 \sim 0.75$ $\delta\epsilon_{1h}$ is still around $0.06mH$ (=0.002eV=0.04Kcal/mol), we choose $\alpha_0 = 0.75$ to be our optimum α_0. Note that the unit in Figure 3.4 is in mH (= 0.027 eV = 0.627 kcal/mol). This indicates that the error is not very sensitive to the choice of α_0 given a sufficiently large N_l(e.g.,$N = 5$).

In Figure 3.4b we show the effect of optimizing β, fixing $N = 5$ and $\alpha = 0.75$. Again, the criterion for E_{tot} would be slightly different from that for ϵ_{1h}. However, for the accuracy requirement of most calculation they can be taken as the same. This leads to the optimum β of 3, for the given N and α_0.

Figure 3.4a and Figure 3.4b demonstrate the sensitivity of the choice of α_0 and β to the accuracy of PP/S. They are not sensitive for most purposes.

So far we have been focusing on ground state configuration $(3d)^6(4s)^2$. For other configurations, BHS/S also reproduces BHS results, as demonstrated in table 1.5.

II. F and Cs atom

Now we consider the largest core atom Cs and the smallest core atom F. Since the Cs core is large, we would need more diffused Gaussians to represent its PP/S. The scale β in even tempered Gaussian series should decreases correspondingly. Using $N = 7$ and $\beta = \sqrt{3}$ (instead of $\beta = 3$) leads to the results in Figure 3.5. It shows that $\alpha_0 = 0.15$ is best. This leads to $\{\alpha_n\} = 0.029, 0.050, 0.087, 0.15, 0.26, 0.45, 0.78$

For F Atom, one of the smallest core in the periodic table, we would need sharper Gaussians. Using $N = 5$ and $\beta = 3$ leads to the results in Figure 3.6a, indicating that $\alpha_0 = 3.0$ is optimum, leading to $\{\alpha_n\} = 0.33, 1.0, 3.0, 9.0,$ and 27.0. Fixing α_0 at 3.0, Figure 3.6b shows that the optimum β for F is $\beta = 3.0$

One can use the above procedure to setup a table of (α_0, β)s, optimized for each element within the even tempered sets. Or, taking the advantage of the insensitivity of the PP/S to the choice of (α_0, β)s, one can try to find a general set of (α_0, β) that works for all the elements without losing much accuracy. In this paper we prefer the simplicity to the accuracy,

III. The Universal Set

Based on the above results we find the following $\{\chi_\mu^{pp}\}$ give the best overall accuracy: $\{\alpha_n\} = \alpha_0\beta^n$, with $\alpha_0 = 0.75$, $\beta = 3.0$, and

$$\{n\} = (-3, -5/2, -2, -3/2, -1, -1/2, 0, 1, 2, 3). \tag{3.39}$$

There are 10 exponents all together. Figure 3.7 shows the s-type radial eigenfunction, $rR_{ns}(r)$, of V^{nl} for F, Fe and Cs atoms. As one can see, the eigenfunctions on this basis (Eq. 3.39) cover the space rather completely for cores ranging from F to Cs. Using the universal basis (Eq. 3.39) for $\{\chi_\mu^{pp}\}$, Table 2 lists the total energy from both V_{cor}^{dir} and V_{cor}^{VW}. With the VW method the accuracy is uniformly better than $0.1mH(\sim 0.003eV)$. In comparison, using the direct method, we have a general accuracy of around $3mH(\sim 0.1eV)$ (with exception for Os, where the near degeneracy of 6s and 5d cause big error for the direct method). This shows (a)that our universal basis (Eq. 3.39) is rather complete for describing the V_{cor}^{nl} of all the elements; (b)VW method is computationally more efficient. In the following we will always refer to the VW method when using $V^{BHS/S}$.

Given an element we can use a subset of 3, 4, 5,.... elements of (Eq. 3.39) to reduce the computational cost. Different $\{n\}$ cover different range of the real-space. For a particular element, the subset should be centered at n_0 and include only integers (except for $n_0 = -2$) at first. If not accurate enough, add in half integers. Examination of the V^{BHS} of all the elements leads to the following rules for selecting n_0:

Element (n_0)
Li(-1), Be–N(0), O–Ne(1)
Na(-1), Mg–Cl(-1,0),Ar(0)
K–Cr(-1), Fe–Cu(0),Zn(1)
Ga–Se(-1),Br–Kr(0)
Rb–Ru(-1), Ag–Cd(0)
In–Xe(-1)
Cs(-2),Ba–Hg(0)
Tl-Rn(-1)

The result of using a subset of (Eq. 3.39) according to the above table is listed in table 3. It shows that for most elements 3 primitive Gaussians ($N_l = 3$) is sufficient to achieve an accuracy of 3mH (0.1 eV) for the total energy and the orbital energy (for states under 1 Hartree =27.22 eV). Some elements, like Zn, Hg, F and Ba, etc. need 4 or 5. Larger N_l would always yield better accuracy.

IV. Procedure For Constructing PP/S

In summary, the general procedure for constructing an accurate PP/S is as follows:

1. *Choose suitable* $\{\chi_\mu^{pp}\}$. Namely, choose the values of $\bar{\alpha}_0$ and β from the above table. The size N_l depends on the accuracy desired but should be larger than 3.

2. *Solve (Eq. 3.21) for the* $\{\theta_i\}$. That is, we diagonalize V_{cor}^{nl} on $\{\chi_{l\mu}^{pp}\}$, using V^{BHS}.

3. *Construct the PP/S* using (Eq. 3.25) or (Eq. 3.29) for $V^{BHS/S}$.

4. *Test PP/S for the atomic states.* Using a basis set $\{\chi_\mu\}$ that allows a good description of the core orbitals, solve (Eq. 3.34) with $V^{BHS/S}$ and compare the result with that of V^{BHS}. If the excited states calculated account to $\epsilon \sim 4h = 108$ eV there should be no concern about ghost states.

3.2.4 Transferability

The real test for transferability of pseudopotentials (PP or PP/S) is an extensive test on a wide range of molecules and solids. A simpler one is to check the condition that the scattering properties of the pseudopotential and full potential(all-electron) have the same energy variation to first order. [8] This condition involves matching the orbitals of the PP with that of the all-electron case. Here we first demonstrate that BHS/S satisfies such conditions to the same accuracy as BHS PP. Then we test the PP/S on several solids.

Figure 3.8 compares the *orbitals* of Fe calculated with BHS/S with the ones from BHS, using 6 functions in $\{\chi_\mu^{pp}\}$. The orbitals are virtually indistinguishable up to 150 eV.

3.3 Applications

Now that we have tested PP/S for each element we can apply them to the general environments. For calculations on crystals we use the Gaussian dual-space (GDS/DFT) approach [15] and use the decontracted valence atomic basis $\{\chi_\mu\}$ of Hay-Wadt. [6] The Ceperley-Alder exchange-correlation scheme [14] used in BHS construction [8] is used throughout.

In Table 5 we compare the crystal properties and band structures of diamond, silicon, and germanium calculated by using BHS/S with those by using BHS. The results agree very well. Therefore the transferability of $V^{BHS/S}$ should be as good as V^{BHS}.

For the cases hitherto unsuccessful, [8] we calculated the crystal structure and the band structure for CdTe. Again BHS/S works successfully. The result is reported in Figure **??**, and structural parameters in Table 6.

Using the BHS/S, Chen *et al.* [15, 16, 17] carried out extensive calculations on the electronic state of II-VI and III-V and IV-IV semiconductors. They calculated the bulk properties, surface reconstructions and interface electronic properties (e.g., heterojunction band offsets). The results show that the BHS/S is accurate in all the cases tested. Therefore we believe that the PP/S approach proposed in this paper is transferable.

3.4 Discussion

Since the V_{cor}^{pp} is short ranged, decaying exponentially away from the nucleus, it is particularly simple to design a Gaussian basis set to form a series of wave packets for its proper representation. We demonstrate that with just three Gaussians one can achieve an accuracy better than 0.1eV for the orbital energy and total energy for most elements. Basis set optimized for each individual element could give even higher accuracy. This makes quite practical calculations on atoms, molecules, and solids.

Improper sampling of the real space in 3.21 could cause the "ghost-state" [12, 13, 14, 21, 22] of higher energy falling into or below the reference valence states. Figure 3.1 shows that the space adequate for representing the V^{BHS} is quite different from the space adequate for representing the orbitals $\phi_i(r)$. V^{BHS} is large in the core region while the orbitals $\phi_i(r)$ are large in the valence region, generally outside the core region. To represent V_{cor}^{nl} one needs tight Gaussian functions (e.g., $\bar{\alpha}_0 \sim 0.75$ for Fe) for the wave packets to cover the core region.

Previous workers employed pseudo-orbitals to represent V_{cor}^{nl}. For example in Kleinman and Bylander's form [11] the PP/S is written as,

$$V_{KB}^{nl} = \frac{|V^{nl}\phi_{val}\rangle\langle\phi_{val}V^{nl}|}{\langle\phi_{val}|V^{nl}|\phi_{val}\rangle}. \tag{3.40}$$

One can view it as using one very diffuse wave packet $|\phi_{val}\rangle$ to represent the core potential V_{cor}^{nl}, albeit in a potential weighted way. Other workers have proposed using more terms [12, 13, 14, 21, 22] in addition to $|\phi_{val}\rangle$. Our analysis indicates that one should choose the Gaussian basis such that the wave packets they form have the same resolution and cover range as required by the core potential V_{cor}^{nl}. Thus we believe that it is the poor sampling of the real space that caused the ill-representation of the

V^{nl}_{cor}, leading to the ghost-states. This is demonstrated by the $3'$ case in Figure 3.2a. Incidentally, as shown in Figure 3.9 , $V^{BHS/S}|\phi_i\rangle = V^{BHS}|\phi_i\rangle$, a condition built into Kleiman-Bylander scheme, is satisfied to good accuracy in our construction. The $V^{BHS/S}$ was constructed on the basis with $\{n\}$=(-2, -1, 0, 1, 2).

In the above tests we employed the MS4 basis set [9] to solve the atomic KS equation (Eq. 3.34). To estimate the error introduced by using this finite Gaussian basis set, we also performed PP-BHS calculations on a numerical grid. For Fe we found that the error in eigenvalues due to the finite basis set are less than 6 mH = 0.18 eV as shown in Table 6.

The accuracy of the PP-BHS in reproducing the relativistic all electron calculation is around $0.1eV$. See table 6 for calculation on Fe atom.

Therefore our compact basis $\{\chi^{pp}_\mu\}$ for BHS/S should be sufficient for most purpose. When higher accuracy is needed, one can always refine the $\{n\}$-grid and increase N_l to any desired level.

The use of Gaussian functions to represent PP/S is very efficient. It provides a systematic and efficient way of describing molecules and solids with results as close to those of non-separable PP as desired but with better scaling (linear instead of quadratic) with the size of the basis set. It also facilitates the tabulation of the pseudopotential and applications to complex systems. For programs using planewaves to calculate total energy, one can use recursion relations to build efficiently the Fourier transforms of Gaussian functions. [15] For programs using Gaussian basis sets, the standard two-center integrals can be used. [15]

3.5 Conclusion

We conclude that general transferable pseudopotentials for all the atoms of periodic table can be constructed in a separable form that makes the computational cost linear in size of the basis set. We achieve that by using a series of wave packets formed from a Gaussian basis set to represent the non-local part of the pseudopotential. We have demonstrated here some successful applications to the BHS pseudopotential. [8] However, the construction of equation (Eq. 3.29) is general and can be applied to any *ab initio* pseudopotential. [29]

Acknowledgments

We thank Dr. G. Bachelet for helpful discussions on the BHS pseudopotentials. We also thank Dr. M. Scheffler for making available to us the report (reference 8) on their studies of separable pseudopotentials.

Bibliography

[1] J. C. Phillips and L. Kleinman, Phys. Rev. **116** 287, (1959)

[2] W. A. Goddard III, Phys. Rev. **182**, 48 (1968) L. R. Kahn and W. A. Goddard III, Chem. Phys. Lett. **2**, 667 (1968)

[3] C. F. Melius, W. A. Goddard III, J. Chem. Phys. **56**, 3348 (1972)

[4] A. Redondo, W. A. Goddard III, and T. C. McGill, Phys. Rev. B **24**, 6135, (1981)

[5] A. K. Rappé, T. A. Smedley, and W. A. Goddard III, J. Phys. Chem. **85**, 1662 (1981)

[6] P. J. Hay and W. R. Wadt, J. Chem. Phys. **82**, 270 (1985); J. Chem. Phys.**82**, 2663 (1985).

[7] D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[8] G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).

[9] A. Redondo, W. A. Goddard III, and T. C. McGill, Phys. Rev. B **15**, 5038, (1977).

[10] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[11] L. Kleinman and D. M. Bylander Phys. Rev. Lett. **20**, 1425, (1982).

[12] A. K. Rappé and W. A. Goddard III, "A Balanced Basis Set for All-Electron Ab Initio Electronic Calculations on Atoms H through Xe. to be published.

[13] X. J. Chen, J-M. Langlois, and W. A. Goddard III, Phys. Rev. B, **52**, 2348, (1995)

[14] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[15] X. J.Chen, X. Hua, J. Hu, J-M. Langlois, and W. A. Goddard III, Phys. Rev. B **53**, 1377 (1996)

[16] X. J. Chen, A. Mintz, J. Hu, X. Hua, J. Zinck., and W. A. Goddard III, J. Vac. Sci. Technol. **13**, 1715 (1995)

[17] J. Hu, X. J. Chen, X. Hua, and W. A. Goddard III, "Properties of Group IV, III-V, and II-VI Semiconductors Using Gaussian Basis Functions with Separable Pseudopotentials", paper to be submitted to Phys. Rev. B15.

[18] P. Blöchl, Phys. Rev. B **41**, 5414 (1990)

[19] D. H. Vanderbilt, Phys. Rev. B **41**, 7892 (1990).

[20] X. Gonze, R. Stumpf, and M. Scheffler, Phys. Rev. B **44**, 8503 (1991).

[21] R. Stumpf, X. Gonze, and M. Scheffler, Fritz-Haber-Institut Research Report No.1, April, 1990 (unpublished).

[22] M. Saito, O. Sugino, and A. Oshiyama, Phys. Rev. B **46**, 2606 (1992).

[23] Jean-Marc Langlois, *Ph.D. Thesis* California Institute of Technology (1993).

[24] M. T. Yin and M. L. Cohen, Phys. Rev. B **26**, 5668 (1982).

[25] V. Fiorentini, Phys. Rev. B **46**, 2086 (1992).

[26] S-H Wei and A. Zunger, Phys. Rev. B **37**, 8958 (1988).

[27] B. Segall, M. R. Lorentz, and R. E. Halsted, Phys. Rev. **129**, 2471 (1963).

[28] D. G. Thomes, J. Appl. Phys. **32**, 2298 (1961).

[29] X. Hua, J-M. Langlois, X. Chen, and W. A. Goddard III, "Separable *Ab Initio* Effective Core Potentials for Calculations on Large Molecules and Solids", to be published.

Table 3.1: The basis set for Fe (MS4 decontracted) using standard normalized Gaussians, equation (Eq. 3.20). The exponents α are listed below.

s	p	d
8612.00	374.500	18.5300
1299.00	87.2800	4.7960
293.000	26.6700	1.3460
79.7500	8.96500	0.3061
18.3400	2.69100	(0.0696)[a]
7.08400	0.82530	(0.0158)[a]
2.07200	0.12790	
0.74570	0.03221	
0.096711		
0.035601		

[a] Added to describe the 4d orbitals.

Table 3.2: The excited configuration total energies and orbital energies. The number in parenthesis are from separable potential BHS/S, using $\{n\}$=(-2,-1,0,1,2).

Occupation	E_{tot}	$\{\epsilon\}$
$(3d)^6(4s)^2$	-20.866037 (-20.866045)	-.273556 (-.273535) (d) -.202323 (-.202321) (s)
$(3d)^7(4s)^1$	-20.867995 (-20.868008)	-.151986 (-.151994) (s) -.089994 (-.090003) (d)
$(3d)^6(4s)^1(4p)^1$	-20.712608 (-20.712606)	-.336168 (-.336168) (d) -.242227 (-.242227) (s) -.084465 (-.084465) (p)
$(3d)^6(4s)^1(5s)^1$	-20.601885 (-20.601925)	-.401898 (-.401999) (d) -.290452 (-.290513) (s) -.122675 (-.122718) (p) -.009380 (-.009448) (s)
$(3d)^5(4s)^2(5s)^1$	-20.536170 (-20.536209)	-.551689 (-.551548) (d) -.284714 (-.284691) (s) -.105729 (-.105708) (p)

Table 3.3: Comparison of the results of using V^{dir} and V^{vw} (in brackets) for the generic basis set (Eq. 3.39). Errors are in mH(1mH=0.027eV).

Element	δE_{tot}	$\delta \epsilon_{occ.}$	$\delta \epsilon_{1H}$
K	0.004(0.000)	0.095(0.001)	2.423(0.211)
Rb	0.006(0.000)	0.062(0.000)	2.881(0.046)
Cs	0.790(0.000)	0.805(0.001)	3.287(0.078)
Fe	0.131(0.002)	0.069(0.000)	0.069(0.001)
Ru	0.049(0.001)	0.119(0.001)	0.057(0.004)
Os	(0.100)	(0.014)	(0.023)
Zn	1.083(0.005)	0.286(0.001)	0.286(0.012)
Cd	0.171(0.265)	0.265(0.104)	0.303(0.104)
Hg	46.991(1.111)	3.122(0.028)	0.473(0.028)
C	5.839(0.042)	0.415(0.011)	2.920(0.057)
Si	3.742(0.006)	1.222(0.002)	2.831(0.008)
Ge	2.318(0.012)	0.829(0.004)	0.829(0.014)
Pb	0.841(0.029)	0.224(0.016)	1.479(0.023)
F	0.115(0.046)	0.956(0.032)	0.990(0.032)
Cl	1.985(0.006)	0.839(0.002)	1.426(0.013)

Table 3.4: Atomic total energy calculated with PP/S (VW method) using the compact basis $\{\chi_\mu^{pp}\}$ according to the recipes. $\alpha_0 = 0.75$, $\beta = 3.0$, n_0 and N_l are listed after the element name as (n_0, N_l). All are in Hartrees (1H=27.2eV).

l	$E_{tot}(BHS)$	$E_{tot}(BHS/S)$	δE_{tot}
Na(-1,3)	-.183575	-.183654	.000079
K(-1,5)	-.155129	-.155143	.000014
Rb(-1,4)	-.149431	-.149431	.000000
Fe(0,3)	-20.866037	-20.866897	.000860
Ru(0,3)	-15.914191	-15.917038	.002847
Os(0,3)	-14.901081	-14.898283	-.002798
Zn(1,4)	-62.588388	-62.589035	.000647
Cd(0,3)	-46.117026	-46.118228	.001202
Hg(0,5)	-41.234831	-41.233215	.001616
C(0,3)	-5.331368	-5.329319	-.002049
Si(0,3)	-3.746177	-3.746037	-.000140
Ge(-1,3)	-3.800031	-3.800140	.000109
Pb(-1,4)	-3.466488	-3.467582	.001094
F(1,4)	-23.991309	-23.988570	.002739
Cl(0,3)	-14.906916	-14.908189	.001273
Br(0,3)	-13.391763	-13.391174	-.000589
I(-1,3)	-11.462261	-11.463525	.001264

Table 3.5: Atomic eigenvalues of the atoms by PP/S

l	$\epsilon_i(BHS)$	$\epsilon_i(BHS/S)$	$\delta\epsilon_i$
		Na(-1,3)	
0	-.102961	-.103020	.000059
1	-.031007	-.030997	-.000010
0	.037950	.037809	.000141
1	.135563	.135876	-.000313
1	1.551917	1.561645	-.009728
0	1.596508	1.581855	.014653
		K(-1,5)	
0	-.089291	-.089303	.000012
1	-.035366	-.035363	-.000003
0	.016788	.016788	.000000
2	.019179	.019184	-.000005
1	.064186	.064186	.000000
2	.167678	.167695	-.000017
0	.820342	.818101	.002241
1	.844938	.843148	.001790
2	1.079678	1.079663	.000015
		Rb(-1,4)	
0	-.086760	-.086758	-.000002
1	-.033788	-.033778	-.000010
2	.007506	.007508	-.000002
0	.016151	.016165	-.000014
1	.031062	.031067	-.000005
2	.053668	.053671	-.000003
1	.467364	.467196	.000168
0	.512990	.512461	.000529
0	1.970645	1.967893	.002752

l	$\epsilon_i(BHS)$	$\epsilon_i(BHS/S)$	$\delta\epsilon_i$
		Fe(0,3)	
2	-.273556	-.274005	.000449
0	-.202323	-.202805	.000482
1	-.052558	-.052233	-.000325
0	.054683	.054708	-.000025
1	.113168	.113330	-.000162
2	.718075	.717866	.000209
1	1.913817	1.791430	.122387
0	1.996359	1.881123	.115236
		Ru(0,3)	
2	-.251496	-.251416	-.000080
0	-.192792	-.193912	.001120
1	-.050677	-.050745	.000068
0	.060487	.060427	.000060
1	.158657	.157971	.000686
2	1.041678	1.041805	-.000127
0	1.413619	1.361864	.051755
1	1.533122	1.490619	.042503
		Os(0,3)	
2	-.233658	-.232049	-.001609
0	-.232744	-.231654	-.001090
1	-.052464	-.050610	-.001854
1	.095980	.096560	-.000580
0	.183877	.183760	.000117
2	.339741	.342667	-.002926
3	.478074	.479035	-.000961
1	1.304068	1.288540	.015528
0	1.540151	1.499417	.040734
3	1.548299	1.549509	-.001210

l	$\epsilon_i(BHS)$	$\epsilon_i(BHS/S)$	$\delta\epsilon_i$
		Zn(1,4)	
2	-.379287	-.379294	.000007
0	-.232140	-.232337	.000197
1	-.043564	-.043069	-.000495
0	.084192	.084045	.000147
1	.264897	.264887	.000010
2	.332551	.332692	-.000141
		Cd(0,3)	
2	-.443909	-.443394	-.000515
0	-.210222	-.210418	.000196
1	-.057348	-.056953	-.000395
0	.083138	.082515	.000623
1	.152332	.152515	-.000183
2	.328091	.328434	-.000343
		Hg(0,5)	
2	-.368390	-.368257	-.000133
0	-.258235	-.258195	-.000040
1	-.043494	-.043410	-.000084
1	.166959	.167007	-.000048
0	.209947	.208641	.001306
2	.447199	.447318	-.000119
1	1.695909	1.694259	.001650

l	$\epsilon_i(BHS)$	$\epsilon_i(BHS/S)$	$\delta\epsilon_i$
		C(0,3)	
0	-.492305	-.493139	.000834
1	-.187658	-.187230	-.000428
1	.406738	.408004	-.001266
0	.494531	.493434	.001097
		Si(0,3)	
0	-.392984	-.393006	.000022
1	-.146700	-.146790	.000090
0	.201698	.201153	.000545
1	.322046	.321488	.000558
		Ge(-1,3)	
0	-.431131	-.431186	.000055
1	-.143287	-.143270	-.000017
2	.094402	.094437	-.000035
0	.168538	.168368	.000170
3	.214043	.214068	-.000025
1	.295014	.295020	-.000006
2	.491271	.491282	-.000011
3	.809618	.809654	-.000036
		Pb(-1,4)	
0	-.439623	-.440162	.000539
1	-.133975	-.133987	.000012
1	.169952	.169985	-.000033
0	.240345	.240369	-.000024
2	.318891	.318753	.000138
3	.451783	.451742	.000041
3	1.288811	1.288767	.000044
2	1.827080	1.824356	.002724

l	$\epsilon_i(BHS)$	$\epsilon_i(BHS/S)$	$\delta\epsilon_i$
		F(1,4)	
0	-1.063674	-1.064174	.000500
1	-.383775	-.383716	-.000059
1	.972455	.972812	-.000357
0	1.410764	1.410274	.000490
		Cl(0,3)	
0	-.752061	-.752467	.000406
1	-.307799	-.307598	-.000201
0	.537648	.534915	.002733
1	.672068	.672105	-.000037
		Br(0,3)	
0	-.737849	-.737796	-.000053
1	-.286862	-.287051	.000189
0	.322350	.325912	-.003562
1	.514578	.514971	-.000393
		I(-1,3)	
0	-.641746	-.642510	.000764
1	-.262270	-.262178	-.000092
0	.242973	.242119	.000854
1	.344502	.344693	-.000191

Table 3.6: Lattice parameters (a), bulk modulus (B), and band gap[d] (Eg) for diamond, silicon, and germanium.

		Exper.	BHS/S	Non-Sep
Diamond	a(A)	3.567	3.573	3.53[a]
	B(GPa)	443	465	473[a]
	Eg(eV)[d]	5.5	4.06	4.05[a]
Silicon	a(A)	5.430	5.420	5.45[b]
	B(GPa)	101.2	97.3	98[b]
	Eg(eV)[d]	1.13	0.52	0.52[b]
Germanium	a(A)	5.658	5.659	5.56[c]
	B(GPa)	77.2	66.9	76[c]
	Eg(eV)[d]	0.76	0.28	0.32[c]

[a] Reference [23].

[b] Reference [2].

[c] Reference [25].

[d] Band gap calculated at the experimental lattice constants.

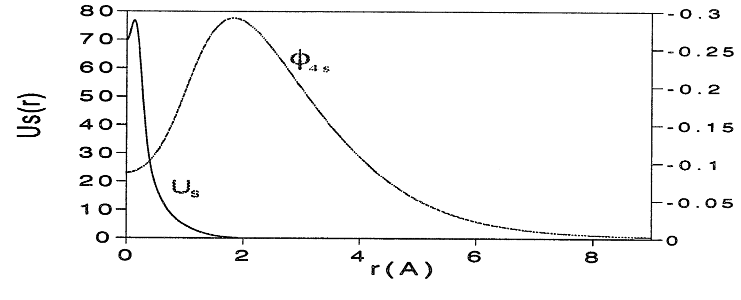


Figure 3.1: The radial part of V^{BHS}, $U_s(r)$, and the radial part of $4s$ pseudo-orbitals $\phi_{4s}(r)$ for Fe atom.

Table 3.7: CdTe structure and band parameters. Band structure calculations used the calculated lattice constants.

	LAPW[a]	Present	Exper.
a(A)	6.47	6.478	6.480[b]
Eg(eV)	0.47	0.58	1.59[c]

	LAPW[a]	Present	All-electron[d]
Γ_{1v}	-11.30	-11.21	-10.32
Γ_{15d}	-8.43	-8.20	-8.24
Γ_{12d}	-8.17	-7.87	-7.66
Γ_{15v}	-0.00	-0.00	-0.00
Γ_{1c}	0.47	0.58	1.72
Γ_{15c}	4.48	4.53	4.82
Γ_{15d}	-8.43	-8.20	-8.24
χ_{1v}	-10.79	-10.67	-9.84
χ_{3v}	-4.44	-4.42	-4.11
χ_{5v}	-1.92	-1.91	-1.93
χ_{1c}	2.45	2.44	2.68
χ_{3c}	2.54	2.67	3.41
L_{1v}	-10.91	-10.80	-9.92
L_{1v}	-4.54	-4.55	-4.37
L_{3v}	-0.80	-0.79	-0.79
L_{1c}	1.60	1.66	2.54

[a] Reference [26].
[b] Reference [27].
[c] Reference [28].
[d] Reference [15].

Table 3.8: Comparison of eigenvalues for Fe atom from solving the non-relativistic KS equation. Listed are result from numerical grid all electron calculation, from MS4 Gaussian basis for all electron calculation and that of PP-BHS (nonseparable form) on MS4 basis.

l	AE(grid)	AE(MS4)	BHS(MS4)
4s	-0.201822	-.199072	-.202323
3d	-0.279096	-.293823	-.273556

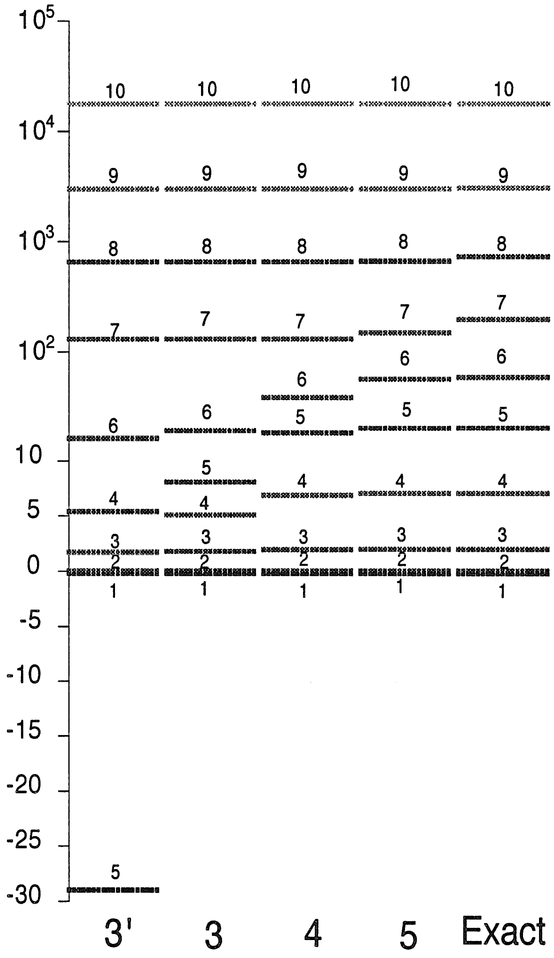
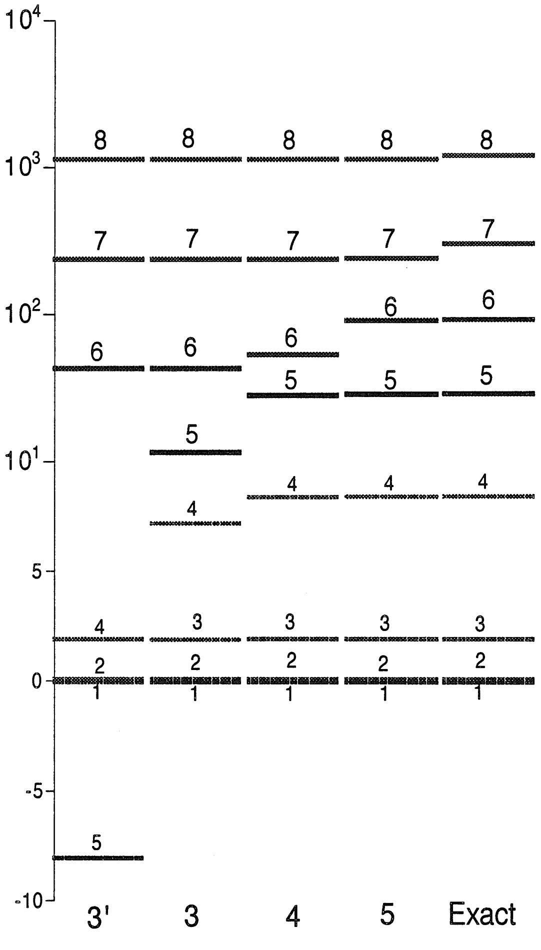


Figure 3.2: The s-eigenvalues of atom Fe KS pseudoorbitals from PP/S represented on various basis. $3'$ is from PP/S represented on the basis having exponents (0.0278/0.083/0.25); 3 is from (0.25/0.75/2.25); 4 is from (0.25/0.75/2.25/6.75); 5 is from (0.25/0.75/2.25/6.75/20.25); *Exact* is from the non-separable PP-BHS. We used logarithmic scale for eigenvalues above 10 Hartree.



(b)The p-eigenvalues of atom Fe.

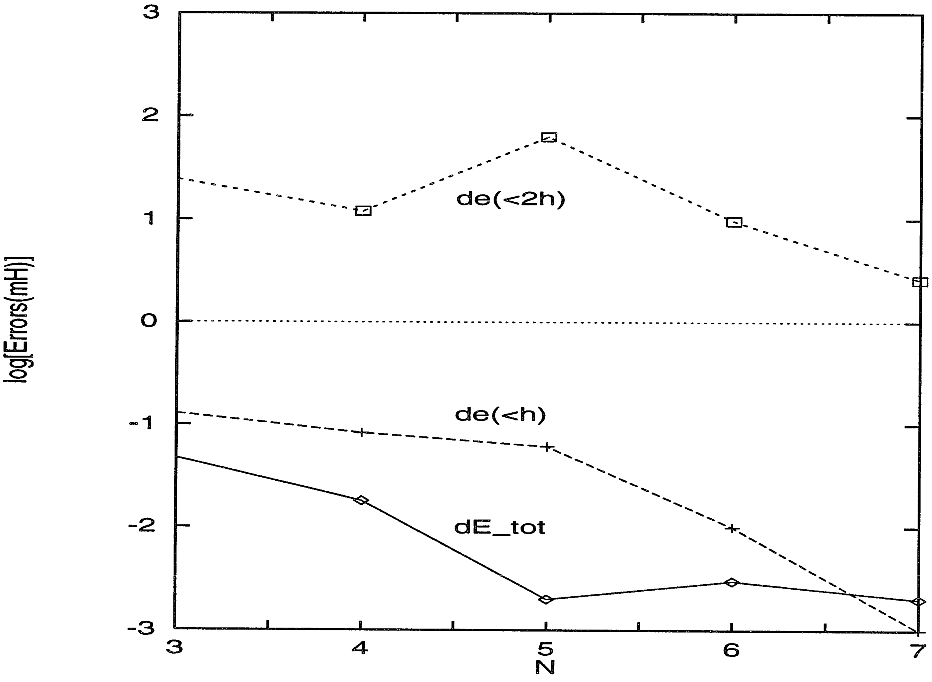


Figure 3.3: Systematic improvement of the accuracy with the increase of the basis set size on Fe atom. Plotted are the error vs. N_l, keeping $\alpha_0 = 0.75$ and $\beta = 3.0$.

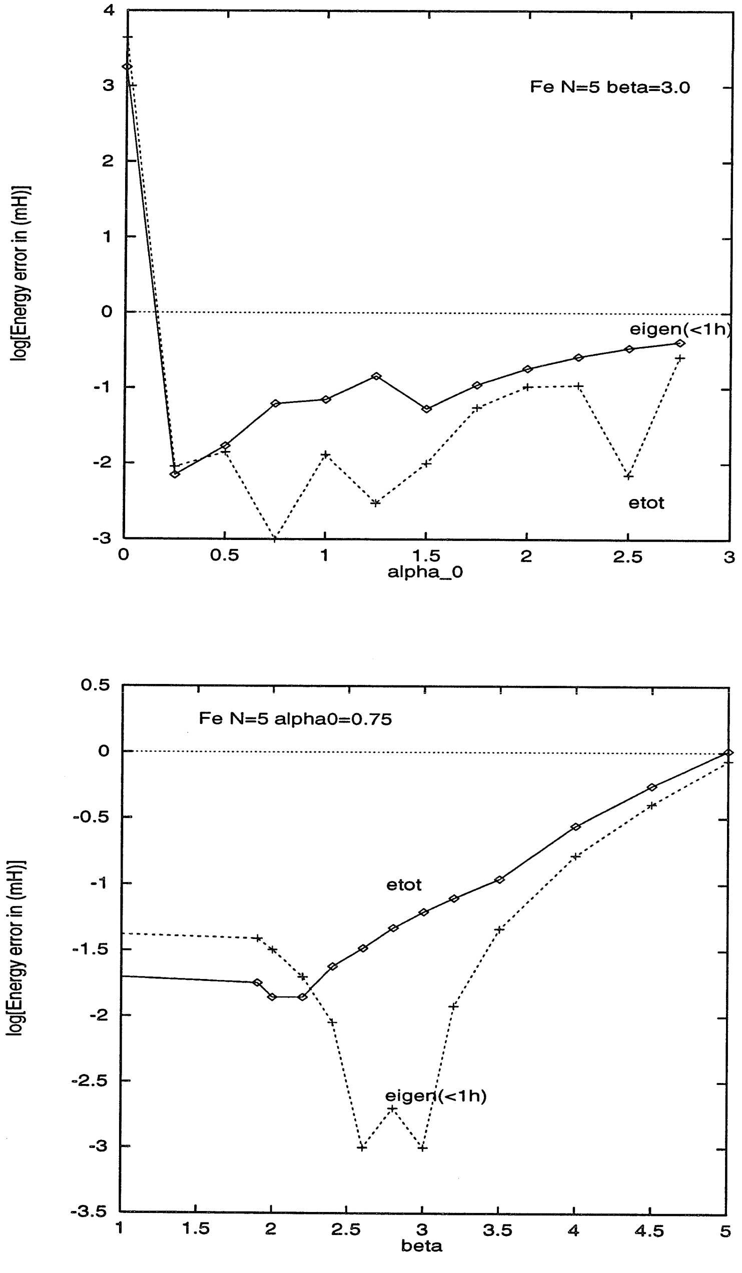


Figure 3.4: For Fe atom, the optimal even-tempered basis for PP/S, i.e. optimal α_0 and β. (a)Error in E_{tot} and eigenvalues ϵ_i plotted against the various values of α_0, fixing $\beta = 3$ and $N_l = 5$; (b)Error vs. β, fixing $\alpha_0 = 0.75$ and $N_l = 5$.

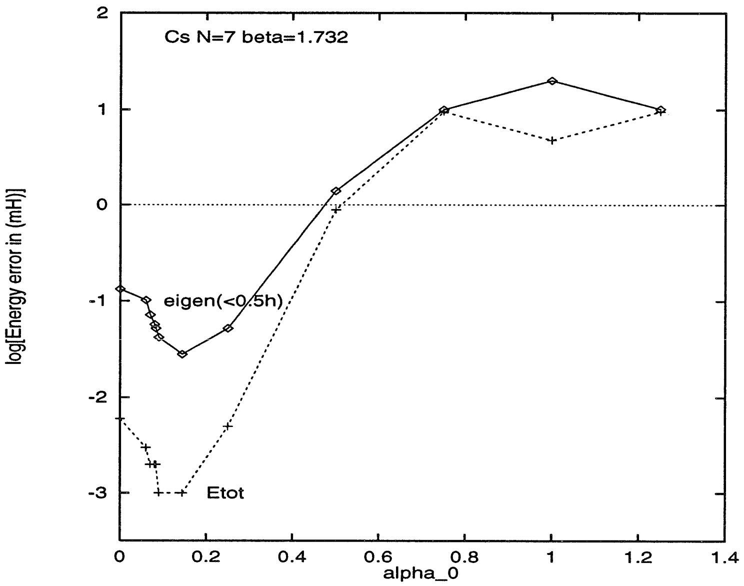


Figure 3.5: For Cs atom finding the optimal α_0. Error in E_{tot} and eigenvalues ϵ_i plotted against the various values of α_0, fixing $\beta = \sqrt{3}$ and $N_l = 7$;

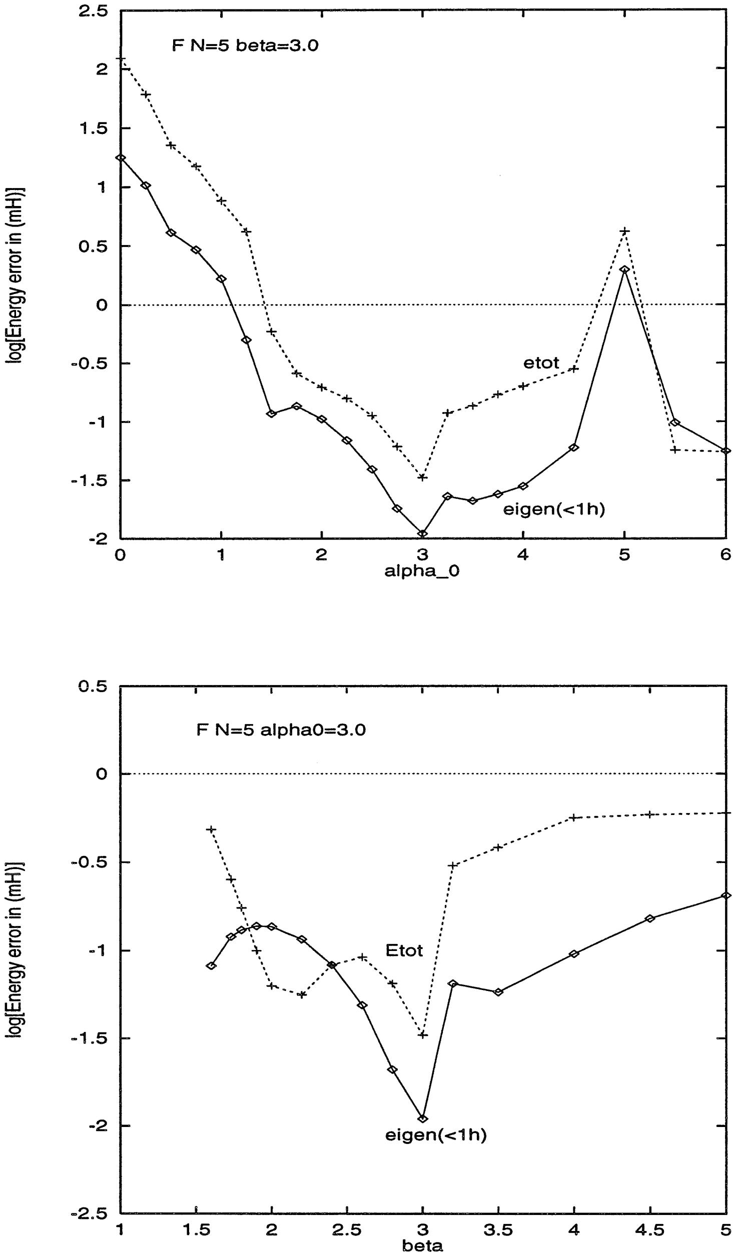
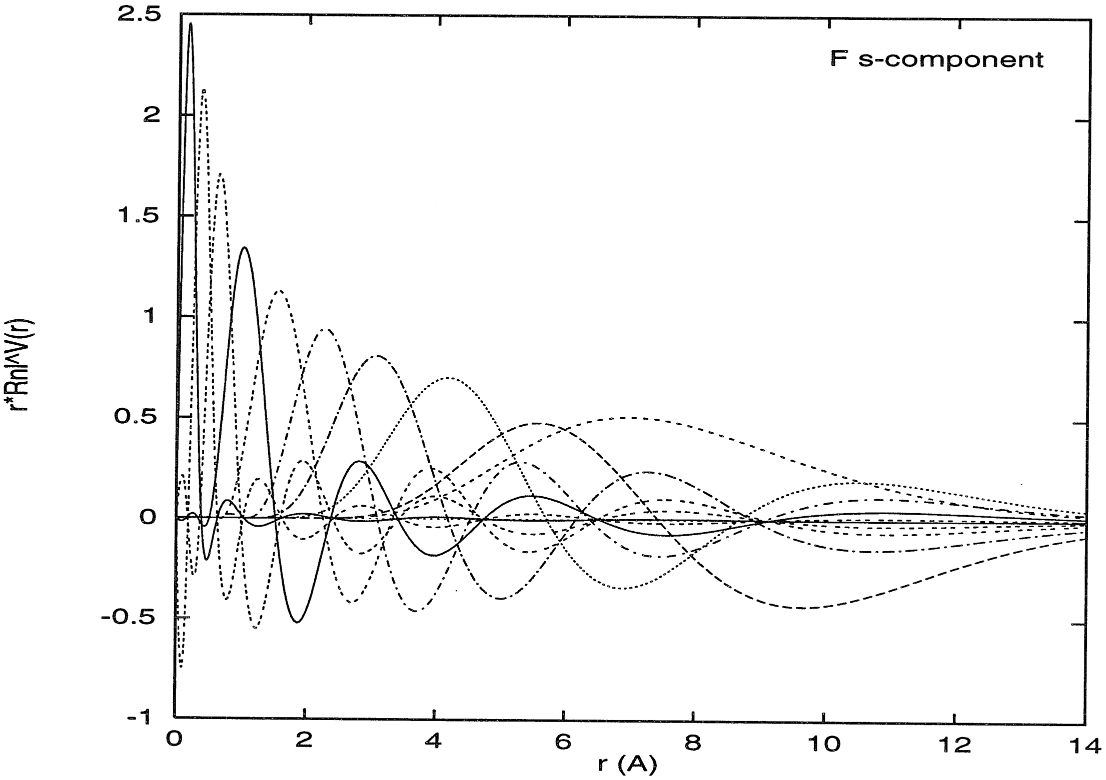
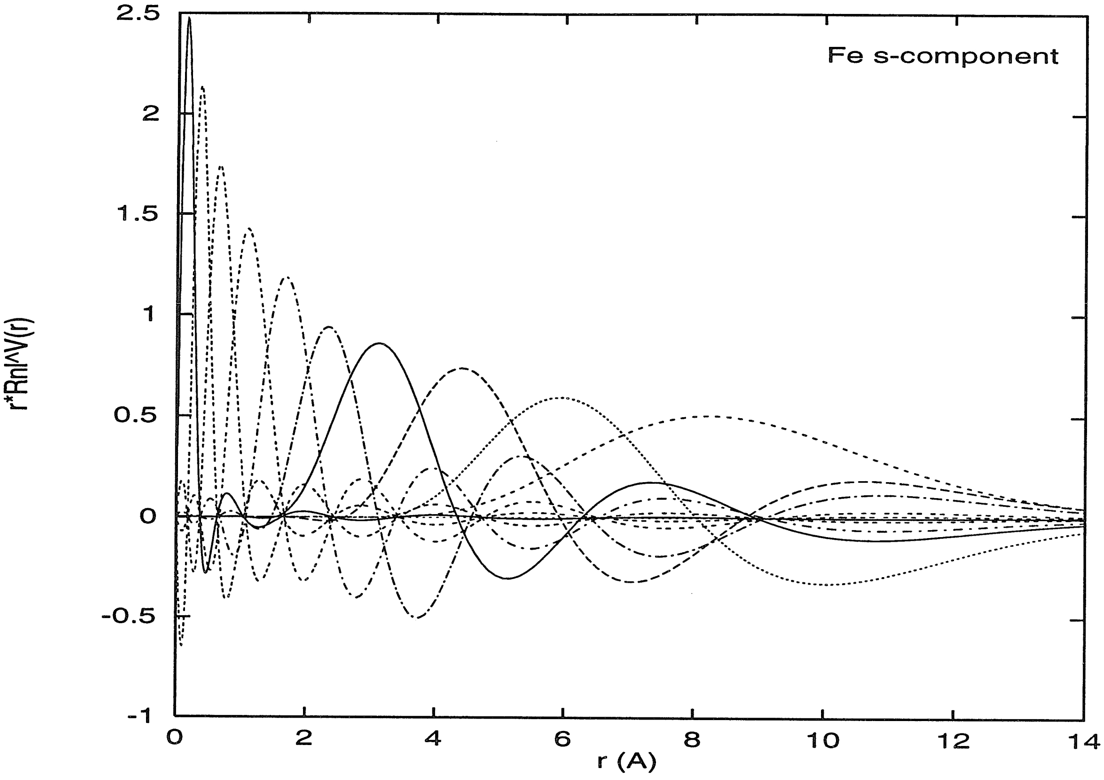


Figure 3.6: For F atom, the optimal even-tempered basis for PP/S, i.e. optimal α_0 and β. (a)Error in E_{tot} and eigenvalues ϵ_i plotted against the various values of α_0, fixing $\beta = 3.0$ and $N_l = 5$; (b)Error vs. β, fixing $\alpha_0 = 3.0$ and $N_l = 5$.

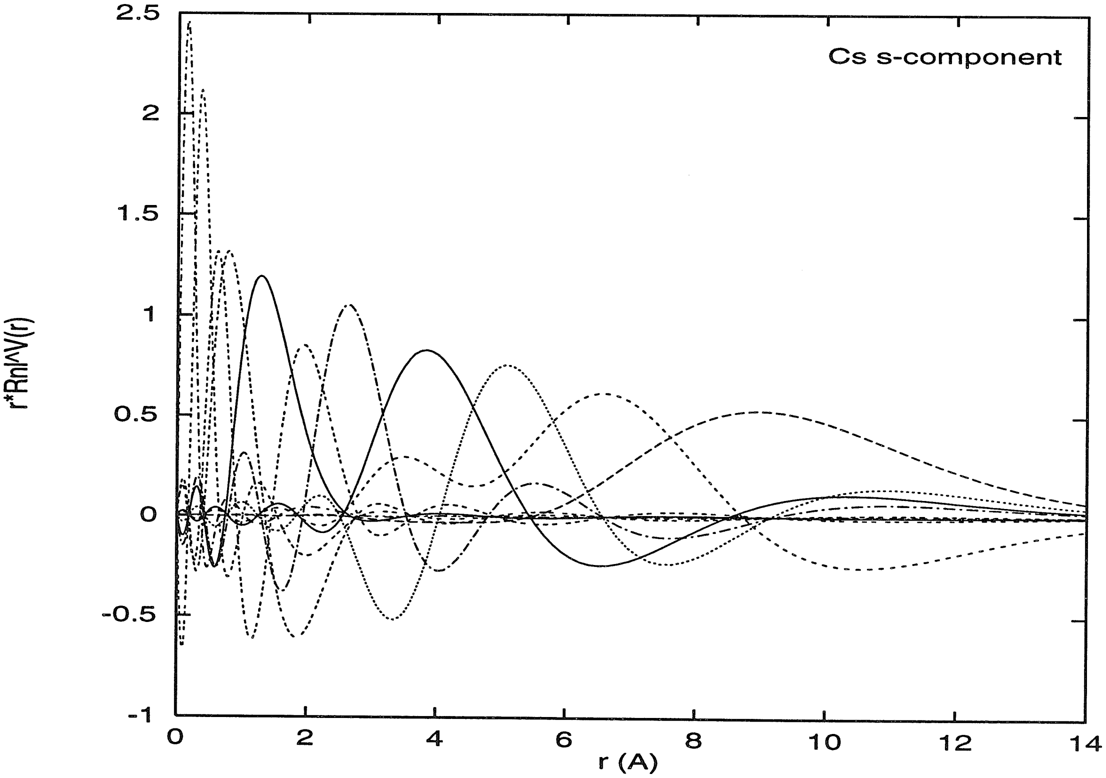


a

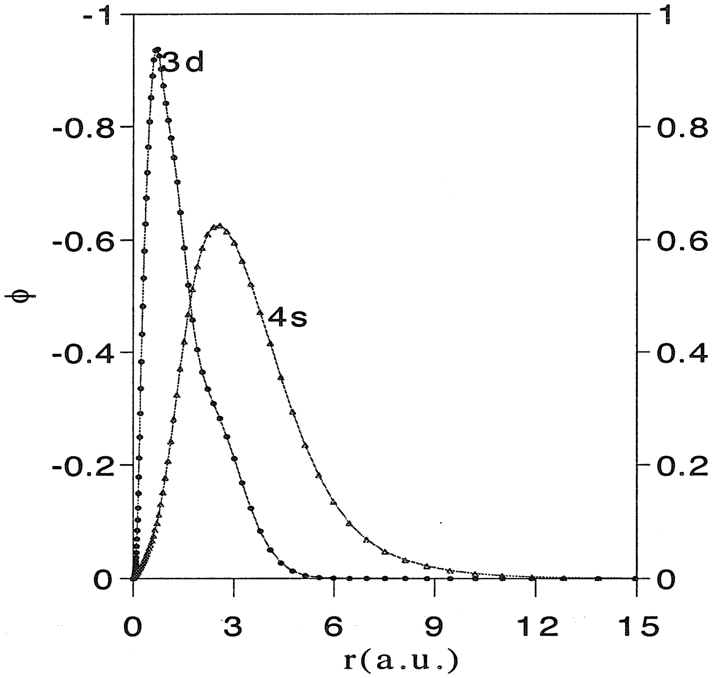
Figure 3.7: (a)F Atom $rR_{ns}(r)$, the s-type radial eigenfunctions of V^{BHS} on the universal Gaussian basis set (Eq. 3.39), i.e. with exponents $\{\alpha_n\} = \{\alpha_0\beta^n\}$. Here $\alpha_0 = 0.75, \beta = 3.0$ and n=(-3, -5/2, -2, -3/2, -1, -1/2, 0, 1, 2, 3). (b)same as (a) for Fe atom. (c)same as (a) for Cs atom.


Fe s-component
r*Rnl^V(r)
2.5
2
1.5
1
0.5
0
-0.5
-1
0
2
4
6
8
10
12
14
r (A)

b

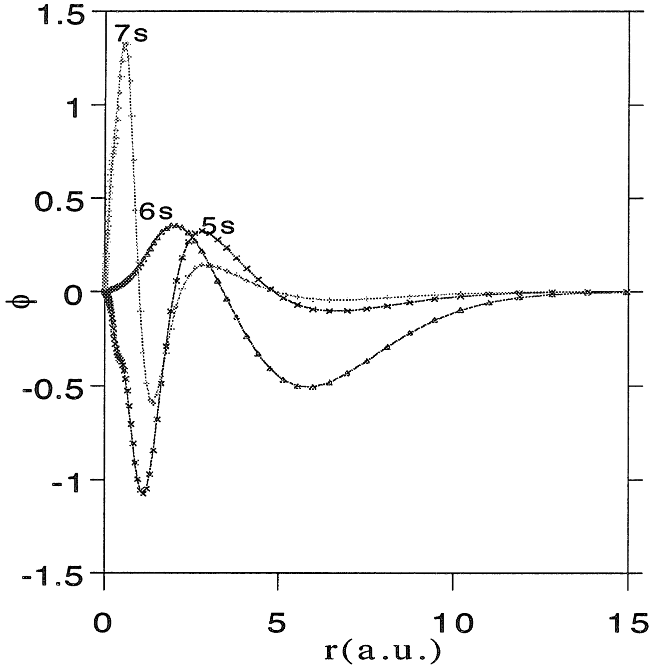

Cs s-component
r*Rnl^V(r)
2.5
2
1.5
1
0.5
0
-0.5
-1
0
2
4
6
8
10
12
14
r (A)

c

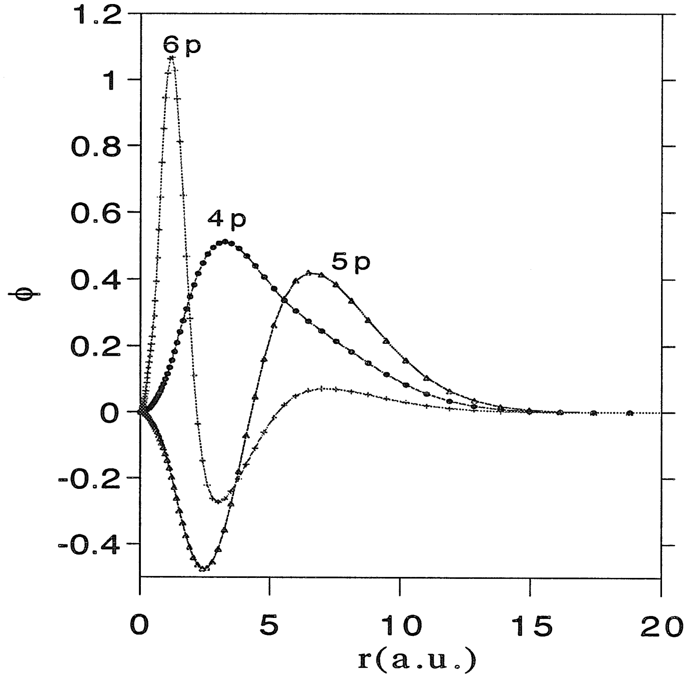


a

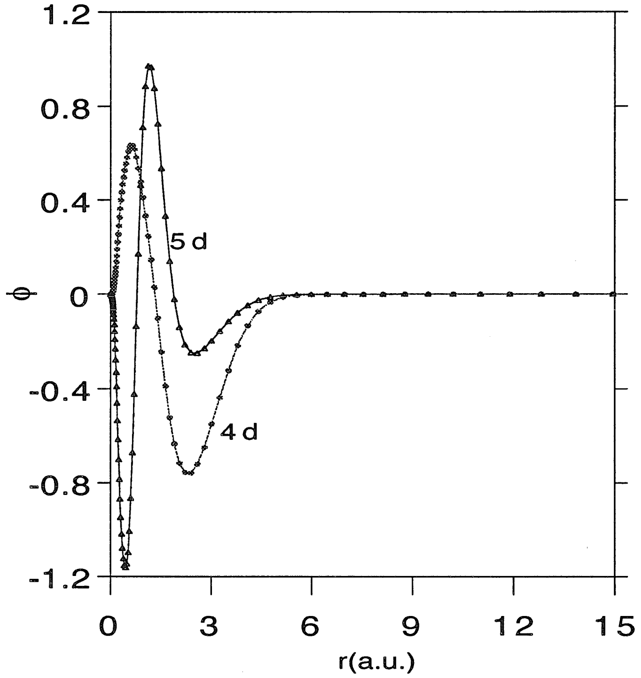
Figure 3.8: Fe atom radial orbitals $rR_{nl}(r)$ from solution of the Kohn-Sham equation on decontracted MSC basis. Solid lines are from the nonseparable BHS pseudopotentials and symbols are from the BHS/S. (a) occupied 4s and 3d orbitals; eigenvalues $\epsilon_{sep}(\epsilon_{BHS}) = -0.2023(-0.2023)$ for 4s and $-0.2735(-0.2735)$ for 3d. (b) s-type of virtual orbitals; eigenvalues $\epsilon_{sep}(\epsilon_{BHS}) = 0.0547(0.0547), 1.9728(1.9963), 6.9260(7.0582)$. (c) p-type virtual orbitals; $\epsilon_{sep}(\epsilon_{BHS}) = -0.0526(-0.0526), 0.1132(-0.1132)$, and $1.9063(1.9138)$. (d) d-type virtual orbitals; $\epsilon_{sep}(\epsilon_{BHS}) = 0.7181(0.7181), 5.2446(5.2446)$. We see that up to the 7s orbital $\epsilon_{sep} = 6.9260$, the separable and BHS nonseparable potentials are indistinguishable.


7s
6s
5s
1.5
1
0.5
0
-0.5
-1
-1.5
φ
0
5
10
15
r(a.u.)

b


6p
4p
5p
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
φ
0
5
10
15
20
r(a.u.)

c



d

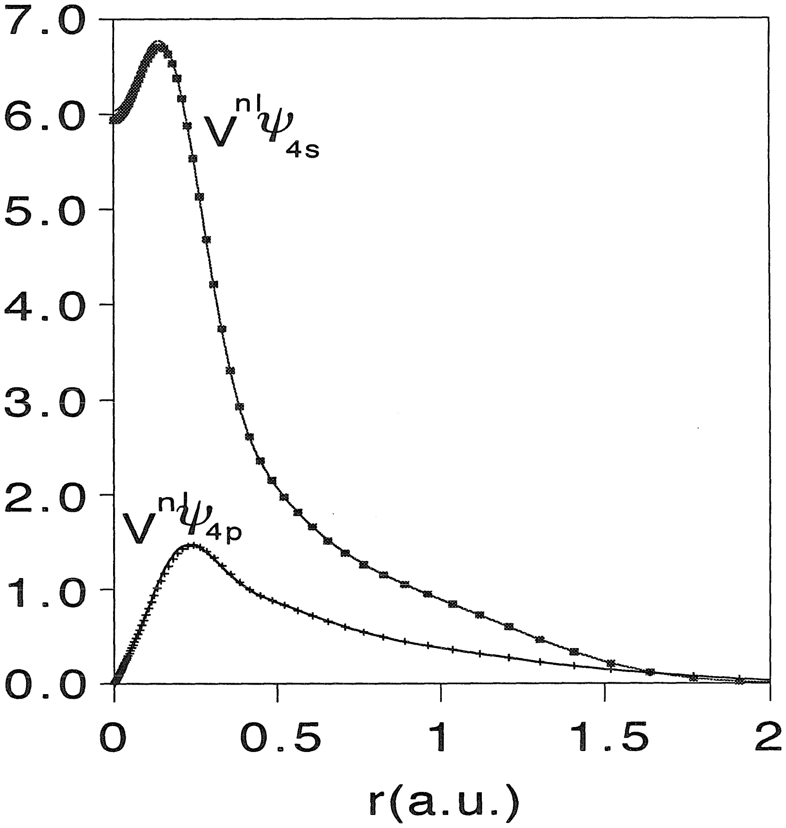


Figure 3.9: $V^{PP/S}|\phi\rangle$ compares with $V^{PP}|\phi\rangle$

Chapter 4 Separable *Ab Initio* Effective Core Potentials

4.1 Introduction

Last chapter we dealt with pseudopotentials for the density functional theory calculations. As we pointed out that the separable technique can be applied to other potentials. In this chapter we develop the *Ab initio* version. Hartree-Fock and DFT are the two main family of first principle method. (see chapter 1) Hartree-Fock based method, GVB, MP2, MP4, MCSCF, CI, etc., start from HFSCF wavefunction and then take up the residue interaction with Shrödinger-Reighley perturbation theory. While DFT map the many-electron problem into an single-particle theory and improve upon such mapping scheme. The issues in effective core in the two family are the same, sort of orthogonal to the exchange and correlation scheme of the many-electron systems. So the concept of separable potential should be implemented in a similar fashion. We will use the effective core potential for the ab initio context and pseudopotential for the DFT context. The abbreviation of the respective separable potentials are ECP/S and PP/S.

Hay and Wadt have implemented Goddard's idea of Ab initio ECP. Their implementation differ slightly from BHS's implementation of DFT PP. The central difference involves the matrix-element of $\langle\chi_{\mu B}|V_A^{nl}|\chi_{\nu C}\rangle$. The coding in ECP/S is a bit more complicated than PP/S. We derived ECP/S based on Melius [6] and implemented it in the PS-GVB program. The detailed derivation is listed in Appendix A_1.

Ab initio method can give a progressively more and more accurate result with higher and higher level of incorporation of electron correlations. Can the accuracy of our fast ECP/S live up to the standard of these high accuracy calculations? To test that we carried out GVB calculation on some very sensitive systems where HF fails to give the right structure.

In section 4.2 we outline the method and give a brief account for the matrix element evaluation. Then in section 4.3.2 we test our ECP/S for various "tough" cases with HF and GVB calculations. Finally in section 4.4 we benchmark the timing for the major component of first-principle calculations.

4.2 Methodology

4.2.1 *Ab Initio* Effective Core Potentials

The Hartree-Fock equations for an atom have the form

$$\left(t + V_{nuc} + V_{core}^{HF} + V_{val}^{HF}\right) \phi_i^{HF} = \epsilon_i^{HF} \phi_i^{HF} \tag{4.1}$$

where the Fock operator has the terms

$$t = -\frac{1}{2}\nabla^2 \tag{4.2}$$

(the kinetic energy operator),

$$V_{nuc} = -\frac{Z}{r} \tag{4.3}$$

(the nuclear attraction),

$$V_{core}^{HF} = \sum_{j=1}^{N_{core}} (2J_j - K_j) \tag{4.4}$$

which describes the field due to the core electrons (e.g. 1s, 2s, 2p, 3s, 3p for Fe), and

$$V_{val} = \sum_{j=1}^{val} (a_{ij} J_j + b_{ij} K_j) \tag{4.5}$$

describes the various coulomb and exchange interactions among the valence electrons (e.g. the 3d and 4s orbitals for Fe). The various eigenstates of (1a) are orthogonal so that the valence orbitals will have one radial nodal plane for each shell of core orbitals who have the same angular momentum (for Fe this leads to three radial nodal planes for 4s, two for 4p, one for 4d, and zero for 4f). For describing most properties of molecules and solids, the core orbitals can be considered as fixed. Thus in describing the chemistry of molecules or solids containing C, Si, Ge, Sn, or Pb, we consider that only the four valence electrons are significant. However, because of the core electrons quantum mechanical calculations of Pb are very considerably more difficult than for C.

For this reason many approaches have been developed to carry out quantum mechanical calculations in which the core electrons are replaced by effective core potentials, or ECP for short. This leads to atomic HF equations of the form

$$\left(t + V_{ecp} + V_{val}\right) \phi_i^{ecp} = \epsilon_i^{ecp} \phi_i^{ecp} \tag{4.6}$$

where the valence pseudo-orbitals, ϕ_i^{ecp} no longer need be orthogonal to the core orbitals. One approach (denoted as effective core potentials, ECP) developed by Goddard and coworkers [2] - [7] starts with the *ab initio* valence orbitals from (1a), makes them smooth in the core region to obtain ϕ_i^{ECP}, and defines the V_{ecp} as

$$V_{ecp} = \frac{1}{\phi_i^{ecp}} \left[\left(t + V_{val} - \epsilon_i^{HF} \right) \phi_i^{ecp} \right], \tag{4.7}$$

a process referred to as *inverting the orbital.* Various flavors of this method have included electron correlation effects in the core orbitals, [2, 3] accounted for changes between V_{val}^{ecp} and V_{val}^{HF}, [6] and ensured that the smoothing of the core region retains conservation of the norm. [5] Equation 4.7 is solved either numerically or by expanding in terms of various functions.

This approach leads to a general ECP of the form

$$V_{ecp} = V_{loc} + V_{nl} \tag{4.8}$$

$$V_{nl} = \sum_{l=0}^{l_{core}} |P_l\rangle U_l(r) \langle P_l|, \tag{4.9}$$

where $\langle P_l|$ is an angular momentum projector operator, and $U_l(r)$ is a radial function. Here l_{core} is generally the maximum angular momentum of the shells in the core. For angular momenta $l > l_{core}$, the potential due to the core electrons becomes V_{loc}, which far from the atom has the form $-Z_{val}/r$ (where Z_{val} is the number of valence electrons) and close to the nucleus has the form $-Z/r$ (where Z is the total charge). On the other hand, for $l \leq l_{core}$ the ECP becomes

$$V_{loc} + U_l. \tag{4.10}$$

Here U_l is very repulsive in the core region in order that the valence-like orbital ϕ_i^{ecp} go smoothly to zero while retaining the valence character.

Using equations (4.8) and (4.9) the Hartree-Fock equation for the valence orbitals become

$$\left(t + V_{ecp} + V_{val}\right) \phi_i^{ecp} = \epsilon_i^{ecp} \phi_i^{ecp} \tag{4.11}$$

where $\phi_i^{ecp} \approx \phi_i^{HF}$ outside the core and $\epsilon_i^{ecp} \approx \epsilon_i^{HF}$. The solutions of equation 4.11 are not required to be orthogonal to the core orbitals; hence, the spectrum of equation 4.11 must include resonances (ghost states) involving significant core character. In order to avoid artifacts in treating molecules and solids, it is essential that such core-like ghost states have energies far above the energies relevant for any physical excitation. It was found [3, 4] that requiring the ϕ^{ecp} to go smoothly to zero at the nucleus leads to U_l sufficiently repulsive in the core region that the ghost states were not a problem.

If relativistic effects are included [8, 6] in the *ab initio* HF equation, then the *ecp* obtained by inverting the orbital, equation 4.7, will also include relativistic effects. This is particularly important for the heavier atoms (particularly, beyond Lu) where differential shielding effects dramatically affect the shape of the valence orbitals (making *s* more stable and *d* less stable).

Based on such procedures a full set of ECP have been developed by Hay and Wadt [8] (for H through Bi) and by Christrensen *et al.* [10] (from Li through Rn (?)) and used for *ab initio* calculations for many molecules.

For molecules and solids the nuclear attraction terms in the all electric-Hamiltonian

$$V_{nuc} = \sum_A -\frac{Z_A}{R_{Ai}} \tag{4.12}$$

is replaced by

$$V_{nuc} = \sum_A V_{ecp,A} = \sum_A V_{loc,A} + \sum_A \sum_l^{l_{core,A}} |P_l^A\rangle U_{lA}(r)\langle P_l^A|. \tag{4.13}$$

The presence of angular momentum projection operators in (6) leads to three-center integrals,

$$\langle \chi_B | V_{ecp,A} | \chi_C \rangle, \tag{4.14}$$

where χ_B and χ_C are (Gaussian) functions centered on atoms B and C. For large systems this leads to costs that scale as,

$$\frac{1}{2} N_{ecp} N_{BF}^2. \tag{4.15}$$

where the number of centers with *ecp*.

In this paper we propose an approach for modifying 4.13 so as to evaluate three-center integrals while avoiding problems due to ghost states.

4.2.2 Separable Effective Core Potentials (ECP/S)

In order to reduce the costs of calculating matrix elements for 4.13, various workers [11, 12, 13, 14] have explored the use of separated potentials

$$V_{sep}^{nl} = \sum_{Am,n} |\theta_{Am}\rangle \lambda_{Amn} \langle \theta_{An}|, \tag{4.16}$$

including several functions θ_{Am} on each atom A.

This was first used for *ab initio* calculations by Huzinaga. Equation 4.16 considerably simplifies the calculations for larger systems; however, it can lead to artifacts. With projection operators, the standard method [8] automatically leads to extremely high energies for core-like states so that the eigenfunctions properly describe valence states. However with a separated potential, it is possible to obtain *ghost states.* These are the core-like states with energies comparable to the valence states. Various authors attempted to generate the separable potential, but all have limited successes. [11, 12, 13, 14].

In this section we provide a general procedure for constructing the separable potentials. We analyze the situation where ghost-states can occur and provide prescription to avoid such pathologies.

We consider a set of Cartesian Gaussian functions $\{\chi_\mu\}$ suitable for describing the pseudo-orbital valence states of an atom,

$$\phi_i^{ecp} = \sum_{\mu=1}^{N_A} C_{\mu i}\chi_\mu,$$

where ϕ_i^{ecp} is the eigenfunction of 4.11, the basis functions have the form,

$$\chi_\mu(r,\theta,\phi) = C_\mu x^m y^n z^p e^{-\alpha r^2} Y_{lm}(\theta,\phi) \tag{4.17}$$

where C_μ is a normalization constant, $m+n+p=l$, and N_a is the number of basis functions on atoms A.

A. The Direct Method (V_{sep}^{dir})

The most direct approach to defining a separable pseudopotential is to represent V^{nl} on the N-dimensional Hilbert space spanned by the atomic basis,

$$V_{sep,A}^{dir,nl} = \sum_{\mu_i\nu=1}^{N_A} |\chi_\mu\rangle V_{\mu\nu}^{dir}\langle\chi_\nu|. \tag{4.18}$$

This can be simplified by solving for the eigenstates of V_{ecp} using this basis

$$V_{ecp}^{nl}\theta_i^{nl} = \lambda_i^{nl}\theta_i^{nl} \tag{4.19}$$

$$\theta_i^{nl} = \sum_{\mu=1}^{N_A} R_{\mu i}\chi_\mu \tag{4.20}$$

which leads to

$$V_{sep,A}^{dir,nl} = \sum_{i=1}^{N_A} |\theta_i^{nl}\rangle\lambda_i^{nl}\langle\theta_i^{nl}| \tag{4.21}$$

$$V_{sep,A}^{dir} = V_{loc} + \sum_{i=1}^{N_A} |\theta_I^{nl}\rangle\lambda_i^{nl}\langle\theta_i^{nl}|. \tag{4.22}$$

The problem with 4.22 is that the potentials U_l are very short range (the size of the core) whereas the pseudo-orbitals are smooth and do not require a good description of the core orbitals. As a result the direct approach requires many additional core-like basis functions.

B. The Potential Weighted Method (V_{sep}^{VW})

An alternative is to write

$$V_{sep,A}^{nl,VW} = \sum_{i=1}^{N_A} |V^{nl}\theta_i^{nl}\rangle\frac{1}{\lambda_i^{nl}}\langle\theta_i^{nl}|V^{nl} \tag{4.23}$$

$$= \sum_{l=0}^{l_{core}}\sum_{i=1}^{N_{Al}} |U_l\theta_{li}^{nl}\rangle\frac{1}{\lambda_{li}^{nl}}\langle\theta_{li}^{nl}|U_l \tag{4.24}$$

where 4.24 emphasizes the angular momentum dependence of 4.23 and N_{Al} is the number of sets of basis functions on A with angular momentum l. Because of the repulsive nature of U_l, we expect $\lambda_i^{nl} > 0$. This leads to

$$V_{sep,A}^{VW} = V_{local} + V_{sep,A}^{nl,VW}. \tag{4.25}$$

The advantage of this potential weighted approach is that the basis set need only be adequate for describing the pseudo-orbital. In the VW approach the U_l are expanded in an analytic form which involves only tight core-like functions.

C. Comparison of V_{sep}^{dir} and V_{sep}^{VW}

Our testing shows that V_{sep}^{VW} is uniformly more accurate than V_{sep}^{dir}, both in terms of eigenvalues and total energies. In Figure 3 we consider the effect of operating on the two valence basis functions with the non-local potential $V^{ecp}\chi_\mu$, the direct separated potential $V_{sep}^{dir}\chi_\mu$, and the potential weighted potential $V_{sep}^{vw}\chi_\mu$. In each case

$$V_{sep}^{vw}\chi_\mu \approx V^{ecp}\chi_\mu \tag{4.26}$$

whereas $V_{sep}^{dir}\chi_\mu$ deviates substantially in the core region.

V_{sep}^{VW} is uniformly more accurate than V_{sep}^{dir} Our conclusion is that V_{sep}^{VW} is the better approach and we will use this approach for the balance of this paper. We will denote V_{sep}^{VW} as ECP/S.

4.3 Results

4.3.1 Atomic Eigenvalues and Orbitals

4.3.2 Molecules

To further test of the reliability of V_{sep}^{VW}, we calculated the structures of several molecules.

		ECP	ECP/S
S=1	E_{ROHF}	-227.971154	-227.971106
	R1	1.8294	1.8294
	R2	1.8294	1.8294
S=0	E_{HF}	-227.974353	-227.974304
	R1	1.8555	1.8555
	R2	1.8555	1.8555
S=0	E_{GVB}	-227.996554	-227.996554
	R1	1.8479	1.8479
	R2	1.8316	1.8316

Table 4.1: Comparison of ECP and ECP/S result on F-Pb-F molecule. ECP/S uses generic ten function to represent V^{ECP}.

$F-Pb-F$

$F-Pb-F$ is linear and has two inequivalent bond. The GVB one σ pair leads to a bond difference of $0.01633\AA$. HF gives identical bond lengths. In both HF and GVB calculations, ECP/S and ECP give identical results. In Table 4.1 we list the E_{tot} and geometries of $F-Pb-F$ calculated using $V^{ECP/S}$ and using V^{ECP} potentials lead to excellent agreement.

$Cd-Te$

The ground state of $CdTe$ is $^3\Pi$, a triplet. It has a shallow bound state of about $0.087eV = 2kcal/mol$ The bond stretch mode of the Cd-Te molecule near the equilibrium is plotted in Figure ??, where we calculated the energy using both ECP and ECP/S. The ECP/S is on the generic basis with six functions for s and p, ten functions for d-component. We can see that the ECP/S gives almost the same mode as the ECP. Even at the extremely close distance R=$1.60\AA$, the ECP/S leads to a very

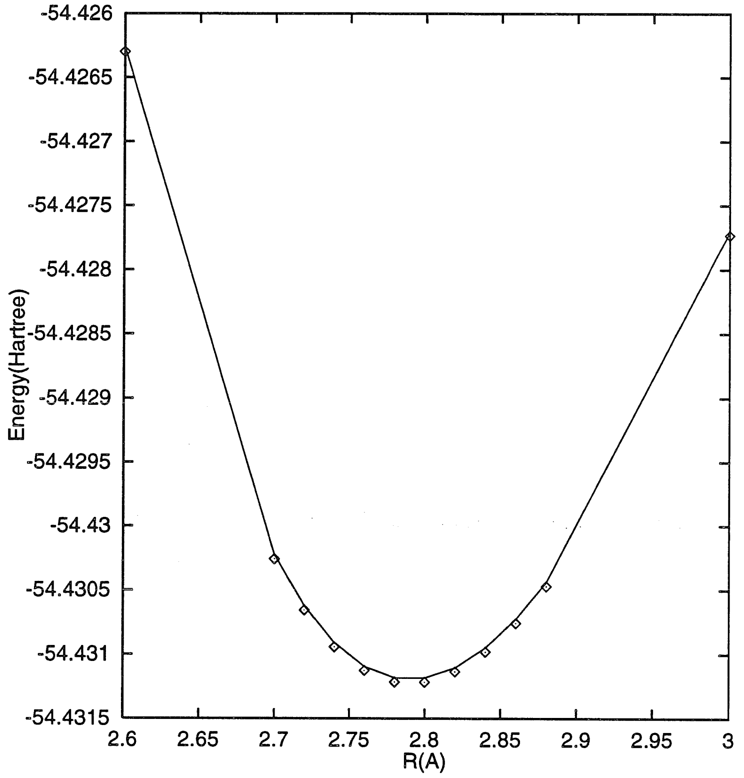


Figure 4.1: Comparing the bond stretching mode of *CdTe*, calculated with ECP/S and ECP respectively. ECP/S is on the generic-10 basis. The symbols are from ECP calculation and the line is from ECP/S.

close result as ECP. See below.

$R(\AA)$	ECP	ECP/S1[a]	ECP/S2[b]	ECP/S-generic10
2.78	-54.431215	-54.431215		-54.431266
1.60	-53.547084	-53.548001	-53.547929	-53.547139

[a]: ECP/S1 has as its basis representation of V^{nl} the following basis functions: {cdte-val, +s(2.1)for Te; +s(4.5/1.5)+p(2.4) for Cd}.

[b]: ECP/S2 has as its basis representation of V^{nl} the following basis functions: {cdte-val, +s(2.1)+p(3.6)+d(15.)for Te; +s(15.0/4.5/1.5)+p(7.5/2.4)+d(15) for Cd}.

We noticed that

> ECP/S-generic10 has the general accuracy of better than $0.1mH$. It is not tailored for any particular cases. ECP/S1 is tailor made. It is more accurate near equilibrium.(Give identical result as ECP. But its accuracy drops at a larger range to $1mH$. Nevertheless, ECP/S1 is more economic in CPU time due to the smaller basis set.

So again we face the choice of algorithmic simplicity versus run-time efficiency. Conceptually, a single basis set, the ECP/S-generic10, good for all the elements in the periodic table is very simple. But it needs ten functions for each angular momentum. On the other hand, tailor made basis for each element is more sophisticated. But, once we build a more compact basis set for each element in the periodic table, it can be used as data base to save time. Then it's the trade off between memory(for loading basis) and speed. For a modern computer, the memory should be sufficient enough to load the basis for all the elements in the periodic table.

Cd_4Te_4

The equilibrium structure of Cd_4Te_4 is show in figure 4.2. Four Cd atoms form a tetrahedral which serve as bases on which to build four tetrahedrals with the four Te atoms at each summit. Again, using the generic-10 basis for ECP/S leads to $0.1mH$

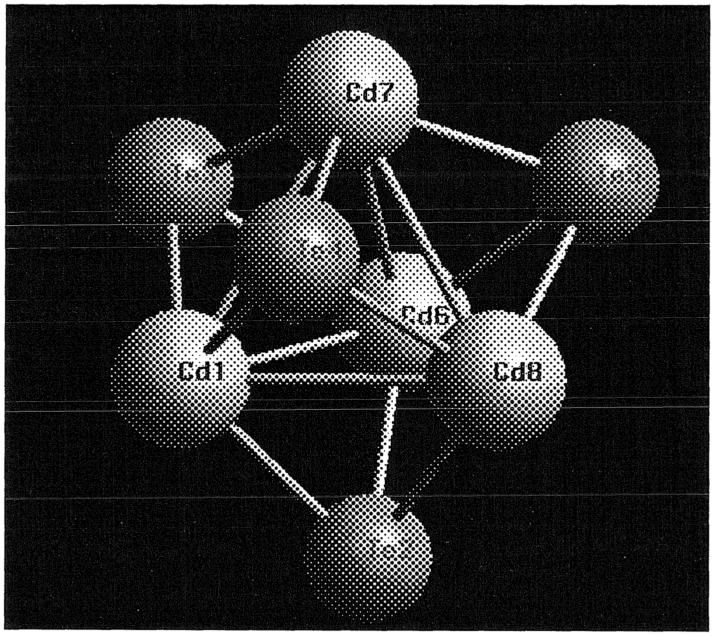


Figure 4.2: The equilibrium structure of Cd_4Te_4 calculated with either ECP or ECP/S.

Si_n	No.of Basis	T_{ECP}	$T_{ECP/S}$	T^{nl}_{ECP}	$T^{nl}_{ECP/S}$
8	72	3.540	2.560	2.370	1.390
16	144	30.720	13.030	21.660	3.970
32	288	311.020	83.320	244.680	16.980
64	576	2779.670	822.010	2062.640	104.980
128	1152	17175.949	6712.220	11310.949	847.220

Table 4.2: CPU second for calculating the ECP matrix element, $\langle\phi_i|\mathbf{V}_{ECP}|\phi_j\rangle$ for silicon cluster Si_n. The calculation is done on the 735 series HP workstations.

difference from the original ECP. ECP leads to -217.9467372 hartrees, while ECP/S leads to -217.9468706 hartree.

4.4 Benchmarking

How does the new method ECP/S fare timewise? We timed the CPU time on a HP 735 series Workstation.(120MHz clock rate and 500M memory) The numbers are listed in Table 4.2.

In Figure 4.3 we plot the nonlocal part of the matrix element cost, the most time-consuming part of ECP matrix element, in CPU second against the number of basis functions. The basis for separable ECP is four for each angular momentum. We see a saving of $13-16$ times.

As the first principle algorithms perform better and better in scaling, it can take the advantage of the hardware progress. First principle algorithm consists of several blocks in implementations.(see section 1.2.4.)

- Initial guess
- Fock matrix evaluation
    - $T[\rho]$
    - $E_{ion-ele}[\rho]$
    - $E_{Coulomb}[\rho]$
    - $E_{xc}[\rho]$

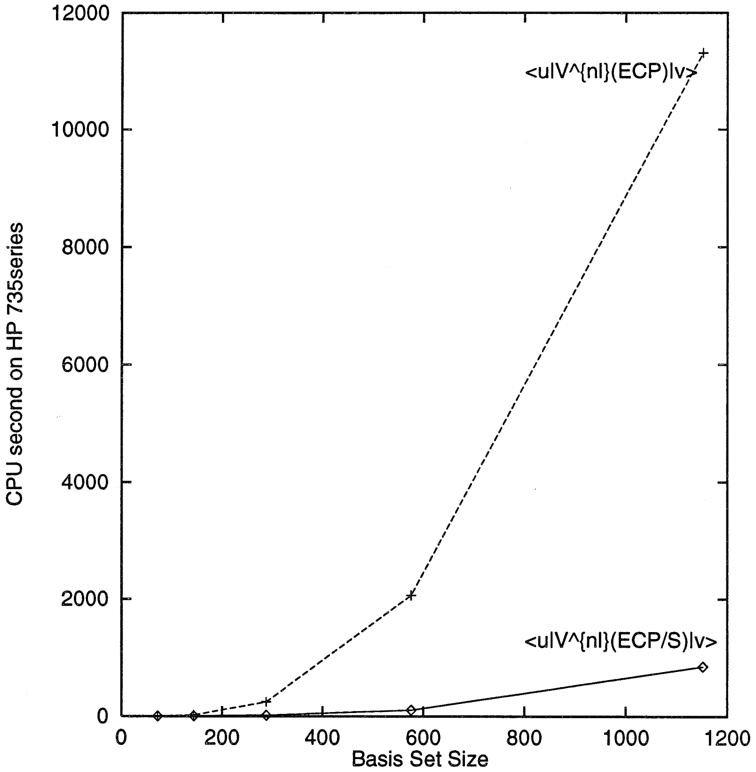


Figure 4.3: CPU second for calculating the ECP matrix element, $\langle\phi_i|\mathbf{V}_{ECP}|\phi_j\rangle$ for silicon cluster Si_n. The calculation is done on the 735 series HP workstations.

- Diagonalization
- Wavefunction update
- Forces on nucleus

When the most time consuming block, the bottleneck, is significantly reduced the second one stick out as the bottleneck. In Figure 4.4 we compare the cpu time of the whole Fock matrix calculated with the state-of-the-art DFT program (GDS/DFT) [15] on a HP workstation with our ECP matrix time cost. We can see that, in the old method of calculating ECP matrix, the non-local part alone has exceeded the total Fock matrix calculation of the new method. Obviously the advances in Coulomb term evaluation has made ECP a bottleneck. Our new ECP/S has successfully broken this bottleneck by reducing the nonlocal part, the most time consuming part of ECP, to 1/13 of its previous cost.

Acknowledgments We would like to thank Dr. Jean-Marc Langlois for assistant in porting the code to PS-GVB program.

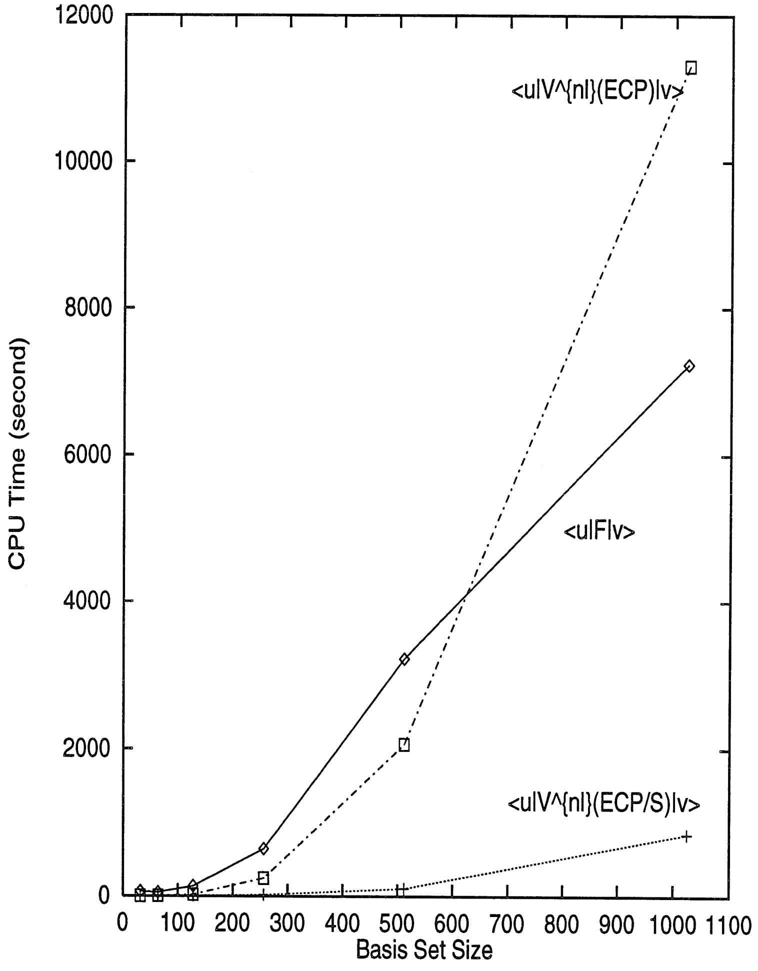


Figure 4.4: CPU second for calculating the ECP matrix element, $\langle \phi_i | \mathbf{V}_{ECP} | \phi_j \rangle$ for silicon cluster Si_n. The calculation is done on the 735 series HP workstations. The size of basis representing the ECP/S is four for each angular momentum.

Bibliography

[1] C. F. Melius, California Institute of Technology, Ph.D. Thesis, (1973)

[2] W. A. Goddard III, Phys. Rev. **182**, 48 (1968)

[3] L. R. Kahn and W. A. Goddard III, Chem. Phys. Lett. **2**, 667 (1968)

[4] C. F. Melius, W. A. Goddard III, J. Chem. Phys. **56**, 3348 (1972)

[5] A. Redondo, W. A. Goddard III, and T. C. McGill, Phys. Rev. B **24**, 6135, (1981)

[6] A. K. Rappé, T. A. Smedley, and W. A. Goddard III, J. Phys. Chem. **85**, 1662 (1981)

[7] J. C. Phillips and L. Kleinman, Phys. Rev. **116** 287, (1959)

[8] P. Jeffrey Hay and Willard R. Wadt, J. Chem. Phys. **82**, 270 (1985); J. Chem. Phys. **82**, 2663 (1985).

[9] AKR-WAG basis set, William A. Goddard III (to be published).

[10] P. A. Christiansen, Y. S. Lee, and K. S. Pitzer, J. Chem. Phys. **71** 4445 (1979).

[11] Leonard Kleinman and D. M. Bylander, Phys. Rev. Lett. **20**, 1425 (1982).

[12] Peter Blöchl, Phys. Rev. B **41**, 5414 (1990).

[13] D. H. Vanderbilt, Phys. Rev .B **41**, 7892 (1990).

[14] Xavier Gonze, Roland Stumpf and Matthias Scheffler Phys. Rev. B **44**, 8503 (1991).

[15] X. J. Chen, J-M. Langlois, and W. A. Goddard III, Phys. Rev. B, **52**, 2348, (1995)

[16] D. R. Hamann, M. Schlüter and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).

[17] G. B. Bachelet, D. R. Hamann, and M. Schlüter, Phys. Rev. B **26**, 4199 (1982).

Part II

Applications of Ab Initio Methods and MD

Chapter 5 Thermodynamics of Carbon Clusters

5.1 Introduction

Why C_{60} fullerenes are so stable? How such highly symmetric molecules are formed? These two questions are the most interesting questions in fundamental fullerene research. This chapter studies the first question. Next chapter addresses the second one.

We use here the density functional theory (DFT), as the best compromise between accuracy and speed for studying these systems. We use the Becke gradient corrected exchange and the gradient corrected correlation functional of Lee, Yang, and Parr. [1] The calculations were carried out using the software package PS-GVB with the 6-31G* basis set [15].

Carbon clusters fall into three categories. One is sp^1 bonded carbon rings, one is sp^2 bonded carbon surfaces and the third is sp^3 bonded carbon which exists mostly only in extended bulk, i.e., diamond. sp^2 bonded carbons form a two-dimensional surface, which could have a topology of a flat plane, curled up tube, or polyhedral/spherical. Several layers of such 2-D surface constitute graphite, multilayered concentric tube, multilayered concentric buckyonions. sp^1 bonded carbons form carbon chains or carbon rings. For clusters smaller than 30 carbons, the ring is the most stable structure. The interests in studying carbon rings revived due to the recent

	Dangling bond	Strain Energy
1-D	Linear	Ring
2-D	Graphitic sheet	Fullerene

results from ion chromatography experiments, which indicate the existence of large monocyclic carbon rings all the way to C_{32}. [4]

5.2 Carbon Rings

Because the carbon rings play a central role, we studied how the structures and energetics of such rings changed with size and extracted a force field (denoted as the MSX FF) that would reproduce the energetics and structures from DFT. This MSX FF would be used later in conjunction with the DFT calculations on various multiring systems to estimate the energetics of the full 60 atom systems without the necessity of DFT on the complete system.

5.2.1 DFT Calculation

We will discuss our DFT results on geometry and electronic structures. The former is from total energy calculation and latter from eigenvalues.

Geometry

The equilibrium geometries are shown in Figure 5.1 and Figure 5.2, with numbers listed in Table 5.1.

Cohesive Energy

The calculated total energies on ring systems up to C_{60} are shown in Figure 5.3. The energies quoted here are cohesive energy par carbon atom. In calculating these energies we used as our reference the triplet C atom, calculated by LSDA.

We did detailed DFT calculations on various carbon rings. Figure 5.4, shows the dimerization energies for C_{20}, C_{24}, C_{32}, C_{60}, and C_{22}. Here the energy reference is that of the non-dimerized geometry. (For C_{4m+2} this reference state is the ground state geometry, while for C_{4m} it is NOT ground state geometry.)

We found that

(i) For $n = 4m$, the minimum energy structure is a polyacetylene structure

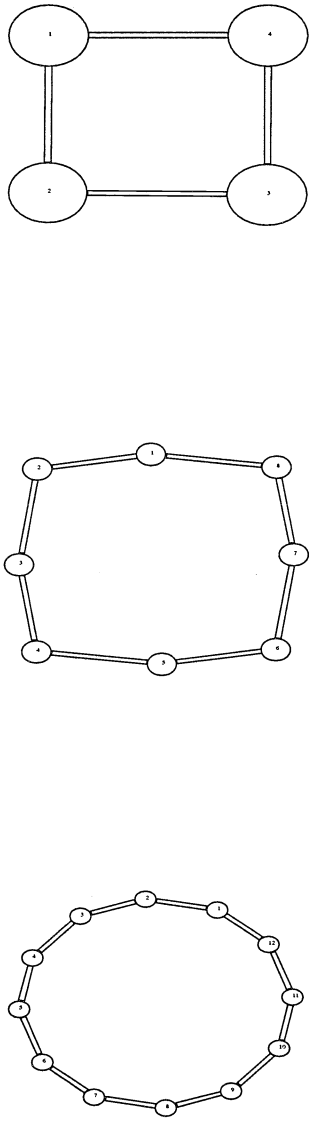


Figure 5.1: Geometries of carbon rings, C_4, C_8 and C_{12}

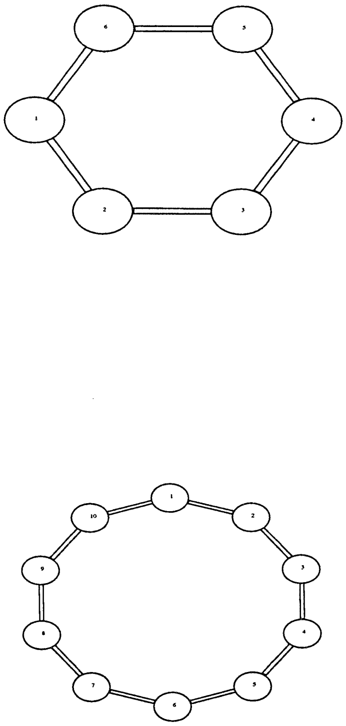


Figure 5.2: Geometries of Carbon Rings, C_6 and C_{10}

C_4		
	$R_{12} = R_{34} = 1.260$	$\theta_1 = 90.00$
	$R_{23} = R_{41} = 1.340$	$\theta_2 = 90.00$
C_8		
	$R_{12} = R_{56} = 1.275$	$\theta_1 = 163.17$
	$R_{23} = R_{67} = 1.390$	$\theta_2 = 107.68$
	$R_{34} = R_{78} = 1.277$	$\theta_3 = 163.17$
	$R_{45} = R_{81} = 1.393$	$\theta_4 = 105.98$
C_{12}		
	$R_{12} = R_{34} = ... = 1.377$	$\theta_1 = 150.00$
	$R_{23} = R_{45} = ... = 1.246$	$\theta_2 = 149.97$
C_{16}		
	$R_{12} = R_{34} = ... = 1.355$	$\theta_1 = 157.56$
	$R_{23} = R_{45} = ... = 1.250$	$\theta_2 = 157.45$
C_{20}		
	$R_{12} = R_{34} = ... = 1.343$	$\theta_1 = 162.5$
	$R_{23} = R_{45} = ... = 1.254$	$\theta_2 = 161.5$
C_6		
	$R_{12} = R_{23} = ... = 1.323$	$\theta_1 = 120.00$
C_{10}		
	$R_{12} = R_{23} = ... = 1.302$	$\theta_1 = 144.00$
C_{14}		
	$R_{12} = R_{23} = ... = 1.297$	$\theta_1 = 154.28$

Table 5.1: Geometry parameter for C_n rings

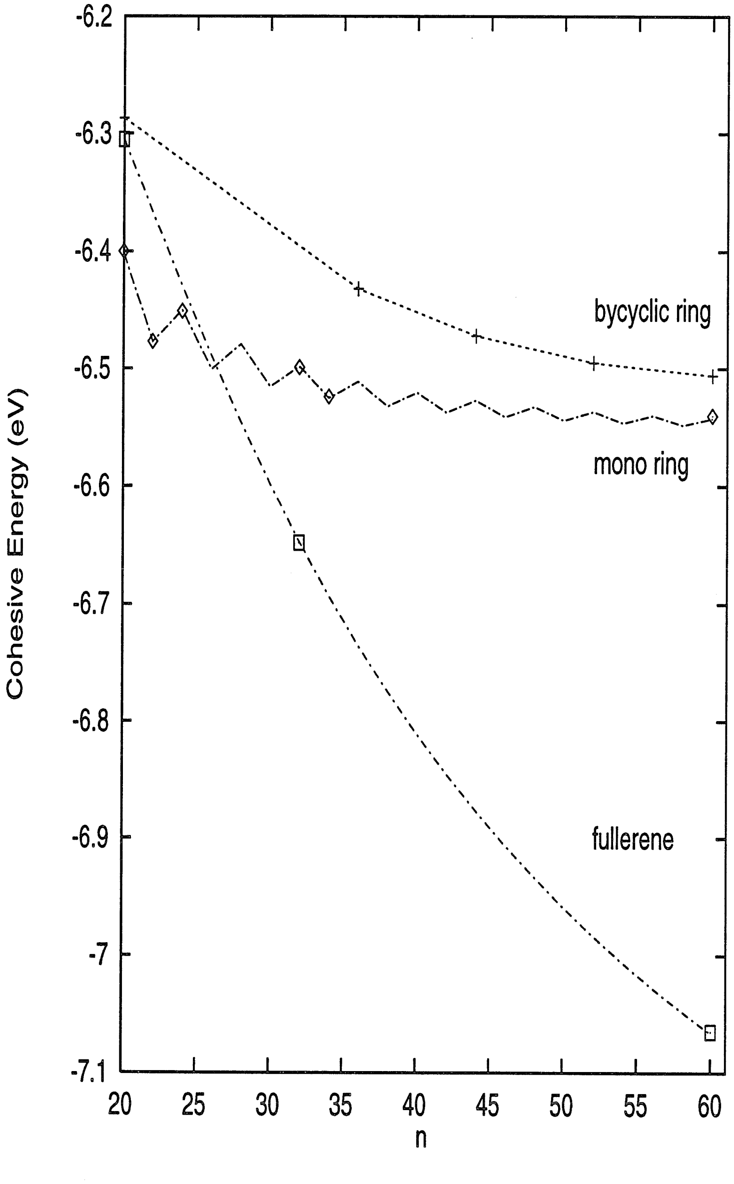


Figure 5.3: Cohesive Energy per atom of Carbon Clusters: monocyclic rings, bycyclic rings and fullerenes

of alternating single and triple bonds. The bond length difference is from $0.06\AA$ to $0.08\AA$. Note that inclusion of correlation reduced the dimerization amplitude, similar to the case in polyacetylene [4]. Comparing to DFT geometry, Hartree-Fock (HF)gives too large of a bond alternation, along with too large of angle alternations. Our HF calculation gives a bond difference of $0.16\AA$, in agreement with that of Feyereisen *et al.* [5] As for angle alternation, for C_{20}, HF gives $160° - 164°$ [6] while DFT gives $161.5° - 162.5°$

(ii) For n=4m+2, the minimum energy structure is a polyene structure of equal bond length. This is due to the resonance between the two structures, involving the π-bond perpendicular to the plane and π-bond parallel to the plane.

(iii) C_{4m+2} is more stable than C_{4m}. As $n \rightarrow \infty$, the energy differences between these two families decrease. However as shown by Figure 5.5 the geometry of the two series do *not* converge. This implies that for a very large monocyclic ring there could be spontaneous transition from polyacetyline structure to polyallene structure.

Both the polyacetylene and polyallene structures involve σ-bonds that are sp^1 hybrids, and hence they prefer linear geometries. Thus we expect a strain energy that is proportional to

$$(\delta\theta)^2 = (\pi - \theta)^2 = (2\pi/n)^2. \tag{5.1}$$

Indeed Figure 5.6 shows that the calculated cohesive energy versus $1/n^2$ leads to an excellent straight line for $n > 6$. The slope gives per carbon $E^{strain} = 63.3eV/n^2$ for n=4m and $E^{strain} = 40.1eV/n^2$ for n=4m+2. The extrapolated cohesive energy for $n \rightarrow \infty$ is $E_{\infty} = 6.56eV$

In Figure 5.7 we listed the strain energy in eV for various ring sizes.

Electronic Structures

The electronic structure for C_{60} monocyclic ring is shown in Figure 5.8, where we plotted all the Kohn-Sham eigenvalues.

The band gap from Kohn-Sham HOMO-LUMO difference is about $0.71eV$ Our approach did not use the assumption of translational symmetry, it is first principle in its true sense. From it we can see *the emergence of the quantum number* k, as the symmetry indices change from $2\pi m/n$ of C_n symmetry to k of space group symmetry.

5.2.2 The sp^1 Force Field

The harmonic force field for our system is:

$$E = E_0 + \sum_{l=1}^{N/2} \left[\frac{1}{2} k_1 q_{r1}(l)^2 + \frac{1}{2} k_2 q_{r2}(l)^2 + k_{rr'} q_{r1}(l) q_{r2}(l) + k_{rr'} q_{r1}(l+1) q_{r2}(l) \right] + \sum_{i=1}^{N} \frac{1}{2} k_\theta q_\theta(l)^2,$$

where $q_{r1}(l) = R_1(l) - R_{10}(l)$ is the bond stretching term, $q_\theta(l) = \theta(l) - \theta_0(l)$ is the bond bending term. Using PBC, $N/2 + 1 = 1$, where N is the total number of atoms in the system and N/2 is the number of unit cells. E_0 gives the energy when there's no strain, i.e., when bond length and bond angle are all at their equilibrium values.

At Γ point, i.e., when $q(l) = q(l + i)$ for an arbitrary i, we have,

$$\frac{E^{strain}}{(N/2)} = \frac{1}{2} k_1 q_1^2 + \frac{1}{2} k_2 q_2^2 + 2 k_{12} q_1 q_2. \tag{5.3}$$

Thus one can use

$$k_1 = \frac{\partial^2 [E^s/(N/2)]}{\partial q_1^2}\Big|_{q_2} \tag{5.4}$$

$$k_2 = \frac{\partial^2 [E^s/(N/2)]}{\partial q_2^2}\Big|_{q_1} \tag{5.5}$$

		C_{20}	C_{22}
Bond Stretch	$R_1^0[\AA]$	1.34	1.30
	$k_1[Kcal/mol/\AA^2]$	1424.0	902.0
	$R_2^0[\AA]$	1.26	1.30
	$k_2[Kcal/mol/\AA^2]$	2096.0	902.0
	$k_{12}[Kcal/mol/\AA^2]$	348.0	316.0
Angle bend	θ_0	180.0	180.0
	$k_\theta[Kcal/mol/rad^2]$	73.9	46.8

Table 5.2: Force Field parameter of sp^1 carbons

$$k_{12} = \frac{1}{2}\frac{\partial^2[E^s/(N/2)]}{\partial q_1 \partial q_2} \tag{5.6}$$

to derive force-field parameter k from DFT calculation of E^s for several q values for the Γ-point mode. As an example, we plot in Figure 5.9 the energies along the asymmetric stretching mode calculated with DFT.

The angle bending term k_θ can be determined directly from Figure 5.6 via,

$$\frac{\partial^2(E^s/n)}{q_\theta^2} = \frac{\alpha}{2\pi^2}(eV/rad^2) \tag{5.7}$$

where

$$\alpha = \frac{d(E^s/n)}{d(1/n^2)} \tag{5.8}$$

is the slope of Figure 5.6

The resulting force field parameters is listed in Table 5.2.

To check the accuracy of our sp^1 force field, we compare the DFT results with FF results in Figure 5.9.

5.3 Cage Structures

The optimum structure for bulk carbon is graphite in which each carbon is bonded to three others (sp^2 bonding) to form hexagonal sheets. These sheets are then stacked up on top of each other to form the layered structure. The fullerenes structures can be considered as finite two dimensional analogue of a single sheet, in which each carbon is distorted (strained) from its preferred planar configuration.

5.3.1 Topology of Fullerene

Identification and categorization of all fullerene polyhedra with a given vertex count $v = n$ is a mathematical problem. From Euler's theorem, the number of vertex v, faces f and edge f satisfy the following relationship:

$$v + f = e + 2. \tag{5.9}$$

Fullerenes are trivalent vertices and each edge has two vertices. So,

$$e = 3n/2. \tag{5.10}$$

Consequently,

$$f = n/2 + 2. \tag{5.11}$$

Let p be the number of pentagons and h be the number of hexagons. For polyhedrons containing only pentagons and hexagons, since each vertex share 3 face, we have

$$(5p + 6h)/3 = n. \tag{5.12}$$

Knowing

$$p + h = f = n/2 + 2, \tag{5.13}$$

we get $p = 12$ and $h = n/2 - 10$. Therefore

> *All fullerenes contain* 12 *five-membered rings and* $n/2 - 10$ *six-membered rings.*

Polyhedra can be flattened onto a plane while preserving the topology. This *isomorphic* transformation from 3-D to 2-D representation is one-to-many, it allows us to represent each 3-D polyhedra in a graph, like a Schlegel diagram. We build all our fullerenes according to the Schlegel diagrams of C_n fullerenes. For illustrations see figure 5.11.

The number of isomers of C_n fullerene can be estimated with the *spiral conjuncture* due to P. W. Fowler and D. E. Manolopoulos. [7]

> The surface of a fullerene polyhegron may be unwound in a continuous spiral strip of edge-sharing pentagons and hexagons such that each new face (after the second) in the spiral shares an edge with both (a) its immediate predecessor in the spiral and (b) the first face in the preceding spiral that still has an open edge.

5.3.2 DFT Calculation

We have performed the DFT calculations on several C_n fullerenes, with $n = 20, 32$ and 60.

C_{20}

C_{20} has a I_h symmetry.(see figure 5.10) It has 12 pentagon faces and 30 edges. The irreducible representation of I_h is

$$\Gamma(I_h) = A_g(1) + T_{1g}(3) + T_{2g}(3) + G_g(4) + H_g(5) + T_{1u}(3) + T_{2u}(3) + G_u(4) + H_u(5) \tag{5.14}$$

The number in the brackets indicate the dimensionality.

All 20 bonds are equivalent, thus by constraining the equal bond length there is only one variable R. However, the 120 electron will fill the orbitals which are

	$1s$			$2s$	
1	-11.674896	A_g	21	-1.736381	A_g
2	-11.674375	T_{1u}/E_{1a}	22	-1.627243	T_{1u}/E_{1a}
3	-11.674375	T_{1u}/A	23	-1.627243	T_{1u}/E_{1a}
4	-11.674375	T_{1u}/E_{1a}	24	-1.627243	T_{1u}/A
5	-11.673630	H_g/A	25	-1.448889	H_g/E_{1a}
6	-11.673630	H_g/E_{2a}	26	-1.448889	H_g/E_{2a}
7	-11.673630	H_g/E_{1a}	27	-1.448889	H_g/E_{1a}
8	-11.673630	H_g/E_{2a}	28	-1.448889	H_g/E_{2a}
9	-11.673630	H_g/E_{1a}	29	-1.448889	H_g/A
10	-11.673116	G_u/E_{1a}	30	-1.204340	T_{2u}/E_{2a}
11	-11.673116	G_u/E_{2a}	31	-1.204340	T_{2u}/E_{2a}
12	-11.673116	G_u/E_{1a}	32	-1.204340	T_{2u}/A
13	-11.673116	G_u/E_{2a}	33	-1.251001	G_u/E_{1a}
14	-11.672394	G_g/E_{1a}	34	-1.251001	G_u/E_{1a}
15	-11.672394	G_g/E_{2a}	35	-1.251001	G_u/E_{2a}
16	-11.672394	G_g/E_{1a}	36	-1.251001	G_u/E_{2a}
17	-11.672394	G_g/E_{2a}	37	-1.043429	G_g/E_{2a}
18	-11.672317	T_{2u}/E_{2a}	38	-1.043429	G_g/E_{1a}
19	-11.672317	T_{2u}/E_{2a}	39	-1.043429	G_g/E_{2a}
20	-11.672317	T_{2u}/A	40	-1.043429	G_g/E_{1a}

grouped according to the irreducible representation of I_h symmetry. As shown in Table 5.3 of the orbital energies of the I_h C_{20}^{2+} ion, the last two electrons of the neutral C_{20} will occupy a four-fold degenerate G_u orbital. As a result the system will undergo a Jahn-Teller distortion, which would lower the total energy E_{tot}. For a Jahn-Teller distorted structure, the last two orbitals (number 60 and 61) are still degenerate. Thus the last two electrons would each occupy one of the two, leading to the $S = 1$ state. In our Hartree-Fock calculations of C_{20} fullerene, the neutral $S = 0$ structure gives $E_{tot} = -756.511163$ *Hartree* while the neutral $S = 1$ structure gives $E_{tot} = -756.529333$ *Hartree*, all done at non-distorted HF minimum R with 6-31G** basis set. The triplet is lower by $\sim 0.5eV$. The Jahn-Teller distortion energy is about $\sim 1.32eV$, calculated from the difference between the constrained minimization w.r.t R and the relaxed minimization.

C_{32}

	occupied			unoccupied	
41	-1.016589	A_g	60	-.362728	G_u/E_{1a}
			61	-.362728	G_u/E_{1a}
42	-.978571	H_g/E_{1a}	62	-.362728	G_u/E_{2a}
43	-.978571	H_g/E_{2a}	63	-.362728	G_u/E_{2a}
44	-.978571	H_g/E_{2a}			
45	-.978571	H_g/E_{1a}	64	-.243398	T_{2u}/A
46	-.978571	H_g/A	65	-.243398	T_{2u}/E_{2a}
			66	-.243398	T_{2u}/E_{2a}
47	-.890786	H_u/E_{2a}			
48	-.890786	H_u/A	67	-.185903	G_g/E_{1a}
49	-.890786	H_u/E_{2a}	68	-.185903	G_g/E_{2a}
50	-.890786	H_u/E_{1a}	69	-.185903	G_g/E_{2a}
51	-.890786	H_u/E_{1a}	70	-.185903	G_g/E_{1a}
52	-.896691	T_{1u}/E_{1a}	71	-.095946	T_{1u}/E_{1a}
53	-.896691	T_{1u}/E_{1a}	72	-.095946	T_{1u}/E_{1a}
54	-.896691	T_{1u}/A	73	-.095946	T_{1u}/A
			74	.059055	T_{1g}/A
55	-.678189	H_g/E_{2a}	75	.059055	T_{1g}/E_{1a}
56	-.678189	H_g/E_{1a}	76	.059055	T_{1g}/E_{1a}
57	-.678189	H_g/A	77	.046224	H_g/A
58	-.678189	H_g/E_{1a}	78	.046224	H_g/E_{2a}
59	-.678189	H_g/E_{2a}	79	.046224	H_g/E_{2a}
			80	.046224	H_g/E_{1a}

Table 5.3: C_{20}^{+2} Hartree-Fock orbital energies in Hartrees. Geometry fixed at I_h, Basis set is 6-31G** (300 total basis), net molecular charge is +2, spin multiplicity is $1(S = 0)$. $E_{tot}^{HF}(C_{20}^{+2})$=-755.93993 Hartree.

Symbols	Symmetry	E^{strain}(Kcal/mol)
C32-6	D_3	665.1977
C32-5	D_{3h}	717.2072
C32-4	C_2	676.0504
C32-3	D_{3d}	691.2343
C32-2	D_2	691.2343
C32-1	C_2	682.9857

Table 5.4: Strain energy for the C_{32} fullerene isomers in Kcal/mol, calculated with the force field of Table 5.5.

C_{32} fullerenes have 12 pentagons, six hexagons, and a total of $n/2 + 2 = 18$ faces. According to the spiral conjuncture (section 5.3.1), there are six different C_{32} fullerenes. Their structure differ by the sequence of the pentagons positions and they have different three dimensional structures. Figure 5.11 shows the Schlegel graph of the six C_{32} fullerenes.

The bonding energy are almost the same since they all have 32 sp^2 bonded carbons. The corresponding strain energies from FF are listed in Table 5.4.

The energy differences are between $0.48eV$ and $2.3eV$. Putting pentagons adjacent to each other raises the strain energy and placing them apart lowers the strain energy. D_3 structure has the pentagons separated as far as possible. We see from our FF calculations that indeed D_3 structure has the lowest energy.

C_{60}

C_{60} has a I_h symmetry. It has 32 faces(12 pentagon faces and 20 hexagon faces). It has 90 edges(60 long and 30 short). See Figure 5.13 for the energies of various bond stretches. We found that the equilibrium bond length for C_{60} fullerene are $R_1 = 1.41\AA$ and $R_2 = 1.46\AA$. Experimental measured values are $R_1^{NMR} = 1.40\AA$ and $R_2^{NMR} = 1.45\AA$. [16]

For a summary of the energetics of carbon clusters, see figure 5.3 where the cohesive energies per carbon atom are shown. Since the strain should be proportional to the square of the planar distortion angle, $\delta\psi$, we expect that the strain energy should scale as $1/n$

		C_∞	C_{60}
VDW (LJ12-6)	$R_v(\AA)$	3.8050	3.8050
	$D_v(kcal/mol)$	0.0692	0.0692
Bond Stretch[b]	R_0 [Å]	1.411	1.414
			1.455
	k_r [Kcal/mol/$\AA^2$]	720.0	823.75
			711.77
	D_0 [Kcal/mol]	133.0	
Angle bend	θ_0(deg)	120.0	120.0
	k_θ[Kcal/mol/rad^2]	196.13	196.13
	$k_{r\theta}$[Kcal/mol/radÅ]	62.71	62.71
	$k_{rr'}$[Kcal/mol/$\AA^2$]	68.00	43.94
Torsion(1 fold)	V_t(kcal/mol)	21.28	21.28

[b] The bond stretches are Morse for graphite and harmonic for C_{60}.

Table 5.5: Graphite Force Field

5.3.3 The sp^2 Force Field

Starting from graphite FF, Y. Guo [16] has derived an accurate FF from fitting the vibration data. The FF has been applied to various crystal made up with C_{60} fullerenes and the results are excellent. [16] Here we use Guo's FF for the sp^2 carbons. We reproduced the GraFF of Y. Guo in Table 5.5

We extrapolate the cohesive energy [1] to $n \to \infty$ to get bonding energy per carbon for sp^2 carbon, $E_{coh}^{sp^2} = 7.71eV$. It can be compared to the cohesive energy of a single graphitic sheet, $E_{coh} = 7.74eV$. The latter is derived from the experimental cohesive energy [9] of graphite of $E^{graphite} = 7.8eV$, and Van der Waals attraction of $E^{VDW} = 0.056eV$ between sheets calculated from MSX force field [16].

5.4 Thermal Stabilities

The absolute cohesive energies for single ring, double ring, and bucky ball structures are compared in Figure 5.3. Here we see that single rings are stable for $n < 26$,

---

[1] $E(1/n) = A(1/n)^2 + B(1/n) + C$, by plugging in $E(C_{20})$, $E(C_{32})$ and $E(C_{60})$, we find A, B and C. $E(n \to \infty) = C$

buckyballs are stable for $n > 26$.

These estimates are for the minimum energy structure at zero temperature, without counting the zero point energy. For finite temperature, we must correct these results for entropy and for temperature differences in the enthalpy. To do this we used the MSX FF to calculate the vibration frequencies, which were then used to estimate the free energies(assuming harmonic vibration modes). These results are tabulated in Table 5.6–5.8 and plotted in Figure 5.14. Figure 5.14 shows the free energy versus temperature for monocyclic rings, bycyclic rings and fullerenes of several sizes n. For C_{20} the monocyclic ring structure has the lowest free-energy. From Figure 5.3 we see that fullerene has slightly lower energy than the bycyclic ring. But the zero point motion has made the bycyclic ring a winner over fullerene. For C_{32}, at zero temperature the stability order is ball, monocyclic ring and bycyclic ring. At around $800K$ the monocyclic ring becomes more stable than the fullerene due to the entropic effect. At around $1200K$ the bycyclic ring also become more stable than fullerene. For C_{60}, the fullerene has such a low energy compared to others that it is the minimum free-energy structure in the whole temperature range between $0K$ and $2500K$.

T(K)	Mono-ring	Bycyclic-ring	Fullerene
0.0	-6.2539375	-6.1646258	-6.1086969
100.0	-6.2683859	-6.1783714	-6.1209549
200.0	-6.2888352	-6.1977244	-6.1357029
300.0	-6.3138593	-6.2216472	-6.1519978
400.0	-6.3429928	-6.2497927	-6.1702125
500.0	-6.3758066	-6.2818159	-6.1906676
600.0	-6.4119261	-6.3173736	-6.2135095
700.0	-6.4510382	-6.3561442	-6.2387511
800.0	-6.4928786	-6.3978363	-6.2663287
900.0	-6.5372220	-6.4421945	-6.2961412
1000.0	-6.5838740	-6.4889950	-6.3280722
1100.0	-6.6326637	-6.5380438	-6.3620037
1200.0	-6.6834419	-6.5891722	-6.3978211
1300.0	-6.7360772	-6.6422329	-6.4354167
1400.0	-6.7904523	-6.6970952	-6.4746915
1500.0	-6.8464629	-6.7536464	-6.5155531
1600.0	-6.9040163	-6.8117836	-6.5579180
1700.0	-6.9630280	-6.8714165	-6.6017090
1800.0	-7.0234226	-6.9324642	-6.6468563
1900.0	-7.0851308	-6.9948532	-6.6932958
2000.0	-7.1480909	-7.0585182	-6.7409671
2100.0	-7.2122455	-7.1233979	-6.7898179
2200.0	-7.2775417	-7.1894374	-6.8397962
2300.0	-7.3439322	-7.2565870	-6.8908567
2400.0	-7.4113723	-7.3248014	-6.9429564
2500.0	-7.4798205	-7.3940354	-6.9960555

Table 5.6: Free energy G(T) for C_{20} isomers

T(K)	Mono-ring	Bycyclic-ring	Fullerene
0.0	-6.3535560	-6.3052448	-6.4554574
100.0	-6.3649048	-6.3158383	-6.4637079
200.0	-6.3824485	-6.3323535	-6.4737218
300.0	-6.4047864	-6.3537684	-6.4853715
400.0	-6.4313860	-6.3796164	-6.4991639
500.0	-6.4617748	-6.4094501	-6.5154011
600.0	-6.4955522	-6.4428609	-6.5341854
700.0	-6.5323874	-6.4794876	-6.5554934
800.0	-6.5720051	-6.5190185	-6.5792344
900.0	-6.6141716	-6.5611846	-6.6052875
1000.0	-6.6586860	-6.6057559	-6.6335231
1100.0	-6.7053733	-6.6525359	-6.6638123
1200.0	-6.7540806	-6.7013544	-6.6960333
1300.0	-6.8046733	-6.7520627	-6.7300722
1400.0	-6.8570323	-6.8045332	-6.7658250
1500.0	-6.9110504	-6.8586512	-6.8031970
1600.0	-6.9666331	-6.9143168	-6.8420999
1700.0	-7.0236956	-6.9714414	-6.8824559
1800.0	-7.0821600	-7.0299424	-6.9241909
1900.0	-7.1419571	-7.0897516	-6.9672406
2000.0	-7.2030231	-7.1508011	-7.0115416
2100.0	-7.2653000	-7.2130337	-7.0570426
2200.0	-7.3287336	-7.2763937	-7.1036891
2300.0	-7.3932775	-7.3408331	-7.1514340
2400.0	-7.4588854	-7.4063075	-7.2002345
2500.0	-7.5255150	-7.4727747	-7.2500498

Table 5.7: Free energy G(T) for C_{32} isomers

T(K)	Mono-ring	Bycyclic-ring	Fullerene
0.0	-6.3949740	-6.4003124	-6.8779842
100.0	-6.4041065	-6.4089095	-6.8828164
200.0	-6.4197344	-6.4240679	-6.8890703
300.0	-6.4403377	-6.4445657	-6.8973241
400.0	-6.4653127	-6.4698070	-6.9080843
500.0	-6.4941544	-6.4992626	-6.9215498
600.0	-6.5264436	-6.5324707	-6.9377411
700.0	-6.5618378	-6.5690343	-6.9565811
800.0	-6.6000543	-6.6086162	-6.9779449
900.0	-6.6408532	-6.6509296	-7.0016908
1000.0	-6.6840294	-6.6957310	-7.0276734
1100.0	-6.7294049	-6.7428136	-7.0557553
1200.0	-6.7768244	-6.7919986	-7.0858067
1300.0	-6.8261508	-6.8431320	-7.1177082
1400.0	-6.8772624	-6.8960802	-7.1513522
1500.0	-6.9300531	-6.9507239	-7.1866400
1600.0	-6.9844256	-7.0069600	-7.2234827
1700.0	-7.0402924	-7.0646954	-7.2617974
1800.0	-7.0975771	-7.1238490	-7.3015119
1900.0	-7.1562083	-7.1843433	-7.3425559
2000.0	-7.2161216	-7.2461141	-7.3848709
2000.0	-7.2161216	-7.2461141	-7.3848709
2100.0	-7.2772587	-7.3090990	-7.4283964
2200.0	-7.3395657	-7.3732434	-7.4730835
2300.0	-7.4029918	-7.4384959	-7.5188839
2400.0	-7.4674949	-7.5048105	-7.5657506
2500.0	-7.5330286	-7.5721430	-7.6136433

Table 5.8: Free energy G(T) for C_{60} isomers

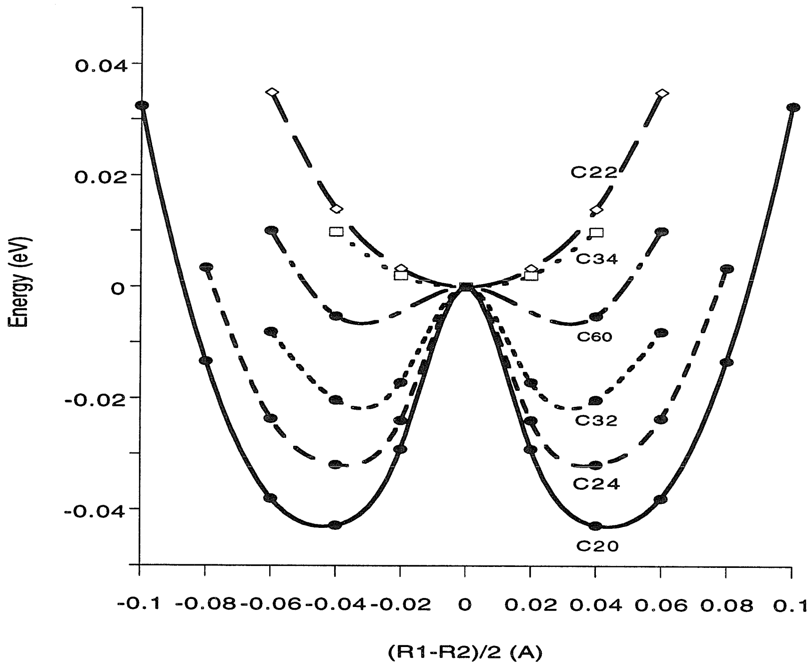


Figure 5.4: The total energy as function of bond alternations for carbon rings. All cases take the equal bond length geometry as the energy referenced point. C_{4m} dimerize, while C_{4m+2} don't

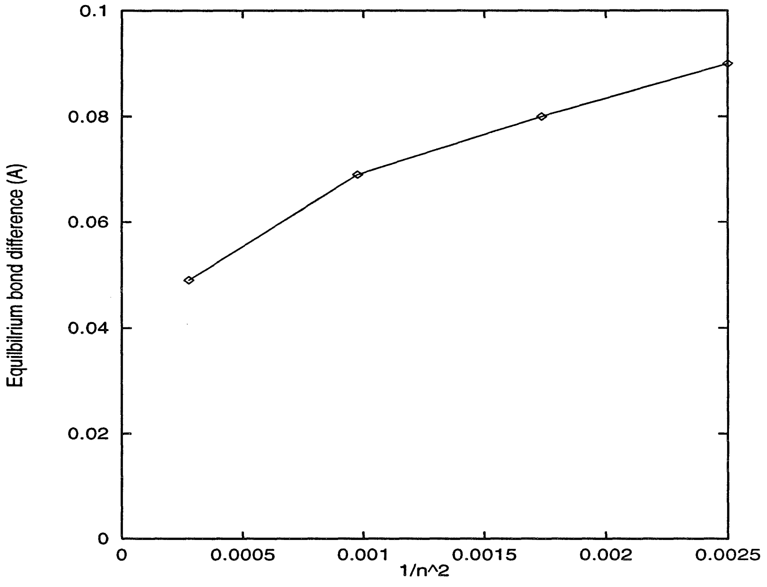


Figure 5.5: The bond difference decrease as the ring size increase. n is the size of a cluster.

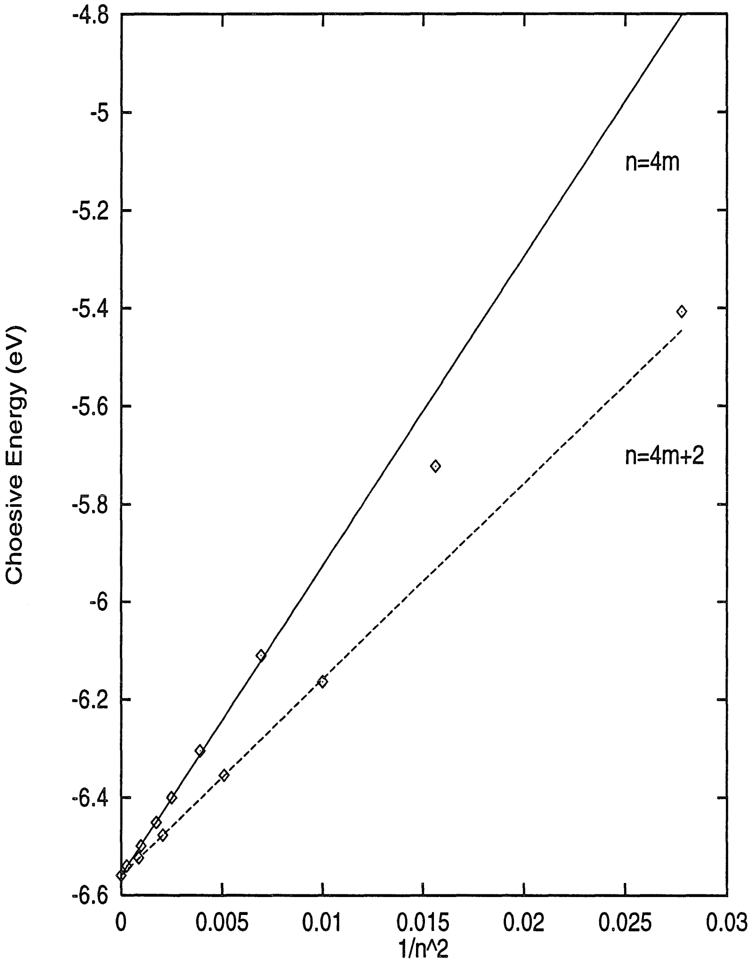


Figure 5.6: Cohesive energy of carbon rings C_n plotted against the $1/n^2$, which is proportional to the bending angle squared θ^2. C_{4m+2} rings, which have equal bond lengths, are more stable than the nearby C_{4m} rings which have alternate bond lengths.

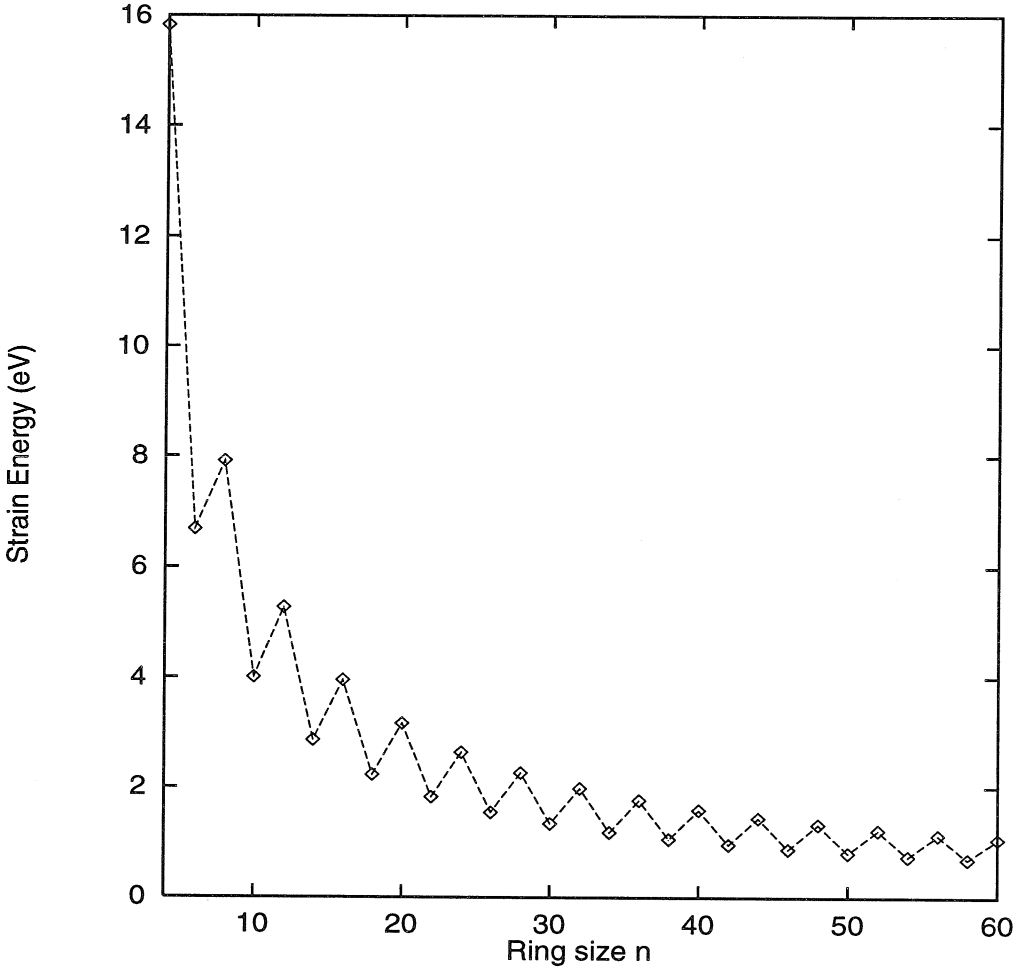


Figure 5.7: Cohesive energy of carbon rings as a function of ring size.

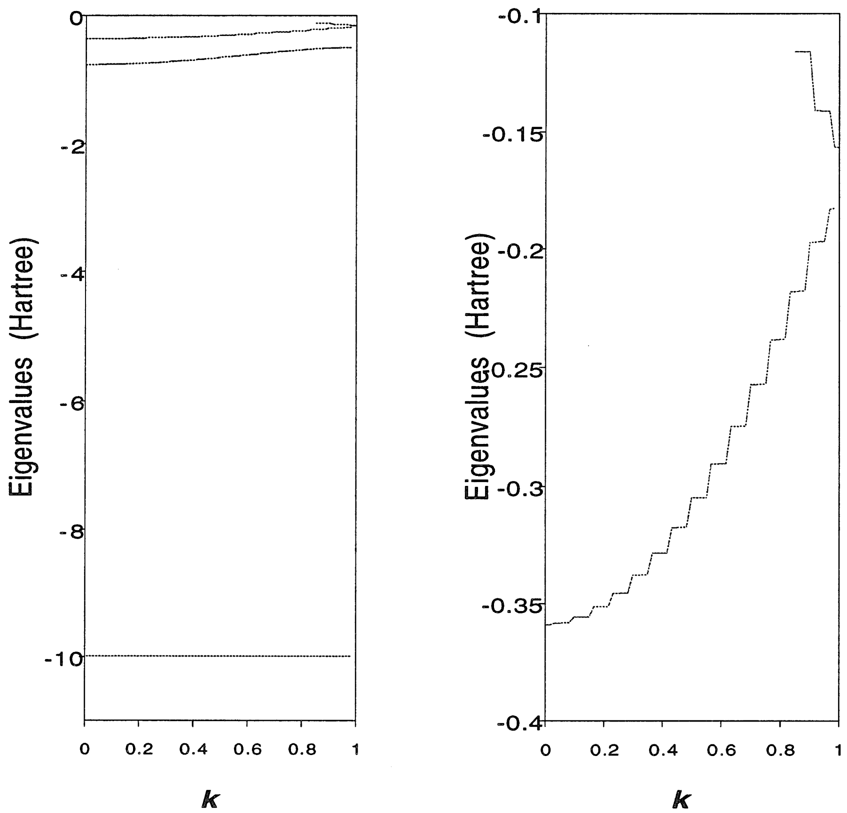


Figure 5.8: The DFT orbital energies of C_{60} ring plotted with equal spacing demonstrated a band-like structure. This is a direct illustration of the translational symmetry, which usually allows one to get the whole eigenvalues spectrum by simply adding a phase factor. See section (1.2.4).

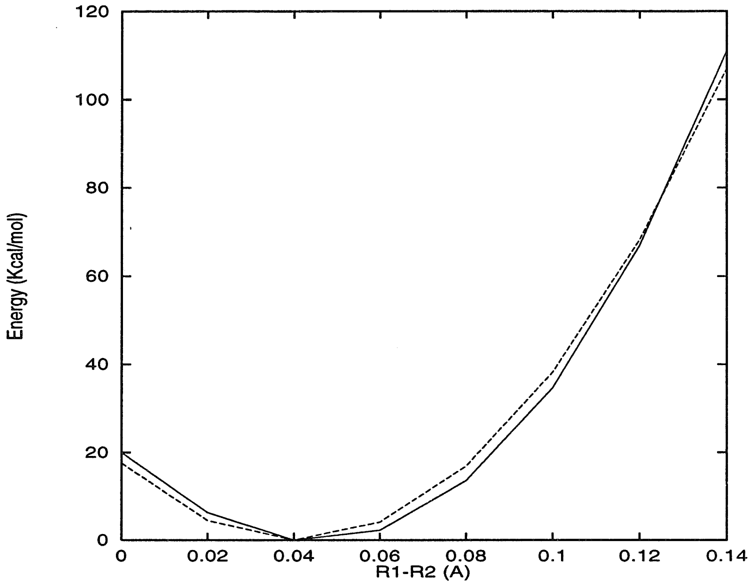


Figure 5.9: C_{20} ring energy as a function of bond length difference. The average bond length is $1.30\mathring{A}$. Solid line is from DFT and dash line from FF calculations.

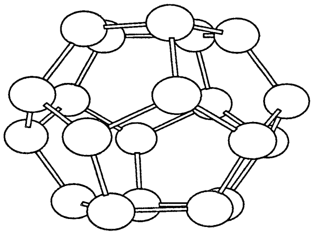
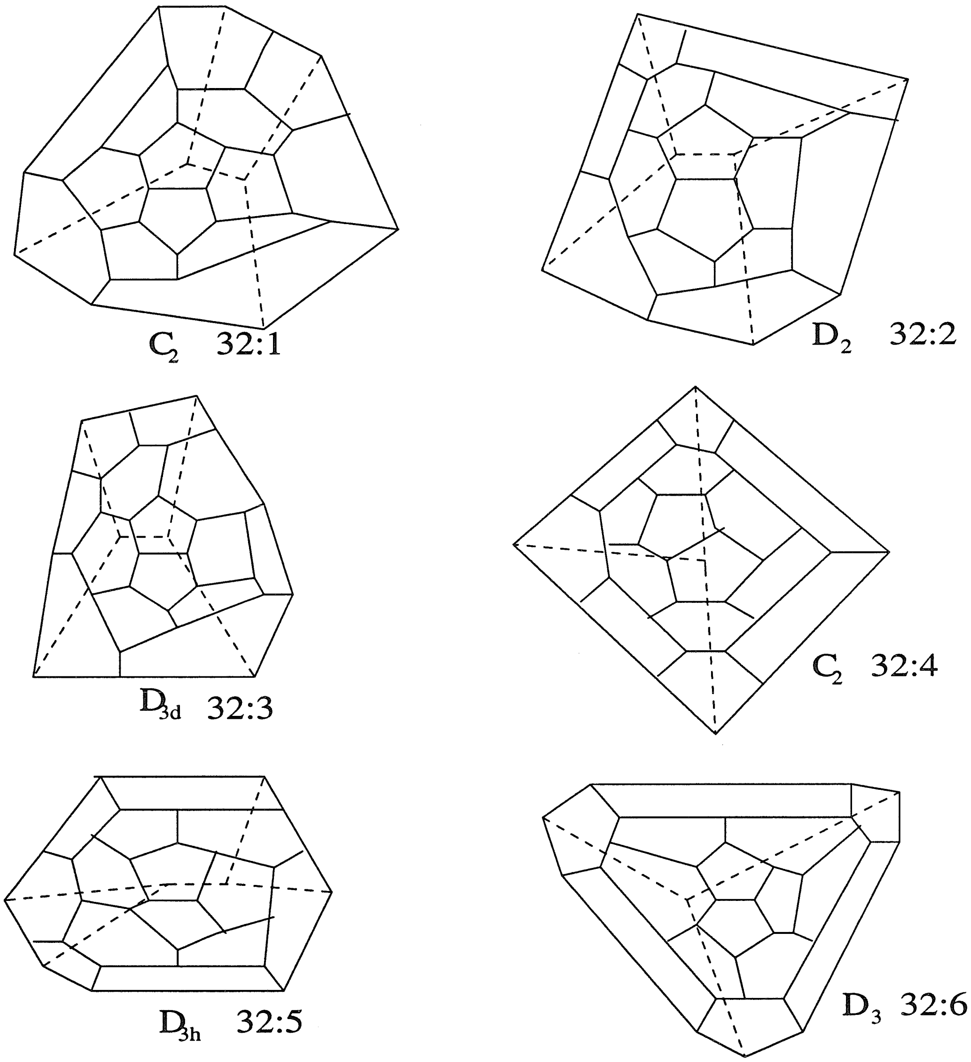


Figure 5.10: The structure of C_{20} with I_h symmetry



C_2 32:1

D_2 32:2

D_{3d} 32:3

C_2 32:4

D_{3h} 32:5

D_3 32:6

Figure 5.11: The Schlegel diagrams of C_{32} fullerenes

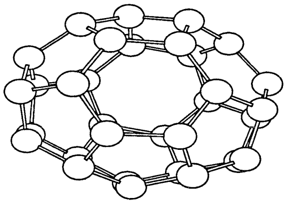


Figure 5.12: The structure of C_{32} with D_3 symmetry

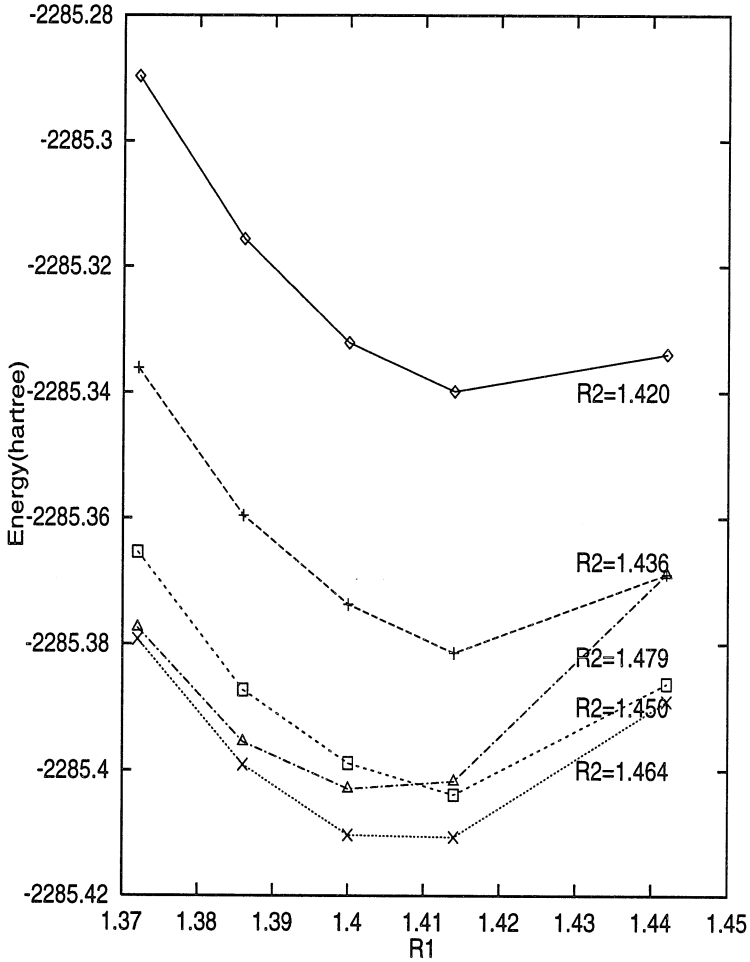


Figure 5.13: Total energy of C_{60} fullerene versus short bond and long bond, DFT calculation with 6-31G** and BLYP functional.

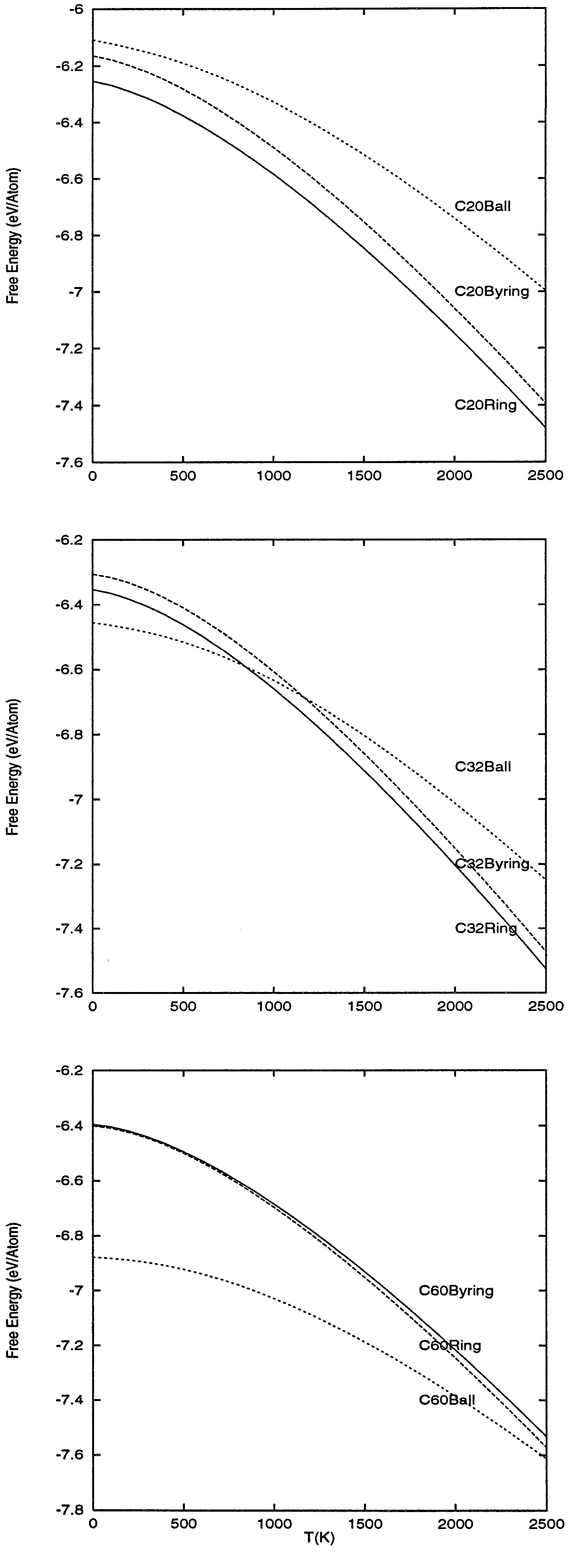


Figure 5.14: The Gibbs free energies for various isomers of various sizes, calculated with DFT within the harmonics approximation.

Bibliography

[1] B. G. Johnson, P. M. W. Gill, and J. A. Pople, *J. Chem. Phys.* **98** 5612, (1993).

[2] Murco N. Ringnalda, Jean-Marc Langlois, Burnham H. Greeley, Robert B Murphy, Thomas V. Russo, Christian Cortis, Richard P. Muller, Bryan Marten, Robert E. Donnelly, Jr., Daniel T. Mainz, Julie R. Wright, W. Thomas Pollard, Yixiang Cao, Youngdo Won, Gregory H. Miller, William A. Goddard III, and Richard A. Friesner, *PS-GVB v2.2, Schrödinger, Inc.*, (1995)

[3] G. V. Helden, N. G. Gotts and M. T. Bowers, *Nature* **363** 60, (1993); G. V. Helden, M-T. Hsu, N. Gotts, and M. T. Bowers, *J. Phys. Chem* **97** 8182-8192, (1993)

[4] G. König, G. Stollhoff *Phys. Rev. Lett.* **65** 1239 (1990)

[5] M. Feyereisen, M. Gutowski, and J. Simons, and J. Almlöf *J. Chem. Phys.* Vo. **96** 2926, (1992)

[6] K. Raghavachari *et al.*, Chem. Phys. Lett. **214**. 357 (1993).

[7] P. W. Fowler and D. E. Manolopoulos, *"An Atlas of Fullerenes"*, International series of monographs on chemistry 30 (1995)

[8] Y. Guo, *Ph.D. Thesis*, California Institute of Technology. (1992)

[9] *CRC Handbook*

[10] H. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curt, and R. E. Smalley, *Nature* **318**, 162-163 (1985);

[11] W. Krätschmer, K. Fostiropolos, and D. R. Huffman, *Chem Phys. Lett.* **347**,354-358 (1990)

[12] S. Iijima *Nature* **354**, 56 (1991)

Chapter 6 Formation Mechanism of fullerenes

6.1 Introduction

One of the most puzzling aspects of fullerenes (C_{60},C_{70}, etc.) is how such complicated symmetric molecules are formed from a gas of atomic carbons [1], namely, the atomistic or chemical mechanism. Are the atoms added one by one or as molecules (C2, C3)? Is there a critical nucleus beyond which formation proceeds at gas kinetic rates? What determines the balance between forming buckyballs, buckytubes, graphite, and soot? The answer to these questions is helpful in manipulating the systems to achieve particular products.

A difficulty in current experiments [3, 2, 4] is that the products can only be detected on times scales of μs, long after many of the important formation steps have been completed. Consequently, it is necessary to use computer simulation, quantum mechanics and molecular dynamics, to determine these initial states. The experiments serve to provide boundary conditions that severely limit the possibilities, making the use of first principles theory practical.

In the original laser evaporation experiment of Kroto and Smalley, [5] in the electric arc experiment of Krätchmer and Huffman, [6] and in the geological fullerene found in the Precambrian Russian rock, [7] which might have resulted from lightening, there is a common condition: pressure and temperature gradient. Thus, in fullerenes formation the intricate balance between space and time played an important role. *Space*, in terms of pressure determines the density of carbon atoms, thus the availability of the source for growth. While *time*, in terms of temperature, determines how long a metastable state would last in existence. Howard [8] produced fullerenes from benzene flame. The crucial difference here is the existence of H atoms. Because the H atoms are agents to terminate the dangling bonds, the energetics in the flame set-up is quite different. The fact that Howard's method produced lower yield than the arc method indicates that the existence of the dangling bonds is actually helpful in the process of finding the fullerene minimum.

The discovery of buckytube [9], [10] suggests that cylinders might be competitive in energy. But both experiment and HF calculations [11] pointed out otherwise. Upon annealing, the buckytubes curl into quasi-spherical particles composed of concentric graphitic shells, suggesting that these "Buckyonions" are more stable than other forms of carbon, including graphite, for systems of finites size. This is the experimental evidence that the tubes are metastable states. Yet they exist and are stable products.

All the above demonstrate that *kinetics pathways*, the temporal sequence for the formation of carbon clusters, more than the energy of the equilibrium states , need to be carefully examined to elucidate the mechanisms by which buckyball, buckytube or graphite form. In this chapter, we investigate the kinetic processes of fullerene formation.

6.1.1 Generalized Phase Diagram

The rate of a complicated reaction depends on how far the product and reactant are separated in phase space and how fast each individual step got accomplished. Again, when talking about kinetics we need to specify a coarse level in our description. Only so can the distance in phase space be meaningfully defined. At the finest grain, the reaction coordinates are bond angles, distances between atoms, the impact parameter, etc.. The output of the theory is the collision cross-sections. This is useful to study the very first moment, when C_2 and C_3 are formed. At the coarsest level, the reaction coordinate could be whatever order parameter $\phi(t)$ there happen to emerge, for example the percentage of sp^2 bonded carbons ρ_2.

If we use a to denote the quantity measuring the distance in phase space, then the macroscopic yield

$$R_{if} = \frac{\Delta a}{\tau_{tot}} = \frac{\Delta a}{\int d\tau}. \tag{6.1}$$

Using $R(\tau) = da/d\tau$ leads to

$$R_{if} = \frac{\Delta a}{\int da/R}. \tag{6.2}$$

Since transition states take up only a negligible fraction of time, it is very natural that integration along τ be divided into segments A, B, C, D, ..., etc., where A, B, C, etc., are distinguished molecular species. This division into reaction steps is the subject of the reaction mechanism. For a proposed reaction mechanism, the above equation, eq. (6.2), becomes summation over reactions:

$$R_{if} = \frac{\Delta a}{\frac{\Delta a_1}{R_1} + \frac{\Delta a_2}{R_2} + \frac{\Delta a_3}{R_3} + \ldots.} \tag{6.3}$$

The PES landscape also reflect the hierarchy structure. As illustrated in Figure 6.1. ϵ_2 is the energy difference between reactant A and product B, but on a coarser grain it is approximately the kinetic barrier between A and C. Because on the larger scale, B is a transition state species.

For the significant minima in PES we define a point (N, ρ_2), where N is the number of carbon atoms in the clusters and ρ_2 is the percentage of sp^2 bonded carbons. Thus we can lay down the cluster species on a two-dimensional diagram, which we call generalized phase diagram. Figure 6.2 illustrates such a diagram. Line segments along constant N represent unimolecular reactions. Whereas paths cross constant N lines represent fusing, growth, fragmentation reactions. In the unimolecular reaction sp^1 bonded carbons are converted into sp^2 bonded ones, a process we call stitch-up. This happens continuously.[1] Another way to represent the formation process is through evolution tree. See Figure 6.3. This tree is analogous to that in biology. A

---

[1] Because it is very hard to imagine that the ring would wind themselves onto the surface of a cage first without having adjacent carbons stitching up.

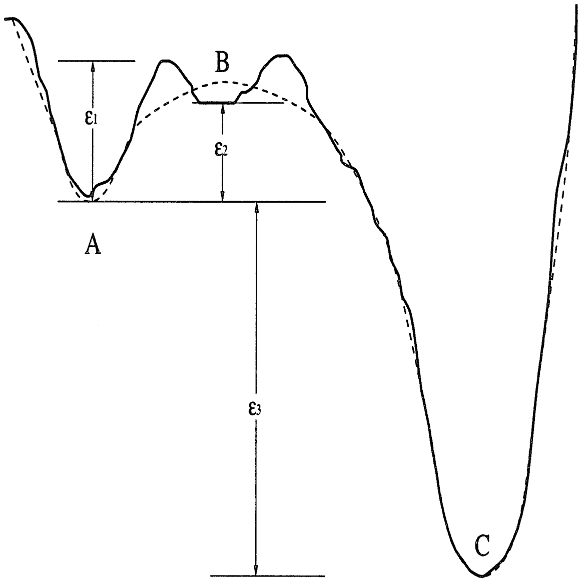


Figure 6.1: The thermodynamics energy difference on one scale, e.g., ϵ_2, could be the kinetic energy barrier of a coarser scale.

single line constitutes one species existing in time. When two individual in a branch combine, they form a different species and thus branch out to become a new species. When two individuals belonging to two species combine they become a third species. The evolution of fullerene from atomic species has an arrow in time. Formation and fragmentation apparently follow different paths.

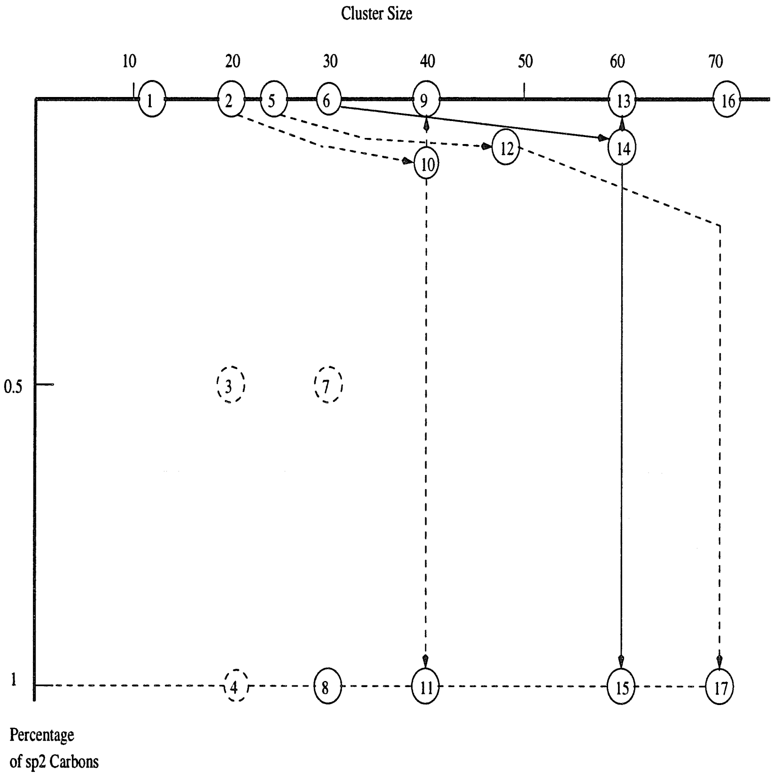


Figure 6.2: The generalized phase-diagram $N-\rho$ for carbon cluster, where N is the cluster size and ρ is the percentage of sp^2 bonded carbons.
(1)C_{12} monocyclic ring; (2)C_{20} monocyclic ring; (3)C_{20} bowl; (4)C_{20} fullerene; (5)C_{24} monocyclic ring; (6)C_{30} monocyclic ring; (7)C_{30} bowl; (8)C_{30} fullerene; (9)C_{40} monocyclic ring; (10)C_{40} bycyclic ring; (11)C_{40} fullerene; (12)C_{48} bycyclic ring; (13)C_{60} monocyclic ring; (14)C_{60} bycyclic ring; (15)C_{60} fullerene; (16)C_{70} monocyclic ring; (17)C_{70} fullerene, etc..

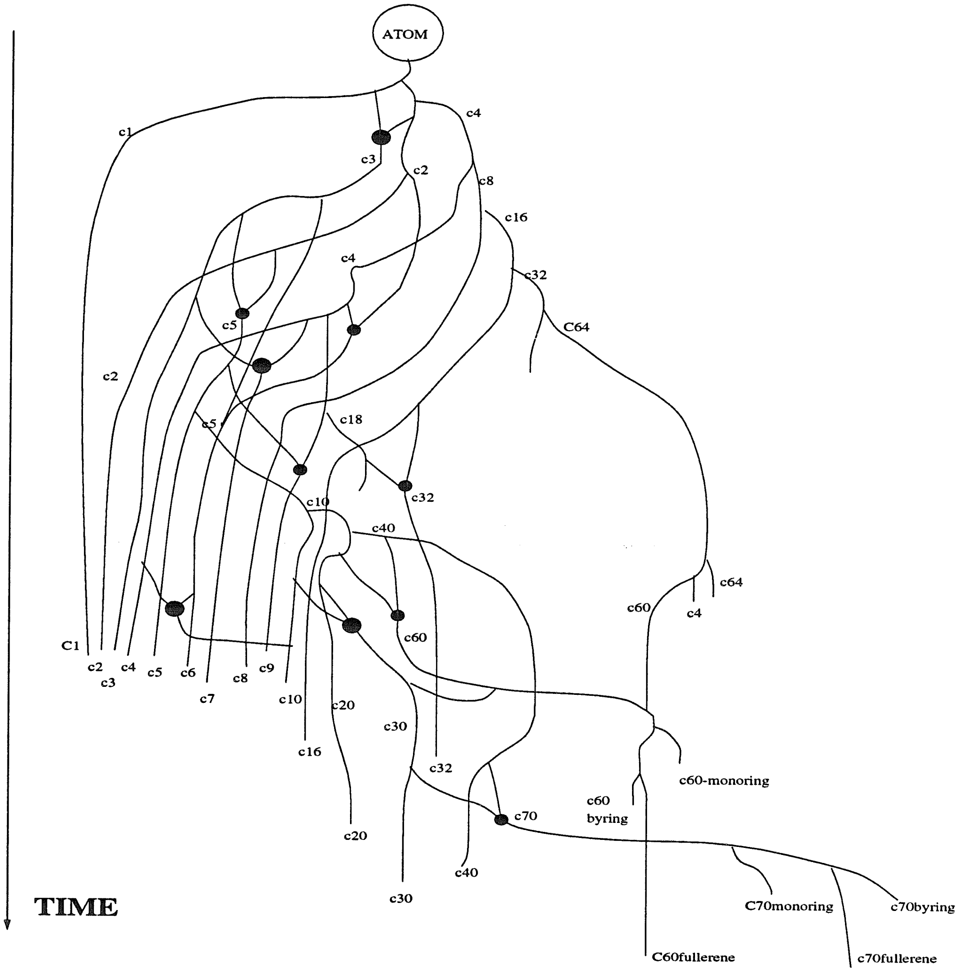


Figure 6.3: The evolution picture of carbon clusters. Massive simplification is for illustration only. Many, many more process are possible.

6.1.2 Experimental Observations

Synthesis Conditions

C_{60} fullerene yields depend only slightly on the buffer gas pressure between 1×10^4 and 5×10^4 Pa (100-400 Torr), but decrease rapidly at lower or higher buffer gas pressures.

Isotopic Scramble Experiment

Ebbesen *et al.* [1] used ^{13}C and ^{12}C respectively for the two graphite gods in the arc experiment. The products showed the mixing ratio of a complete scramble. No magic numbers, no two peaks at the respective locations for C_{60} and C_{70}, but only one single peak with the Poisson distribution. So the carbon clusters must be formed from atomic mixing. Namely, graphite has been broken down to atoms and clusters are assembled from atomic carbons.

Ion Distribution

The earliest work on carbon cluster that show the hint of C_{60} might be the ion abundance distribution of Exxon group in 1984. [12] We noticed that there is a gap between C_{26} to C_{36}.

Ion Chromatography Experiment

Bowers and Jarrold [13] have developed the ion chromatography technique to differentiate the structure of isomers. The method measures the mobility of ions as they drift across a chamber filled with buffer gas. Since the mobility is related to the cross-section, one can measure the cross-section of carbon cluster with this method. The ion chromatography can determine the cross-section to such an accuracy that one can identify whether the ion is monocyclic ring or bycyclic ring or fullerene. Ion chromatography thus serves as a tool to distinguish isomers of various compactness. See Figure 6.5

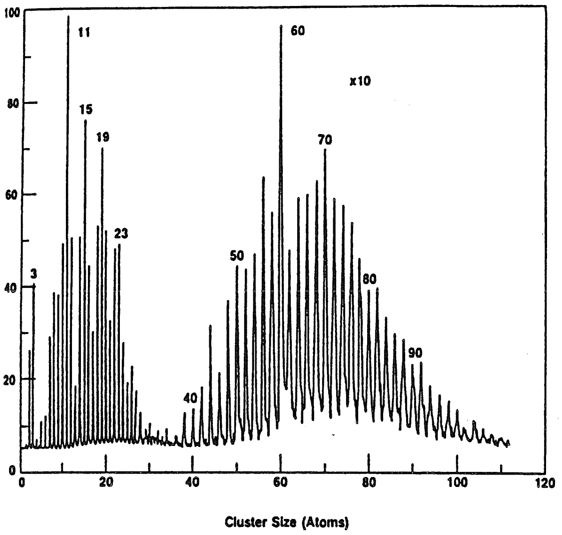


Figure 6.4: The measured carbon cluster distribution by laser vaporization of a solid graphite rod within the throat of a high pressure pulsed nozzle. The experiment showed two regions, odd number small clusters and even number large clusters, with a gap in between around $n = 30$

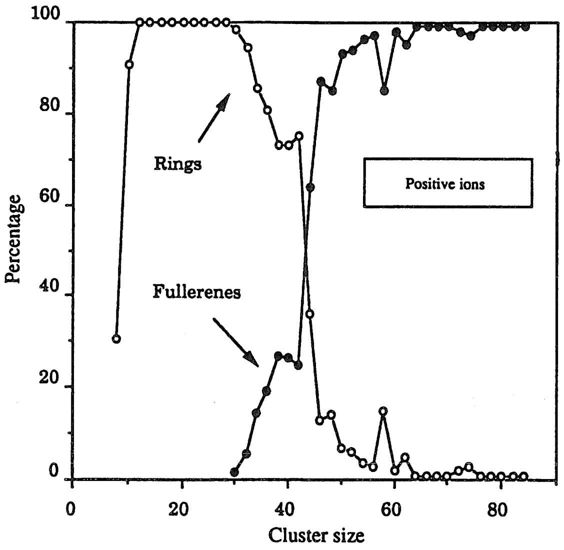


Figure 6.5: Bowers analysis of the isomer populations from ion drifting experiment for cations. We can see that below n=30, the carbon clusters are ring-like, above n=30, they prefer fullerene structures

Both Bowers group and Jarrold's group demonstrated that isomerization reaction can convert C_{60} bycyclic ring into C_{60} fullerene. Also, the observed behavior of C_{40}^{+}, C_{46}^{+}, etc., are very much like C_{60}^{+}, see Figure 6.6 and Figure 6.7.

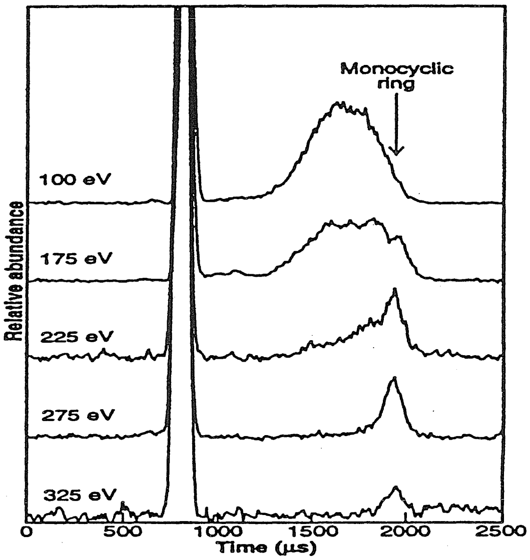


Figure 6.6: Conversion of C_{60} bycyclic rings to C_{60} monocyclic rings and fullerenes when heated.

Thus it is hard to imagine that the C_{60} fullerenes would follow a completely different pathway as that of C_{40} fullerenes. So the distinguished abundance of C_{60} is not due to the kinetic pathways, but due to the stability of the molecule itself, which in turn is determined by $\partial E_{tot}(N)/\partial N$, a structural property. Thus the formation mechanism has a reduced task of explaining the pathway or pathways to reach that thermodynamically very favorable state.

6.2 Methodology

Synthesis of fullerenes involves high temperature and high pressures. In forming C_{60} fullerenes, huge energy is released. Indeed, as we showed in the figure 5.3, the iso-

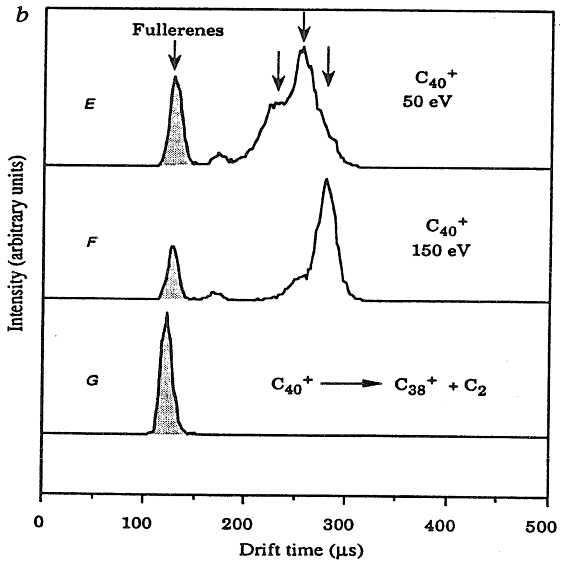


Figure 6.7: Conversion of C_{40} bycyclic rings to C_{40} monocyclic rings and fullerenes when heated.

merization from bycyclic rings to fullerenes releases $30eV$ of energy, not to mention the total atomization energy of $423eV$ from the very beginning. To handle this large energy scale, we used two levels of modeling. For instance, in studying the isomerization reactions from bycyclic rings to fullerenes, we will use a fine grained one for the crucial initial steps, and a coarse grained one for the subsequent stitching up steps.

6.2.1 Fine Grain

Ab initio calculation will give the accurate energies. However, for a molecule of 60 atoms, complete ab initio calculation with correlation included is still too expensive, except for some very symmetric configurations. Notice that the ab initio scheme is not needed in all the calculations. Contineous deformations could be accurately evaluated relative to a reference point using MD with a well calibrated force field. Hence we have devised a combined DFT and MD method that direct the ab initio power to where they are really needed, e.g., where bonds are formed or broken. Take as an example the fusing of two C_{30} monocyclic rings into a C_{60} bycyclic ring. The energies were computed with the combined DFT and MD as follows, see Figure 6.8 for an illustration.

(1) The reaction from two C_{30} ring (**I**) to C_{60} bycyclic ring molecules (**II**) is achieved via an intermediate **III**. For each of **I, II and III**, the system is divided into two parts, part **A**, which involves bond change, and part **B**, which involves just continium deformation.

(2) The energy of **I** is calculated with DFT, see Figure 5.3.

(3) The energy difference between **I** and **III** can be calculated by MSX FF, since it's just strain energies.

(4) (a) For energy difference between **III** and **II** we do it by part. For part **A** we do DFT calculations on the reaction from two C_6H_2 molecules separated afar to the 4-member ring molecule $C_{12}H_2$. (b) For part **B** we calculate the corresponding change in going from **III** to **II** using MSX FF.

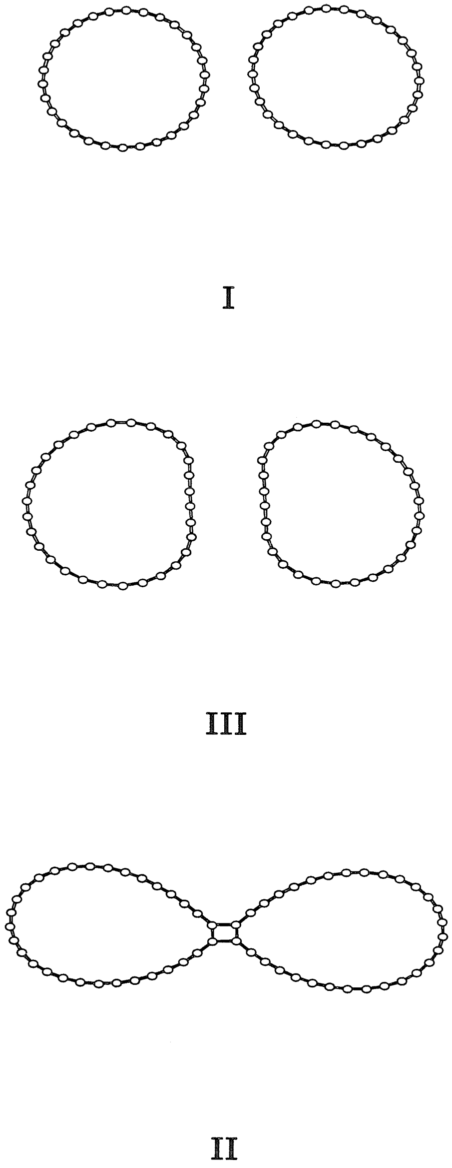


Figure 6.8: We break down the calculations into two parts. One part can be calculated accurately by molecular dynamics with MSX force field, the other part is small and can be calculated accurately DFT

Then we combine **A** and **B** to get the energy difference between **III** and **II**.

(5) Thus the energy of C_{60} bycyclic ring (**II**) can be calculated by II–III–I.

Once one configuration is accurately calculated, it can be used as a reference point for other configurations which are slightly different in bonding. This way, we can devise a contiguous path connecting two structures, each point on this path differs only slightly in bonding from the structure ahead of it. Thus the part that needs ab initio treatment is local and small. A cluster model will be sufficient for it. This scheme substantially cuts the computation cost. One criterion of accuracy of this approach is to check the path independence. The error accumulated in summing up all the differences along the path should be small.

All the DFT calculation is done with the BLYP functional [14] on a 6-31G* basis set using PS-GVB program. [15]

6.2.2 Coarse Grain

The Extended MSX FF for the general sp^1 and sp^2 Carbon Clusters

The isomerization from C_{60} monocyclic rings to fullerenes releases $30eV$ of energy. To handle reactions on this energy scale we shall use a model, whose parameters are calculated with DFT. The key components are the additive energy terms for the dangling bond and the energy cost for bending a triple bond to form a 1,2-benzene. Our FF are defined as follows:

$$E_{tot}(n_2) = E_{bond} + E_{radical} + E_{strain} = n_2(\epsilon_1 - \epsilon_2) + d_1 n_R + d_2 n_{\sigma\pi} + E^{str}(n_2). \quad (6.4)$$

We have chosen $E_0 = 60\epsilon_1$ as zero point. Here, n_2 is the number of sp^2 bonded carbons, $n_2(\epsilon_1 - \epsilon_2)$ gives the energy gained by converting sp^1 bonded carbon into sp^2 bonded carbon, with $\epsilon_1 = -6.56eV$ and $\epsilon_2 = -7.71eV$. d_1 is the energy of a radical relative to the bonded state, n_R number of of such radicals (dangling bond); d_2 is the energy of an atom participating bended planer π-bond relative to the $\sigma-$bonded state

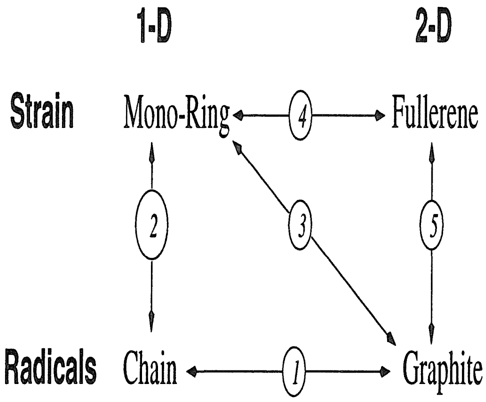


Figure 6.9: The balance between three major forces: σ-bonding, strain energies and dangling bonds

and $n_{\sigma\pi}$ is the number of such atoms. We used the Benson-like scheme to evaluate d_1 and d_2, [16] and found $d_1 = 2.32eV$ and $d_2 = 1.64eV$. The n_R and $n_{\sigma\pi}$ as a function of n_2 depends on the growth sequence, see Table 6.2 for an example. $E^s(n_2)$ is the strain energy evaluated at the FF minimum energy structure of a given bonding scheme.

6.3 Energetics for Initial Growth

The physics under carbon clustering is three type of forces,

- *kill the dangling bonds.*
- *minimize strain energy.*
- *converte $\pi-$bond to $\sigma-$bond.*

For example, carbon dimer C_2 is of course linear having two dangling bonds. Linear C_3 has two dangling bonds. Triangle C_3 has no dangling bonds but huge strains. Linear C_4 has two dangling bonds. Rectangular has no dangling bonds but huge strains. So the formation of clusters is characterized by the relative importance of these three driving forces, see figure 6.9.

As shown in this figure, there are several interconversions among chains, rings and fullerenes.

(1) Branching off leads to one extra dangling bond d_1, the gain in getting one more σ-bond $\epsilon_1 - \epsilon_2$ is not enough to offset it.

(2) Bending to form a ring, on the other hand, would lead to large strain energy E^{strain}. But the larger the ring, the smaller the strain energy. So this linear growth mode continues until the strain energy balance out the dangling bond $2d_1$. Then the closed rings become the thermodynamically most stable species. The prevalence of ring structures are results of thermodynamics. The pathways by which these rings are formed are the subject of detailed kinetics. See our discussion on C_{12} ring formation in last section.

(3) The energy balance between the dangling bonds and σ-bonding, $\epsilon_1 - \epsilon_2$, is reached at cluster size n_0, which can be estimated as follows:

$$\begin{aligned} \pi[R(n_c)]^2 A(\epsilon_2 - \epsilon_1) - 2\pi R(n_c) B\Delta \\ > -E^{strain}_{ring}(n_c) \end{aligned} \tag{6.5}$$

where, A is the area density of carbon atoms in a graphitic sheet. B is linear density of carbon atoms in the peripheral of the sheet. $R(n)$ is related to the sheet size simply by $n = \pi R^2 A$. From section 6.2.2 $\Delta = 2.32eV$, $\epsilon_2 - \epsilon_1 = 7.71 - 6.56 = 1.15eV$. We get $n_0 \sim 26$. Therefore all the way to C_{26}, it is energetically unfavorable to have graphitic sheet.

(4) Both monocyclic ring and fullerenes have zero number of dangling bonds. They interconvert because of the balance between strain energy and σ-bonds. For a given cluster size, monocyclic rings would have smaller strain energy, for they are more extended. But fullerenes enjoy more σ-bond, each gaining $\epsilon_2 - \epsilon_1$. From Figure 5.3 we see that the two forces balanced out at $n = 26$.

(5) Any closed fullerenes has 12 pentagons. (See section 5.3.1) The strain these pentagons incur is constant and decreases on the per carbon basis. So no matter how large a single graphitic sheet gets, the dangling bonds at its boundaries always lost it to the closed fullerene. However, large single graphitic sheet can curl up and form tube whose strain approaches zero at very large size rather than a constant. In reality, the bonding *among* sheets along the direction perpendicular to the sheet will become important. Consequently, for very large clusters, the multilayered structure is most favorable.

(6) We didn't mention the conversion between linear chain and the fullerenes. They are far separated in configuration space and thus don't have direct conversion.

6.4 The Energetics for Isomerizations

Ion chromatography experiments implied the way C_{60}^{BF} is formed. Theoretical analysis of the last section also shows that the history of C_{60} must be that of growth and isomerization. First, grow in forms of extended rings and followed by isomerization into compact fullerenes. This also fits the angular momentum picture. In this section we focus on the energetics of the isomerization process. In particular, how a bycyclic ring, which is the most popular isomer from a growth history, isomerizes into a fullerene. We shall use the methodology described in section 6.2.

6.4.1 Reactive Part of the Molecules

We performed the DFT calculations on the model clusters representative of the reactive part of C_n clusters.

(A) $C_6H_2 + C_6H_2 \rightarrow C_{12}H_4$

The intermediate geometries are shown in Figures 6.14 through Figures 6.10

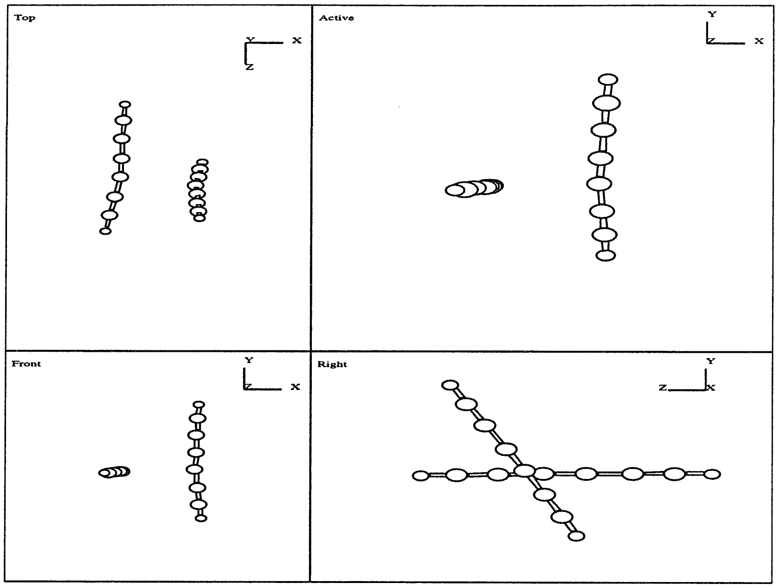


Figure 6.10: The geometry of the intermediates for C_6H_2+C_6H_2 → $C_{12}H_4$. Constraint R_{23} at 3.00$\AA$, starting with ∼ 75^o.

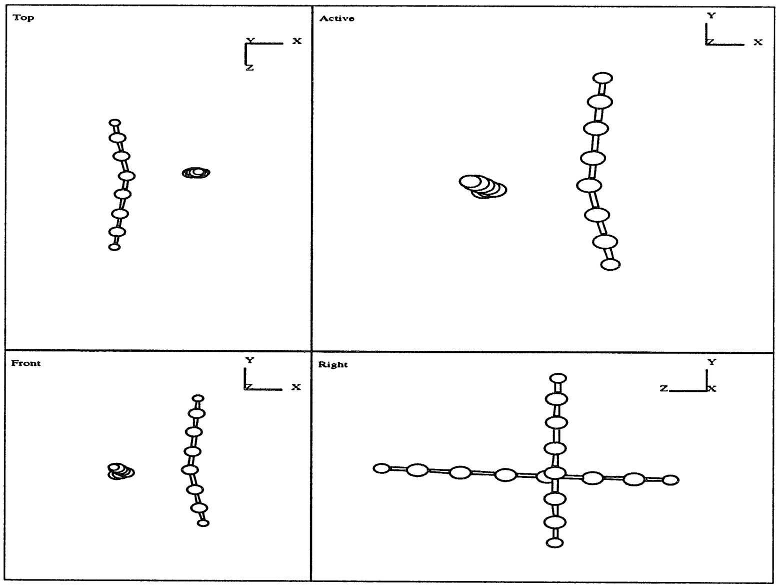


Figure 6.11: The geometry of the intermediates for C_6H_2+C_6H_2 → $C_{12}H_4$. Constraint R_{23} at 2.50$\AA$, starting with ∼ 90^o.

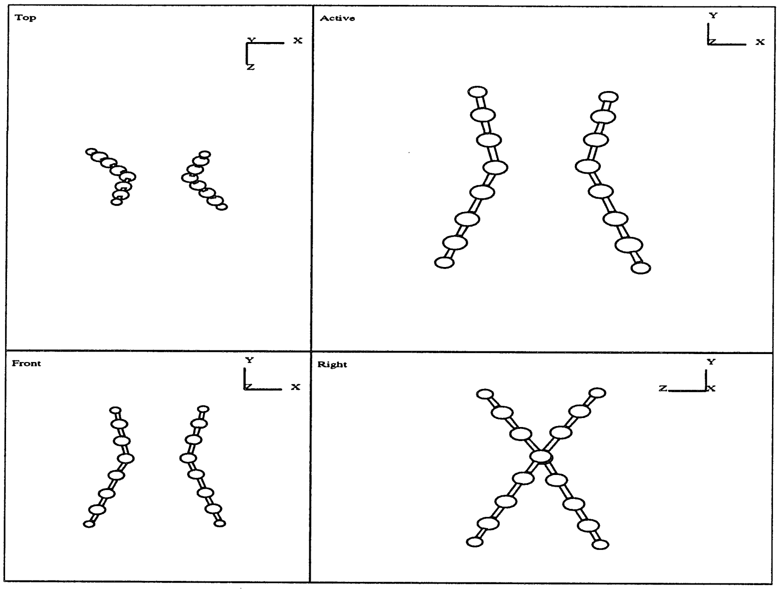


Figure 6.12: The geometry of the intermediates for C_6H_2+C_6H_2 → $C_{12}H_4$. Constraint R_{23} at 2.35$\AA$ and start with 20^o twisting angle.

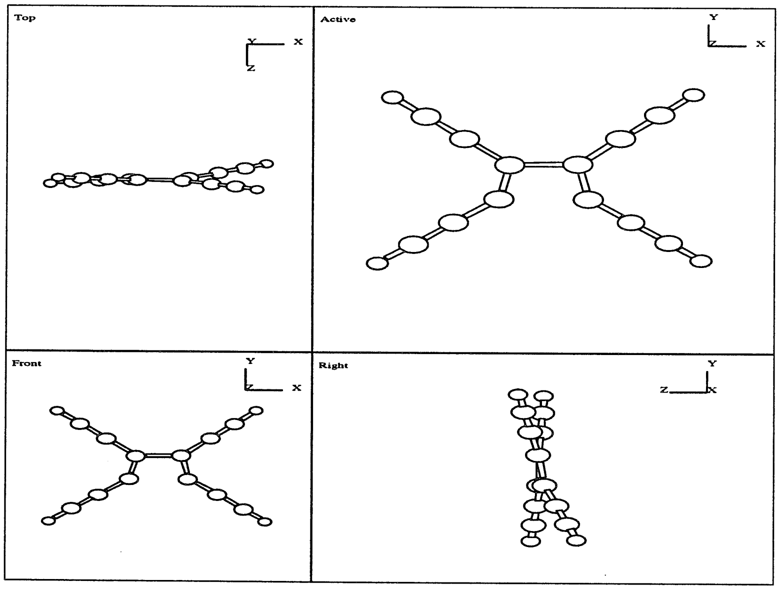


Figure 6.13: The geometry of the intermediates for C_6H_2+C_6H_2 → $C_{12}H_4$. Constraint R_{23} at 2.00$\AA$ and start with 20^o twisting angle.

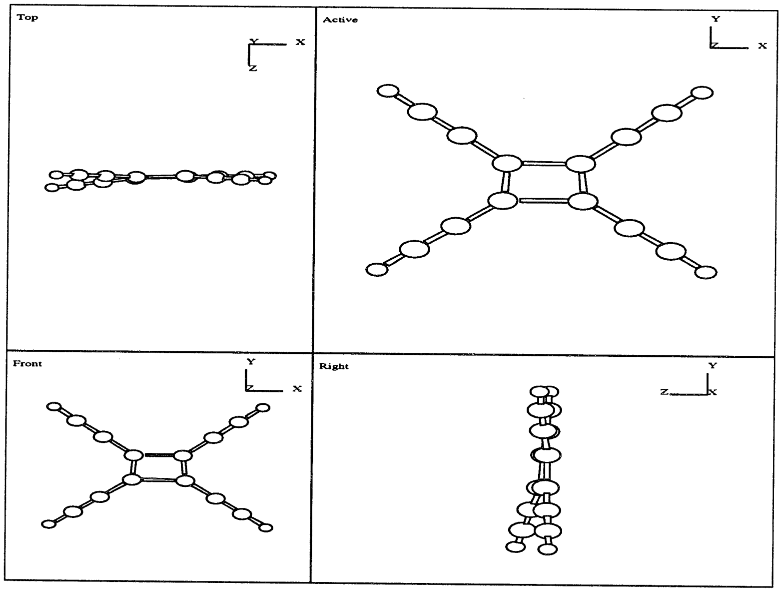


Figure 6.14: The geometry of the intermediates for C_6H_2+$C_6H_2 \rightarrow C_{12}H_4$. Constraint R_{23} at 1.70Å.

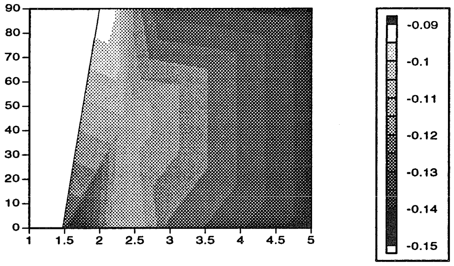


Figure 6.15: The energy contour for C_6H_2+C_6H_2 $\rightarrow$ $C_{12}H_4$. The x-axis is R_{23} and y-axis is the dihedral angle ϕ.

The energies are shown in Figure 6.16. We found that the barrier in going from two separate C_6H_2 chains to $C_{12}H_4$ is $19.6 Kcal/mol$ and the reaction is exothermic by $3 Kcal/mol$. Figure 6.15 gives the contour plot in terms of two reaction coordinates (R_{23}, ϕ). R_{23} is the distance between atom 2 on one chain and atom 3 on the other. ϕ is the twisting angle between the two C_6H_2 planes. We noticed that the transition state geometry has $\phi = 45^o$. The final state has planer geometry.

(B) Bergman Cyclization

The Bergman cyclization [17] is made easier by bending the carbons ahead of the cyclization. Here we examine two configurations. One without bending, one with bending. See figure 6.17. We see that prebending lowered the energy difference from $1.9 eV$ to $0.5 eV$, making the cyclization much easier to occur.

(C) Opening the Four-membered Ring

The four-membered ring in $C_{12}H_4$ has large strain energy. By breaking one σ-bond between atom 1 and 2, (see Figure 6.18) and forming two π bonds, thus opening the four-membered ring, one gets the vase-shaped molecule (referred to as C_{12}-vase) and

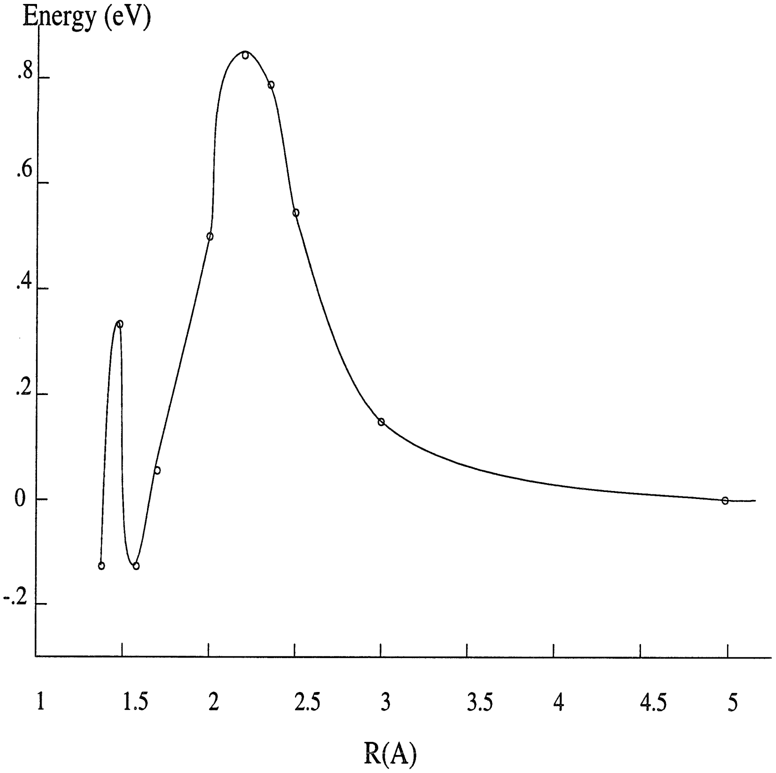


Figure 6.16: The reaction path for C_6H_2+C_6H_2 → $C_{12}H_4$. Note that there is an energy barrier of $0.85ev$. This is used for evaluating the energy of fusing two monocyclic rings

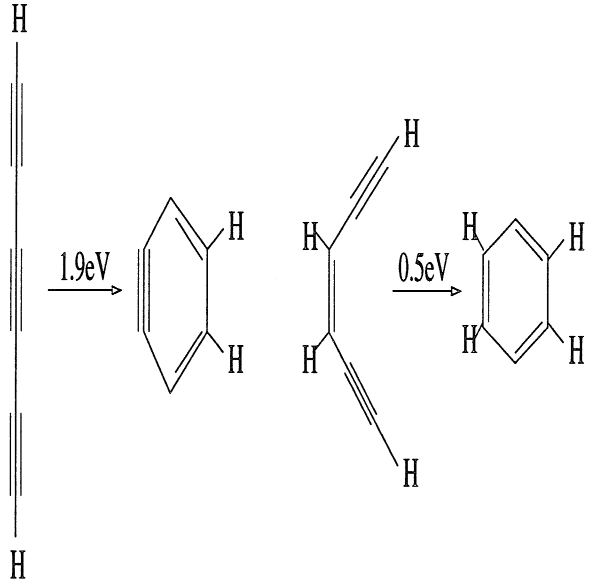


Figure 6.17: The pre-bending of the carbon chain facilitates the Bergman cyclization.

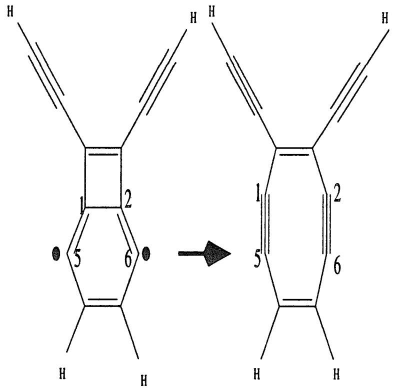


Figure 6.18: It is very easy to cleave the σ bond between atom 1 and atom 2 and form two π bonds between 1 and 5, and 2 and 6

releases the strain. This isomerization is downhill by $\sim 1eV$.

(D) Forming the Second Ring

Figure 6.19 shows the process of forming the second six-membered ring. This process is uphill by about $1.2eV$.

6.4.2 Annealing the Polycyclic Rings

After getting the energetic for the reactive part of the C_n isomers, we add in the chain part calculated with MD to get the total energy of the whole molecules, using the reference point and path integration method prescribed in Section 6.2. Since the central part of molecules, the part that participate bond changes, are the same for all the $C_n(n = 34, 40, ..., 60)$ in the conversion reaction from bycyclic ring to monocyclic rings, their energetics are readily calculated.

One C_{16} monocyclic ring and one C_{18} monocyclic ring collide into a C_{34} bycyclic ring, which subsequently isomerized into a C_{34} monocyclic ring. We use

$$C_{16} + C_{18} \rightarrow C_{34}^{BR} \rightarrow C_{34}^{MR}$$

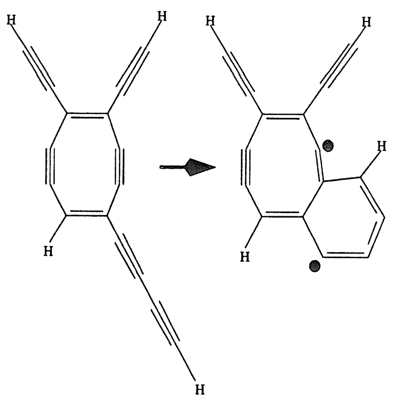


Figure 6.19: One arm swings around and attaches one of the triple bonds, forming the second ring, a six-membered ring. The process is uphill by $\sim 1.2eV$

to denote the sequence.

Similarly, we have the isomerization reaction for C_{40} and C_{60}:

$$C_{20} + C_{20} \rightarrow C_{40}^{BR} \rightarrow C_{40}^{MR}$$

$$C_{30} + C_{30} \rightarrow C_{60}^{BR} \rightarrow C_{60}^{MR}$$

Their energetics are compared in Figure 6.20

From Figure 6.20 we can see that there are two types of bycyclic rings, one is with a four-membered ring and the other is with an eight-membered ring. The 8-member-ring isomers are formed after a Bergman cyclization from the four-membered ring, which leads to the diradical **7**. **7** relaxed into **8**. The four-member bycyclic rings **4** have their kinetic origins from colliding two monocyclic rings **0**. Upon annealing they are converted into monocyclic ring **6**. This is exactly what the Jarrold group found in their ion chromatography experiment. [3, 2] See Figure 6.21 for their relative abundance analysis for annealing carbon clusters with size n =34, 40.

Energetics of several reaction products in the initial steps of forming sp^2 network are list in Table 6.1

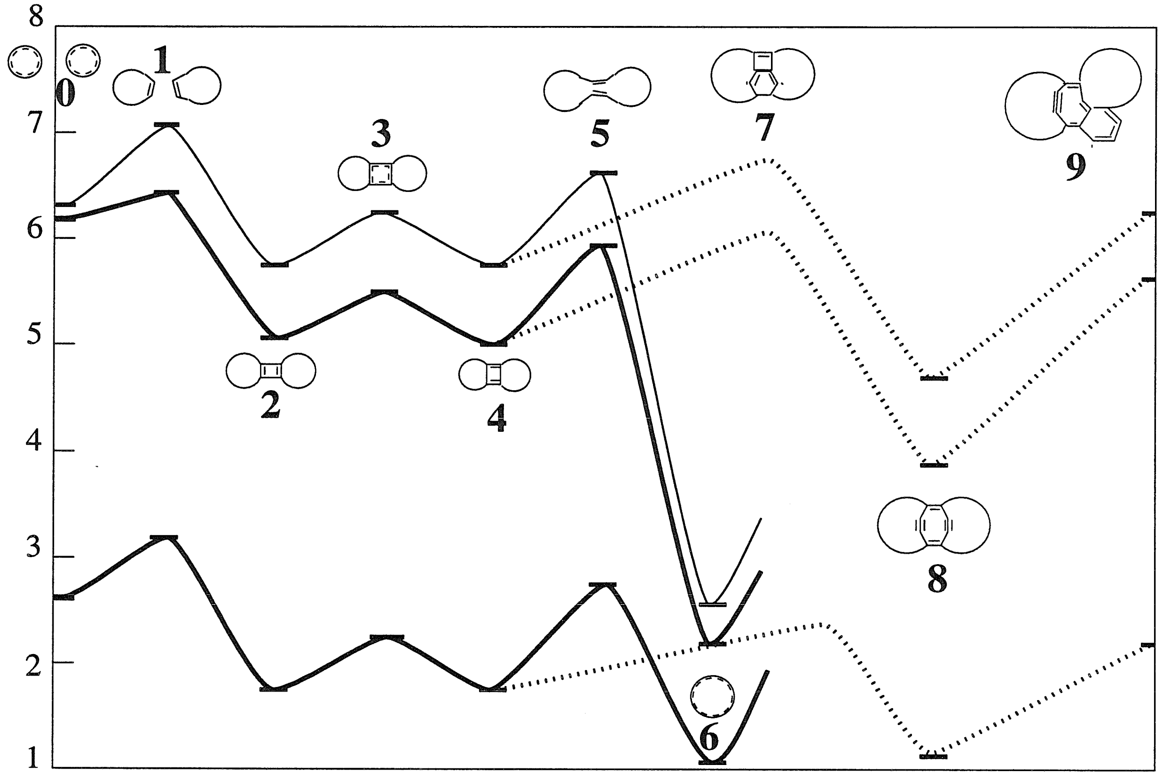
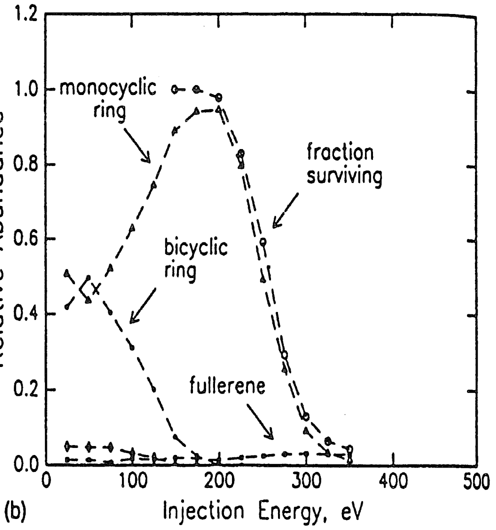


Figure 6.20: The energetics of isomerization reactions that convert bycyclic ring to monocyclic rings for several clusters C_n with, n=30, 40, 60. Their reference points have been set to $30\epsilon_1$, $40\epsilon_1$ and $60\epsilon_1$, respectively, with ϵ_1 being the atomization energy of an infinite carbon chain. The energy is in unit of eV.



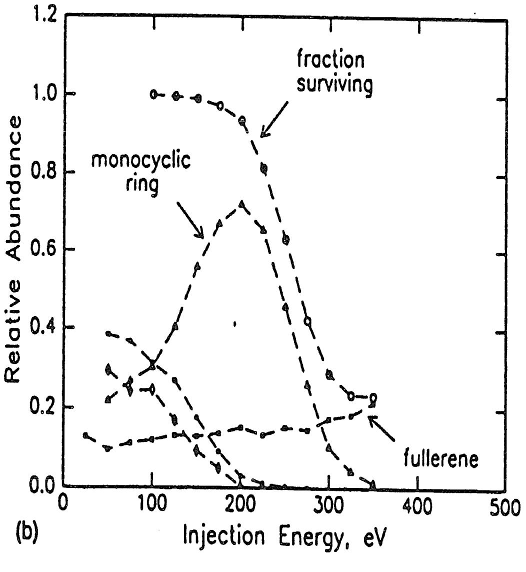


Figure 6.21: The relative abundance of the isomers and fragments as a function of injection energy observed in ion drifting experiment of Jarrold's group. The upper one is for C_{34}^{+} and the lower one for C_{40}^{+}.

Cluster Name	E_{tot}	$[E_i - E_{i-1}]\|_{path}$
0	2.674	0
1	3.25	0.576
2	1.717	-1.533
3	2.171	0.454
4	1.717	-0.454
5	2.72	1.003
6	1.055	(from 5) -1.665
7	2.325	(from 4) 0.608
8	1.086	-1.239
10	2.16	(from 7) -0.165
12	-0.01	-2.17
13	3.07	(from 10) 0.91
14	2.95	-0.12

Table 6.1: The energetics of the initial steps in C_{60} isomerization. The reference energy E_0 is $-60\epsilon_1$. ΔE_i is calculated within each path. The unit is in eV.

6.5 Kinetics for Fullerene Formations

We discuss the reaction kinetics in two parts. Section 6.5.1 and 6.5.2 study the sequence of reaction steps without a measure of time.[2] While Section 6.5.3 studies the reaction rates, which uses a measure of time. It relates the reaction rates to laboratory time scales.

6.5.1 Clustering Kinetics

At the beginning atomic carbons combine themselves to form dimers and trimer, C_2, C_3. These would then collide into linear chains for cluster size $n < 10$. Beyond $n > 10$ the carbon chains collide into some intermediates which subsequently relax into monocyclic rings. This is because thermodynamically monocyclic rings is the most stable species at that size range, as demonstrated by both the experiment [4] and theory. [18] As shown previously, the lowest energy state for carbon clusters of less than ten is linear chain. For C_n of $n > 10$ the lowest energy states are monocyclic rings. How are these monocyclic rings formed? Are they formed by fusing smaller chains, e.g., in the middle, and then going through isomerization, or are they formed by closing the two ends of a long chain? Because the chain already has too high an energy in having two radicals at the ends, it is very unfavorable to bend it, adding strain energy. Also, the event of two ends coming to within bonding range is rare. So it is clear that small rings are formed by fusing two chains in the middle. Take C_{12} monocyclic ring for example. Center of mass consideration tells us that colliding two equal masses results in the maximum energy transfer of kinetic energy into internal energy. The most efficient channel would be colliding two C_6 chain in the middle to form an four-membered ring intermediate. Then via Bergman cyclization the two ends would close successively, forming a pinched monocyclic ring, which would subsequently relax. See Figure 6.22.

Both experiments (Figure 6.5) and theory (Figure 5.3) indicate that at around $n > 30$ the ring structures give way to fullerene structures. This is because replacing

---

[2] A space of discrete points without metric!

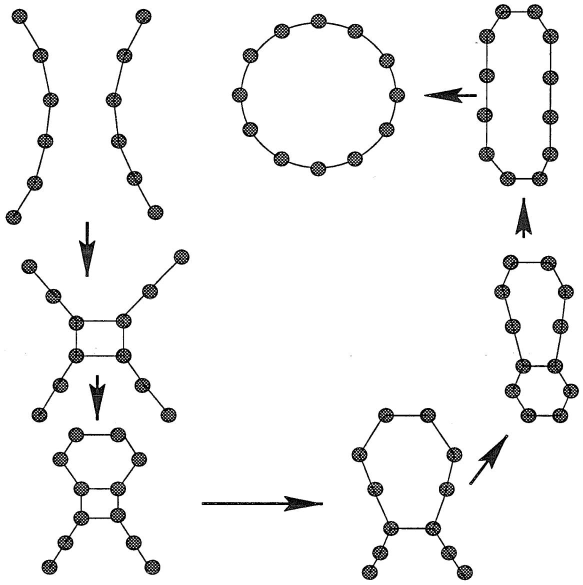


Figure 6.22: The process by which a C_{12} monocyclic ring is formed. Colliding two C_6 chains in the middle to form a four-membered ring, which serves as a seed for Bergman cyclization to close the arms on the two sides.

more π-bond by σ-bond can now over compensate for the strain of folding the 2-D net as the latter decreases. One process of C_{60} formation, as suggested by Jarrolds experiments, [4, 2] is to combine two C_{30} rings to form a C_{60} fullerene. This is an important channel, as is supported by other experimental and computation observations, most noticeably, the cluster distributions of Rohlfing [12] and McElvany's laser desorption experiment. [19] In Rohlfing's cluster size distribution experiment(figure 6.4) we noticed that C_{30} are very reactive toward making higher clusters. McElvany's experiments provided the direct evidence that following the laser desorption of $C_{30}(CO)_{10}$, a monocyclic C_{30} ring coalesces with a C_{30}^{+} ion to form a fullerene ion C_{60}^{+}.

6.5.2 Isomerization Kinetics – The Spiral Growth Model

The Sequence of Isomerization Reactions

We have calculated the energetics of the spiral growth model for the isomerization from a bycyclic ring to a fullerene proposed by Jarrold's group [2] In this section we shall describe the sequence and the energetics. As a mnemonic for referring to the various structures, we will simply denote the ring sizes of a structure. Thus the simple C_{60} ring is denoted as $\{60\}$, while the double ring system, **4**, is $\{30+4+30\}$. This notation does not uniquely describe a structure, but it is for the species we will consider. We will take the reference energy to be $E_o = 60\epsilon_1$, where $\epsilon_1 = -6.56eV$.

(i)**4**=$\{30+4+30\}$ → **7**=$\{30+4+6+30\}$. This is a Bergman diyne cyclization which forms a 6-membered 1,4 benzene-like ring from two triple bonds. This leads to two isolated radical sites (sp^2-like orbitals in the plane, that cannot form a bond), and we find that this increases the energy by about $0.6eV$.

(ii)**7**=$\{30+4+6+30\}$ → **8**=$\{30+8+30\}$. This process kills two dangling bonds by breaking one σ bond and forming two π bonds. This process is downhill by about 1.2eV.

(iii)**8**=$\{30+8+30\}$ → **10**=$\{30+8+6+22\}$. This involves breaking an in-plane π bond and forming a σ bond. In the process there is bending of one triple bond to from a 1,2-benzene-like ring which includes a new radical site. This process is uphill

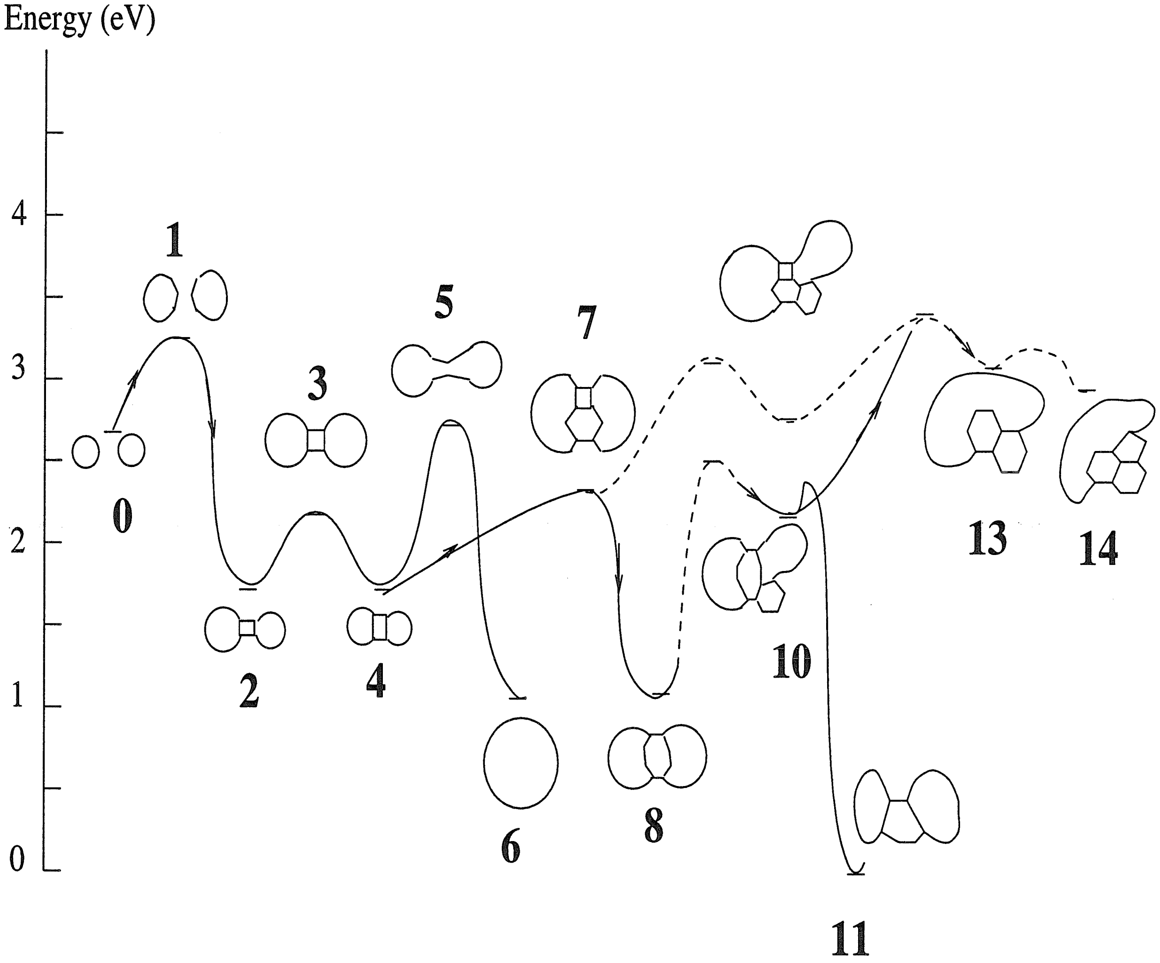


Figure 6.23: The initial steps in Jarrold model. The solid lines are real calculations while the dashed lines are from estimates.

n_2	n_R	$n_{\sigma\pi}$	n_2	n_R	$n_{\sigma\pi}$
10	2	4	40	8	0
12	2	4	42	8	0
14	2	4	44	8	0
17	3	4	45	7	0
20	4	4	46	6	0
22	4	4	48	6	0
24	4	4	50	6	0
26	4	4	52	6	0
28	6	2	54	6	0
30	6	2	55	5	0
32	6	2	56	4	0
34	8	0	57	3	0
36	8	0	58	2	0
38	8	0	59	1	0

Table 6.2: The number of dangling bonds for various isomers in the growth sequence.

by 1.1eV.

(iv)**10**=$\{30+8+6+22\} \rightarrow$ **13**=$\{6+6+55\}$. This involves twisting open the original 4-membered ring. Then it is followed by relaxing the 50 carbon chain to reduce the strain energy. This $\{6+6+55\}$ contains two dangling bonds. This process is uphill by about 0.9eV. See Figure 6.23 for the energetics of these initial steps.

(v)Spiral growth around the **13**=$\{6+6+55\}$. As a first step **13**=$\{6+6+55\} \rightarrow$ **14**=$\{6+6+5+53\}$. This uses one of the sp^2 orbitals of the 1,2-benzene-like ring to attack a triple bond of the carbon chain and form a new 5-membered ring. This process is downhill by 0.13eV.

(vi) Continue the spiral growth to form C_{60} fullerene. The energies of this process is calculated with the extended MSX FF (see section 6.2.2). The sequence of stitching up with the corresponding dangling bond counts are shown in Table 6.2. In Figure 6.24 we show the energy changes as the C_{60} intermediates grow into the closed fullerene cage. The energy are monotonically downhill. The overall gain of energy from **13**=$\{6+6+55\}$ to C_{60} is about $30eV$. No barriers are expected to impede these steps. Figure 6.25 illustrates some of the intermediates between **14** and the fullerene.

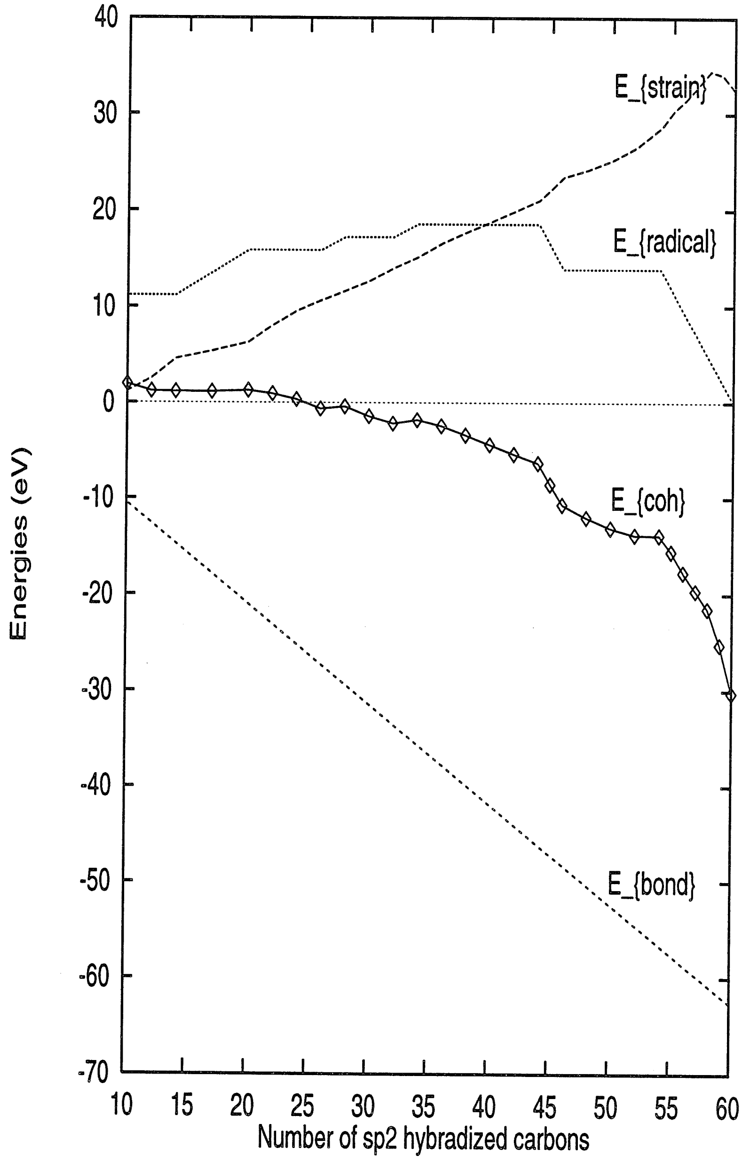


Figure 6.24: The energetics of stitching up, the spiral growth, calculated from the extended MSX force field (coarse grained field field).

We show in Figure 6.24 the various contribution to the total energy. We can see that the driving force for the growth is the gain in forming sp^2 σ bond. The opposing forces are the energy increase caused by the radicals created along the way and the increasing strain energies.

Annealing Into Position

Naturally, one may wonder how the carbons in the ring know their pentagon-hexagon sequence. Namely, when should they decide its time to form a pentagon, not a hexagon? The rule is: *making pentagons but avoiding adjacent pentagons.* Building pentagons are the shortest way of converting sp^1 carbons to sp^2 carbons. So pentagon formations are attempted most frequently. On the other hand pentagon incurs more strain energy than hexagons, two adjacent pentagons are energetically too unfavorable. In any case the sequence above is meant for the idealized case. In reality, the growth is a statistical event and the sequences need not be unique. The energy released upon stitching up heats up the molecule. Through mechanisms like Stone-Wales transformation, [20] or sp^3 intermediates [21] rearrangement, the molecules will anneal themselves into various isomer structures, with isolated pentagon structure being the lowest energy (strain effect). Figure 6.26 illustrates Stone-Wales transformation to move around pentagons in two-dimensional net.

6.5.3 Time Scale of Reactions and the Rates

The energetics of reaction intermediates along the fullerene formation pathways discussed so far is only one factor in determining the yield, the other factor is the energy barrier along the way and the distance in configuration space. The barrier determines the probability of getting over to the product for a given reactant energy. While the detailed separation in configuration space determines the attempting rate in a chemical reaction. *Both the attempting frequencies and energy barrier contribute to the reaction rate of a chemical reaction.* For a complicated reaction like C_{60} fullerene formation, the rate is determined by the time scale of many reaction steps along the

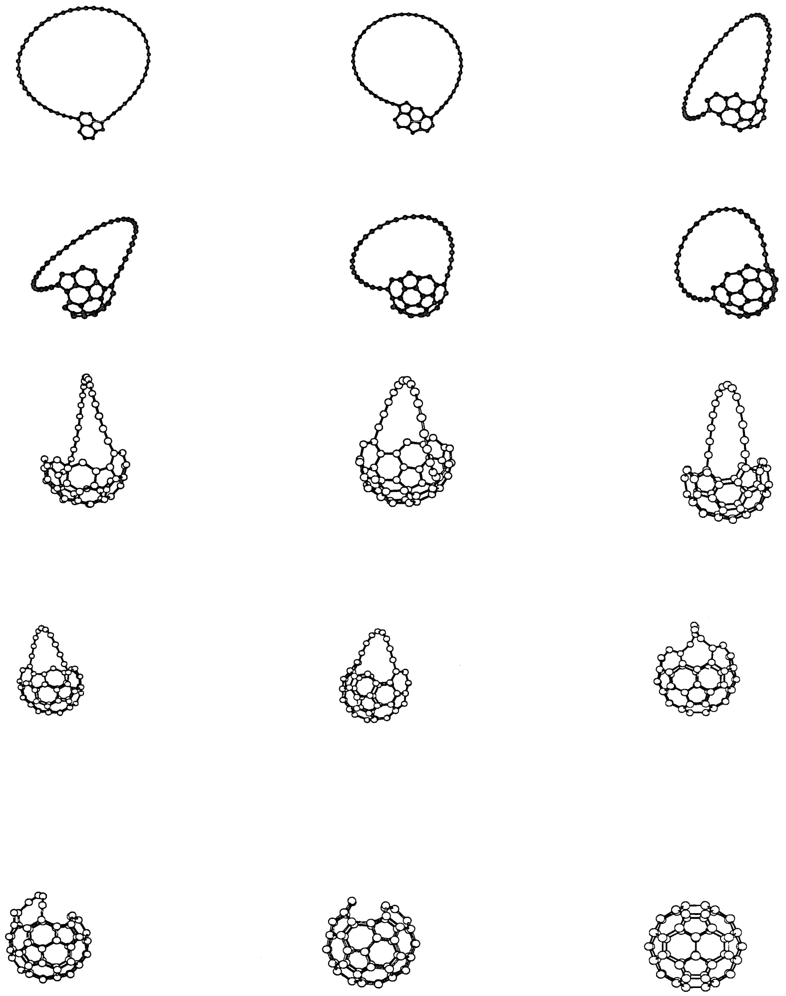


Figure 6.25: Some intermediates from the species **13**={6+6+55} having 10 sp^2 bonded carbons to the C_{60} fullerene.

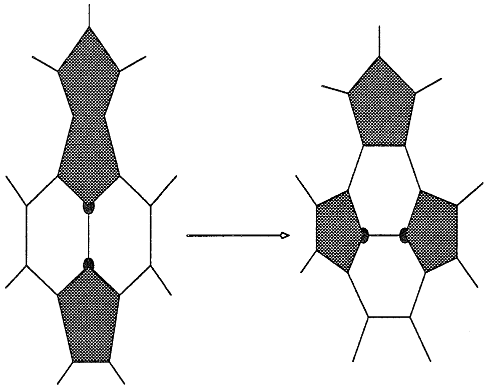


Figure 6.26: The Stone-Wales transformation by which two carbon atoms move their positions and lead to the change in bonding, resulting in the switched position for the pentagons.

reaction pathway, especially the slowest step. At each step one needs to consider both the barrier height and the attempting frequencies. In this section, we give, for the first, estimation of the time scales of the various fundamental reactions in the fullerene formation.

A. Theories

For simple $A + B \rightarrow C$ type of reaction, there has been various theories regarding how to calculate the reaction rate.

Transition State Theory

Transition State Theory(TST) assumes,

> (1) There is an equilibrium between reactant and transition state; and (2)all systems at the transition state go into the product without coming back, and they linger at the transition state for h/k_BT amount of time.

(Classical physics has continuous phase-space volume $d\mathbf{p}d\mathbf{q}$. It needs the introduction of h for counting the state. At temperature T, each degree of freedom has an energy $k_BT/2$. So the average time the system spends in that state is h/k_BT.)

TST is an equilibrium theory, neglecting all the dynamical feature of a reaction.

Kramers' Brownian Motions

Kramers [22] in 1940 derived from Focker-Planck type of equation the two limiting cases of Brownian motion of a particle under an external field, the high viscous limit and the low viscous limit. The behavior of a Brownian particle in external field were used to model chemical reactions. In 1983, Carmeli and Nitzan extended the model to all viscosities, bridging the two limits of Kramers. [23] Kramers' model improves over the transition state theory in incorporating the curvature of the potential energy surface, and thus is a dynamical theory. However, it's too simplistic for our complicated reactions, which involve many degrees of freedom.

NTV Molecular Dynamics

To tackle the complicated reaction like winding spiral chains into fullerenes, we used the full molecular dynamics at finite temperature T. Our MSX force field generates the PES along which systems evolve in real time. Running NTV MD over a period of 2ns, we can determine all the configurations that occurs within two nanoseconds. Also we can determine what must be absent in that time scale. For our 1 femtosecond time step, 2ns consists of 2,000,000 steps. Counting the frequencies of a configuration gives accurate measure of attempting frequency.

B. How Does the Spiral Arm Stitch-up

What is the time scale for the spiral chain to stick to the nucleation center? We did a MD simulation using our MSX force field. At T=3000K, the six-membered ring comes to within $3\mathring{A}$ closure in every $3ps$, while the five-membered ring $0.5ps$. So the attacking frequency is about $0.3 \times 10^{12}/s$ and $2 \times 10^{12}/s$, for forming a six-membered ring and a five-membered ring, respectively, see figure 6.27. So the spiral growth is fairly rapid.

C. The Rate of the Isomerization Reactions

To get the rate for the whole isomerization reaction, we need to to know the rate of each isomerization steps, see eq. 6.2. If $R(\tau)$ for each reaction are approximately the

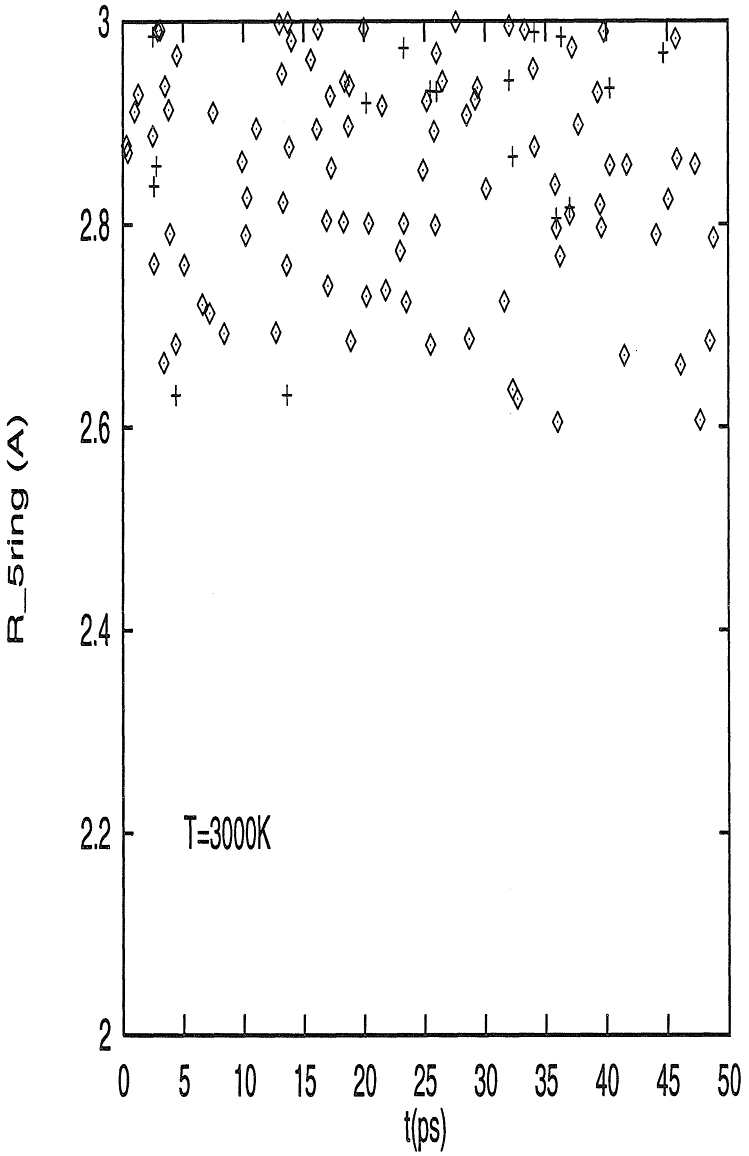


Figure 6.27: The distance between reacting carbons, reactions that lead to five-member ring or six-member ring, at T=3000K for C_{60} when there are 40 sp^2 bonded carbons. The symbol cross is for forming a six-membered ring and diamond is for forming a five-membered ring.

same, then we can use a one-dimensional diffusion model to calculated how long it takes to drift from $a(sp^2) = 12$ to $a = 60$. This will give us the rate of the overall isomerization reaction.

6.6 Comparison with other models and hints for experiments

6.6.1 Comparison with Other Models

For making fullerenes from atomic carbons, there could only be three major roads: the one-dimensional road proceeds by colliding rings; the two-dimensional road proceeds by growing graphitic sheets and the three-dimensional road proceeds by growing small fullerenes.

A. The Pentagon Road

Smalley in 1992 proposed a mechanism which he called the Pentagon road. [24]

> *When the carbon cluster size passes about 30 atoms, graphitic sheets begin to form. As smaller species are ingested by the graphitic sheet, it attempts to follow the low energy path minimizing the number of dangling bonds by including pentagonal at convenient places among the hexagonal ones. The incorporation of pentagons allows the sheet to minimize energy by curling up, reducing the number of dangling bonds.* [25]

It is designed to account for the high yield of C_{60} fullerene synthesis. The explanation is actually only one possibility and it contradicts the small yield of C_{70}. The most serious trouble of this model, though, is its obvious contradiction with ion chromatography experiments which showed unambiguously that there are no graphitic sheet as intermediates. Though one can argue that it's the flux rather the concentration that is important for clusters serving as intermediates, the argument is a week defensive one.

B. The Fullerene Road

Heath in 1991 proposed the Fullerene road: [26]

> *Graphitic sheets growing from small carbon radicals do not stay open. They close to form fullerenes as soon as possible (possibly starting with* C_{28}*). These small fullerenes then grow by addition of small clusters, e.g.* C_2*, until* C_{60} *is reached. The addition of small clusters to the fullerene cage proceeds in a manner which removes adjacent pentagons. The* C_{60} *atom cluster size is rarely skipped over.*

It assumes the coexistence of cage and small C_n, like C_2 and C_3, etc. McElvany *et al.* [19] have shown that following laser desorption of $C_{30}(CO)_{10}$, a monocyclic C_{30} ring coalesces with a C_{30}^{+} ion to form a fullerene ion C_{60}^{+} without any evidence for growth by C_2 addition. More serious trouble is its difficulty in explaining the readiness of producing endohedral metallofullerene molecules. For if the fullerenes are closed at C_{32} and grow from there, they cannot allow the metal atom to enter the cage without huge energy.

C: The Ring Road

The ring road, suggested by Jarrold, [2] is consistent with all the observations.

Isomerization make the cluster most compact, effectively cuts down the collision frequency, thus reduces the chance for further growth. This accounts for why under the condition of high yield of C_{60}, there aren't many fullerenes much larger than C_{60} and C_{70}. At temperature T the collision frequency is

$$\nu = \frac{1}{2}\rho^2 S \left(\frac{16k_BT}{\pi m}\right)^{1/2} \tag{6.6}$$

where S is the area of the cluster. [27] So the collision frequency is reduced by the factor of S_{cage}/S_{ring}. For C_{60} it's (38/222) $\sim$ 0.17. Endohedral metallofullerenes can be produced rather readily, since most of the time the clusters are baskets ready to scoop any nearby metal ions or atoms.

The ring road is so far the only theory that agree with the experimental observations in synthetic conditions, cluster size distributions, ion chromatography, no overshooting, endohedral metallofullerenes, etc.

By assuming the existence of the intermediate like C_{30} fullerene or C_{30} graphitic sheet, both the fullerene and pentagon road fatally missed the history of carbon clustering process of the fullerene synthesis. Namely, birth in atomic carbon, and evolve under huge density and temperature gradient. In terms of the underlying physical forces, the pentagon road missed the dangling bond history, while fullerene road missed the fact of high strain, as well as the dangling bond history. Both of them have a hard time explaining the overwhelming yield of C_{60} over others. The ring road is consistent with both the dangling bond and strain energy considerations. It incorporates the gain in the $\sigma-$bonding gradually, following the history of the growth path. Ring road is the only theory to date that is consistent with all the experimental observations.

6.6.2 Hints for Experiments

Based on our analysis, we arrived at the following suggestions for synthesizing fullerenes.

> (1)*Avoid too large a density/pressure, either by too high a laser intensity or by too large a buffer gas pressure. Because then the drop in the collision frequency upon isomerization into cage structure won't have dominate effect in stopping the further growth into giant fullerenes and nonfullerenes, which are not stable.*

The crucial step is to initiate the conversion from sp^1 carbons into sp^2 carbons. The collision of two monocyclic rings initiates the first four sp^2 carbons. The Bergmen cyclization is an important mechanism for the subsequent conversions. Warner *et al.* have demonstrated that in 1,2-bis(diphenyl phosphinoethynyl)benzene systems, adding $PtCl_2$ or $PdCl_2$ can accelerate the Bergman cyclization by a factor of $> 30,000$ while adding $HgCl_2$ can inhibit it. [28] We could envision $PtCl_2$ or $PdCl_2$ as molecular pliers in assisting the closure of the six-carbon rings. So our second suggestion is

(2)*Incorporate $PtCl_2$ or $PdCl_2$ in the arc environment, maybe by soaking the graphite rod with them, to accelerate the fullerene formation.*

6.7 Discussions and Summary

Neutral Atom Versus Ions

Experiments can explore only the charged particles. Positive ion and negative ion have different behaviors. Diederich's group have shown that the population distribution arising from the coalescence of cycle-carbons are different for positive and negative ions. [19] "Previous experience [29] indicates that the cluster distributions seen with residual *cations* is more similar to that of the *neutrals* than is the residual *anion* distribution. Thus, using cations seems more appropriate to a discussion of the neutral distribution." [25] Our computer simulations can explore the neutral species, as well as ions. Most of the cluster ions are open-shell system. The quantum chemistry methods for the open shell system are readily available. Their computation cost is, on average, always higher than that of closed shell. We leave the investigations of the ions similar to the neutral clusters to future work.

Kinetics versus Thermodynamics

The distinction between the kinetics and thermodynamics quantities is dependent upon the level of description of the system. Using the hierarchy models, we arrived at energetics of a complete pathway from two C_{30} monocyclic ring to C_{60} fullerene. Why C_{60} is so stable and how C_{60} fullerenes are formed are two problems distinguished from but related to each other. Together they explain the existence of C_{60} fullerenes. We concluded that the formation of fullerenes in the expansion of the super-heated carbon vapors is an evolutionary process. History is as important as the thermodynamics of the species. We have provided the energetics of the species at various era of this evolution. Kinetics determines, *in a statistical way*, the source availability, the time each metastable species last in existence, while thermodynamics determines which metastable species would survive and make it to the soot.

For a complete picture of this evolution process, one needs to have the density, temperature or pressure profile from the center of the arc to the wall of the vacuum chamber. This profile $\rho_{C_n}(\mathbf{r})$ relates the microscopic source availability to the macroscopic control. This density profile should be derived from a diffusion type of equation that takes into account the possibility of chemical reactions, i.e., change in the particle numbers.

Methodology

In this study we used a combined the DFT and MD method. We found that this is the most efficient scheme to tackle the large system problems. As we have shown, fine grained theories, e.g., *ab initio* method, are used for getting parameters for the coarse grained models. As our computation power increases, we can calibrate larger and larger systems for their more and more detailed behaviors. However, the change of paradigms, the change of scales, the efficient incorporation of existing knowledge of the system, should always start from modeling. The combination of modeling with first principle method will remain the most efficient and intelligent approach for studying material properties.

Acknowledgments

We would like to thank Dr. Jean-Marc Langlois for various help in using PS-GVB and Dr. C.-H. Kiang for valuable discussions

Bibliography

[1] T. W. Ebbesen, J. Tabuchi and K. Tanigaki, Chem. Phys. Lett. **191** 336, (1992)

[2] J. M Hunter, J. L Fye, E. J. Roskamp, and M. F.Jarrold *J. Phys. Chem.* **98** 1810-1818, (1994);

[3] J. M Hunter, J. L Fye, and M. F.Jarrold *J. Chem. Phys.* **99** 1785-1795, (1993)

[4] G. V. Helden, N. G. Gotts and M. T. Bowers, *Nature* **363** 60, (1993); G. V. Helden, M-T. Hsu, N. Gotts, and M. T. Bowers, *J. Phys. Chem* **97** 8182-8192, (1993)

[5] H. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curt, and R. E. Smalley, *Nature* **318**, 162-163 (1985);

[6] W. Krätschmer, K. Fostiropolos, and D. R. Huffman, *Chem Phys. Lett.* **347**,354-358 (1990)

[7] P. R. Buseck, Tsipursky and R. Hettich, Science **257** 215-217, (1992)

[8] J. R. Howard, J. T. McKinnon, Y. Makarovksy, A. Lafleur and M. E. Johnson, *Nature Lond.* **352** 139-141 (1991)

[9] S. Iijima *Nature* **354**, 56 (1991)

[10] CHK and WAG Buckytube

[11] G. E. Scuseria, Science, **271** 942, 1996

[12] E. A. Rohlfing, D. M. Cox and A. Kaldor, J. Chem Phys. **81** 3322, (1984)

[13] M. F. Jarrold and J. E. Bower, J. Chem. Phys. **96**, 9180 (1992);

[14] B. G. Johnson, P. M. W. Gill, and J. A. Pople, *J. Chem. Phys.* **98** 5612, (1993).

[15] Murco N. Ringnalda, Jean-Marc Langlois, Burnham H. Greeley, Robert B Murphy, Thomas V. Russo, Christian Cortis, Richard P. Muller, Bryan Marten, Robert E. Donnelly, Jr., Daniel T. Mainz, Julie R. Wright, W. Thomas Pollard, Yixiang Cao, Youngdo Won, Gregory H. Miller, William A. Goddard III, and Richard A. Friesner, *PS-GVB v2.2, Schrödinger, Inc.*, (1995)

[16] Y. Guo, *Ph.D. Thesis*, California Institute of Technology. (1992)

[17] R. G. Bergman *Acc. Chem. Res.* **6** 25, (1973)

[18] J. Hutter, H. Lẗhi, and F. Diederich, J. Am. chem. Soc. **116** 750-756 (1994)

[19] S. W. McElvany, M. M. Ross, N. S. Goroff, F. Diederich, Science **259**, 1594-1596 (1993)

[20] A. J. Stone and D. J. Wales, Chem. Phys. Lett. **128**, 501,(1986)

[21] R. L. Murry, D. L. Strout, G. K. Odom, and G. E. Scuseria, Nature **366** 665, (1993)

[22] H. A. Kramers, Physica **VII**,no.4 p284, (1940)

[23] B. Carmeli and A. Nitzan, Phys. Rev. Lett. **51**, 233, (1983)

[24] R. E. Smalley, Acc. Chem. Res. **25**, 98-105, (1992)

[25] R. F. Curl, Phil. Trans R. Soc. Lond. A **343** 19-32 (1993)

[26] J. R. Heath, ACS Symp. Ser. **481** 1-23, (1991)

[27] L. E. Reichl, *A Modern Course in Statistical Physics* University of Texas Press, Austin, p459 (1980)

[28] B. P. Warner, S. P. Millar, R. D. Broene, and S. L. Buchwald, Science, **269**, 814 (1995)

[29] S. C. O'Brien, J. R. Heath, H. W. Kroto, R. F. Curl, and R. E. Smalley, Chem. Phys. Lett. **132** 99-102 (1986)

Chapter 7 Molecular Dynamics Simulation of the $YBa_2Cu_3O_7$ High-Tc superconductors

7.1 Introduction

High-Tc material are characterized by their complex structures which have delicate balances between various forces of its constituents. The understanding of the molecular dynamics properties is desirable for tailoring the materials to improve its mechanical properties and thermal properties. [1] It also sheds light on the superconducting mechanism, which is crucial for synthesizing materials of higher superconducting transition temperature. There has been some DFT calculations on YBCO and LaCuO that employ frozen-phone technique [2] which can give the pretty good vibration at γ-point for systems at the equilibrium. However to study the stability in the whole Brillouin zone and to study the properties away from equilibrium one needs more expensive computation. This is currently beyond the reach of DFT level. Due to the complexity of the HTSC system, empirical force field approach as a much more economic approach is still the choice.

A good force field should be able to model the multi-dimensional potential energy of atoms (PES) well. But the only way to check the PES for system more complicated than diatomic molecules is by comparing it to the *ab initio* PES, which is not available for large complicated systems like YBCO family. In practice the quality of a force field is judged by its predictions for a finite set of experiments, and by its simplicity. Based on physical arguments one employs a functional form and tune the parameters by matching the results of FF to those from experiment. One can then use the force field to study situations unaccessible for experiment or to gain insight about the microscopic mechanism of the crystal structure. This is the approach we take in the chapter.

In section 7.2 we derive the force field for YBCO system, we propose the functional and fit the parameters. In section 7.3 we survey the lattice dynamical properties governed by this force field. In section 7.4 we run Molecular-Dynamics simulation to study thermal properties, with some comments on implications about the structural stability. We focus on $YBa_2Cu_3O_7$ system since it is the best characterized system among all the high temperature superconductors(HTSC).

7.2 The Force Field

7.2.1 Introduction

In view that $YBa_2Cu_3O_7$ is neither pure ionic nor pure covalent, we proposed an ionic-covalent force field. Considering the electronegativity [3] we assume bond only between Cu-O pairs. The ionic part is pair-wise potential which consists of a long-range Coulomb $1/r$ term, an instant dipole-dipole terms $-(1/r)^6$ and a Pauli term $(1/r)^{12}$. The covalent part consists of a bond stretching term and an angle bending term which takes up the leading three-body interaction. Our model has the following functional form:

$$E = \sum_{i<j} \frac{Q_i Q_j}{R_{ij}} + \sum_{i<j} D_{ij0} \Big[\Big(\frac{R_{ij0}}{R_{ij}} \Big)^{12} - \Big(\frac{R_{ij0}}{R_{ij}} \Big)^{6} \Big] + E_{bond} \tag{7.1}$$

where Q_i is charge; D_{ijo} R_{ijo} are van der Waals parameters;

$$E_{bond} = \frac{1}{2} k_\theta (cos\theta - cos\theta_0)^2 + \frac{1}{2} k_r (r - r_0)^2 \tag{7.2}$$

where k_r, r_0 are bond parameters and k_θ, θ_0 are angle parameters. Ewald sum is employed for the non-bonding terms.

The parameter are determined by requiring that the force field reproduce the observed structure, the structure be stable, and give best fit to the available vibration frequencies. More specifically we require that

(I) Zero force on every atom in the unit cell. Strictly speaking one should fit to the zero-temperature lattice parameter, which can be obtained by extrapolating the thermal expansion data. [4] However, this would require later either fitting zero temperature Raman spectra which is not available. So we fit the room-temperature lattice structure. [5] and allow the force field parameters to take up some of the thermal energies. namely, entropic contributions at room temperature.

(II) Zero stress on the unit cell. Stablization of the twin-free crystal requires that there be no stress on each unit cell. Since we are modeling perfect crystals we should put this constraint.

(III) No negative frequencies in the whole Brillouin zone. Below 700K we did not see structural instability in $YBa_2Cu_3O_7$. For the temperature range we are investigating there are no soft modes. Thus we require non-negative frequencies in the whole BZ.

(IV) Reproduce the experimental Raman frequencies(Γ point). The availability of good quality untwinned single crystal allows the measurement of the polarized Raman scattering [6], giving us a complete set of Raman data for fitting the FF parameters. [1]

The simulated structure is built on SGI work-station with Molecular Dynamics package POLYGRAF of MSI. According to their chemical properties in the crystal we employ four types of O ions, two types of Cu ions, one type of Ba and one type of Y. See Figure 7.1 for the unit-cell structure. As our initial guess we take the value from the work of Chaplot. [9] Further adjustments proceed as follows. From Raman mode involving *only* Y and Ba we get the charges for Y and Ba, since for such modes the charge on Y and Ba set the scale of the frequencies. At each Y and Ba charge value, we adjust other parameters in the force field to maintain the stable structure. The effective charge on Y and Ba thus determined are 1.56 and 1.45, respectively. The

---

[1] The exact procedure would be to calculate the frequency via atomic correlation functions at room temperature. But this procedure is expensive and may not be accurate enough.

Charge	$Q_i(e)$
Y	1.55600
Ba	1.45100
Cu1	1.35200
Cu2	0.83860
O1	-1.11800
O2	-0.93780
O3	-0.93780
O4	-1.50000

Table 7.1: Charge parameters for YBa2Cu3O7 The atom labels are in Figure 7.1.

charge on O and Cu are adjusted subsequently to preserve the neutrality. Once the effective charges are determined, all the other short-ranged terms in Eqs.(7.1) and (7.2) are further tuned to meet requirements (I), (II), (III) and (IV).

The non-linear fitting of the force field parameters can be cast into a multi-dimensional optimization problem. Here we use singular-value-decomposition (SVD)method for the optimization, implemented by Terumasa Yamaguchi. In non-linear fitting, the dimensionality of the parameters space is more complicated than that of the linear fitting, where fitting N parameters requires N and only N data points.

7.2.2 Results

Parameters that meet all the above requirement are listed in Table 7.1 - 7.3 .

Two distinguished features we discovered are as follows.

(a)There is a huge bonding term between Cu1 and chain oxygen O1 along b axis. Cu1-O4 is absent in $YBa_2Cu_3O_6$, the insulator cousin of $YBa_2Cu_3O_7$, since there is no O4 in that system. The ab anisotropy characteristic of $YBa_2Cu_3O_7$ manifested itself clearly in this extra large Cu1-O4 bond. This large bonding of Cu1-O4 indicates the tendency for O4 atom to leave the system upon heating which is just what the phase diagram of $YBa_2Cu_3O_{6+x}$ showed. Wille *et al.* [11] showed that at high temperature ($\sim$ 1600K) the stable phase is tetragonal(P4mm), see also the work of Nozaki. [12]

VDW Parameter	$R_{0ij}(cm)$	$D_{0ij}(kcal/mol)$
O3 - O3	2.88626	0.09774
O2 - O2	2.69486	0.09278
O4 - O4	3.55326	0.09997
O1 - O1	3.17630	0.09678
Cu1 - Cu1	2.42398	0.10884
Cu2 - Cu2	2.47550	0.10631
Y - Y	3.80504	0.09879
Ba - Ba	4.04807	0.09849
O2 - O3	3.50751	0.10520
O4 - O3	3.57322	0.10080
O4 - O2	3.64945	0.09684
O1 - O3	3.20242	0.09867
O1 - O2	2.98598	0.10360
O1 - O4	3.70546	0.10637
Cu1 - O3	2.95250	0.09830
Cu1 - O2	2.67473	0.10042
Cu1 - O4	2.33823	0.10530
Cu1 - O1	2.56456	0.10363
Cu2 - O3	2.35880	0.10520
Cu2 - O2	2.30665	0.10160
Cu2 - O4	3.29345	0.10475
Cu2 - O1	3.01087	0.09991
Cu2 - Cu1	2.42007	0.09297
Y - O3	3.34383	0.09313
Y - O2	3.44952	0.09535
Y - O4	3.90925	0.10310
Y - O1	3.73971	0.09573
Y - Cu1	3.10892	0.10004
Y - Cu2	2.97369	0.10172
Ba - O3	2.68787	0.09481
Ba - O2	3.74382	0.09602
Ba - O4	4.00692	0.09856
Ba - O1	3.92018	0.10205
Ba - Cu1	3.49319	0.10140
Ba - Cu2	4.41050	0.11323
Ba - Y	3.59763	0.10237

Table 7.2: vdW Force field parameters for YBa2Cu3O7 The atom labels are in Figure 7.1.

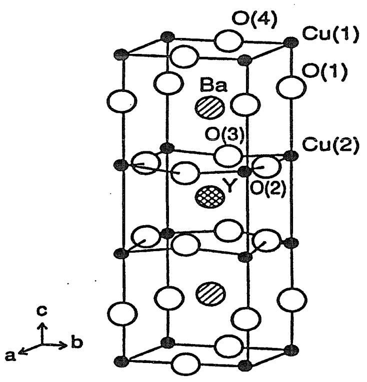


Figure 7.1: The unit cell of $YBa_2Cu_3O_7$. Two different types of Cu and four types of O atoms are used.

Bond Parameter	$k_r[(kcal/mol)/a^2]$	$r_0(a^2)$
Cu1 - O4	190.8385	1.914561
Cu1 - O1	0.9965	1.945591
Cu2 - O3	50.6204	1.971260
Cu2 - O2	31.1643	2.010588
Angle Parameter	$k_\theta[(kcal/mol)/rad^2]$	$\theta_0(deg)$
Cu2 - O3 - Cu2	0.0894	179.1897
Cu2 - O2 - Cu2	1.1871	152.1073
O1 - Cu1 - O4	0.0005	90.0000
O3 - Cu2 - O3	0.0218	141.1687
O2 - Cu2 - O3	104.4444	69.6829
O2 - Cu2 - O2	0.0075	167.6002

Table 7.3: Force field parameters for the bonding terms of YBa2Cu3O7. The atom labels are in Figure 7.1.

Symbol	Experimental[a]	The fitted
a	3.820	3.82
b	3.885	3.885
c	11.68	11.68
Y	0.500	0.500
Ba	0.184	0.184
Cu_1	0.000	0.000
Cu_2	0.355	0.355
O_1	0.158	0.158
O_2	0.378	0.378
O_3	0.377	0.377
O_4	0.000	0.000

Table 7.4: Our force field reproduce the lattice constants and fractional coordinates for atoms in $YBa_2Cu_3O_7$ unit cell.
[a] Reference [10].

(b)The most important three body term is the bond angle bending in Cu-O plane, the O3-Cu-O2 angle. This is exactly the internal coordinate that is involved in the celebrated $335cm^{-1}A_{1g}$ mode that softens upon superconducting transition. [13, 14] See the eigen-mode calculation in the following section. This signifies the importance of the O3-Cu-O2 local structure in the superconductor $YBa_2Cu_3O_7$.

7.3 Lattice Dynamics

Employ standard lattice dynamics we diagonalize the dynamic matrix.

$$D_{\alpha\beta}(\mathbf{q}|\kappa\kappa') = \frac{exp\{i\mathbf{q}\cdot[\mathbf{r}(\kappa') - \mathbf{r}(\kappa)]\}}{(m_\kappa m_{\kappa'})^{1/2}} \sum_L \phi_{\alpha\beta}(0\kappa|L\kappa')e^{i\mathbf{q}\cdot\mathbf{r}(L)} \tag{7.3}$$

by solving

$$\sum_{\kappa'\beta}[D_{\alpha\beta}(\mathbf{q}|\kappa\kappa') - \delta_{\alpha\beta}\delta_{\kappa\kappa'}\omega^2(\mathbf{q})]\epsilon_\beta(\kappa'|\mathbf{q}) = 0 \tag{7.4}$$

where

$$\phi_{\alpha\beta}(0\kappa|L\kappa') = \frac{\partial^2 E}{\partial u_\alpha(0\kappa)\partial u_\beta(L\kappa')} \tag{7.5}$$

with E from equation (7.1) and (7.2) and $u_\alpha = x_\alpha - x_{\alpha 0}$. L is Brava lattice index and κ is atomic index within a unit cell, and α, β denote Cartesian coordination.

7.3.1 Vibration Mode

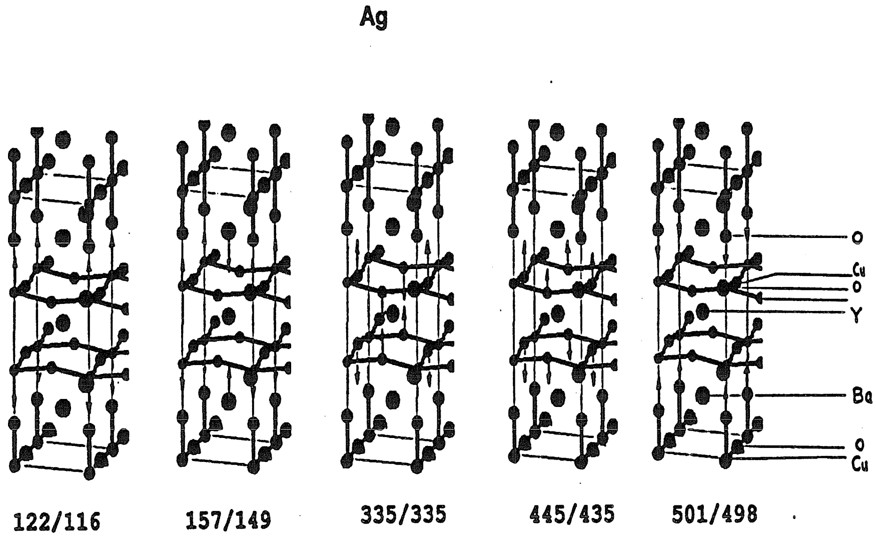


Figure 7.2: Ag Raman modes of $YBa_2Cu_3O_7$ vibrations. The first one quoted is from our calculation and the second ones are from experiments. The frequencies are in unit of cm^{-1}.

First we list the vibrations at the long-wavelength limit, Γ point, namely, $\mathbf{q} = \mathbf{0}$. The frequencies are shown in Table 7.5. Eigenvectors are shown in Figure 7.2 and 7.3. Table 7.5 are the comparison between our calculation and experimental data of $YBa_2Cu_3O_7$ [6, 7] for two isotopes of oxygen, ^{16}O and ^{18}O. [8] Only ^{16}O vibrations were used in the fitting. We can see that fitted force field reproduce the Raman frequencies of ^{16}O isotope very well and the predicted isotopic shifts for ^{18}O agree

Symmetry	^{16}O Raman freq. Theory	Expt.^{a}	Isotope shift theory	Expt.^{b}
A_{1g}	122	116	0.8	
	157	149	0.6	
	335	335	5.7	5.5(0.6)
	445	435	5.4	5.5(0.4)
	501	498	5.4	
B_{2g}	70	70	0.0	
	142	142	2.1	
	213	210	4.2	
	403		5.0	
	552	579	5.4	
B_{3g}	106	83	0.0	
	172	140	1.2	
	317	303	4.7	
	545		5.5	
	595	526	5.5	
$B_{1u}(LO)$	114	147	1.8	2.0(0.7)
	173	192	0.6	1.6(0.5)
	197		0.0	
	301		5.3	
	486		4.3	
	512		4.7	
	709		5.5	
$B_{2u}(TO)$	93		2.2	
	179		0.0	
	232		0.0	
	333		4.8	
	380		5.0	
	611		5.4	
	662	575	5.6	4.4(0.2)
$B_{3u}(TO)$	71		1.4	
	144		2.1	
	193		0.5	
	231		4.3	
	300		4.0	
	438		5.0	
	564		5.5	

a Reference [6]. b Reference [8].

Table 7.5: Calculated and measured Raman and IR frequencies of $YBa_2Cu_3O_7$. Also included are the isotope shifts upon substitution of ^{16}O by ^{18}O . Vibration Frequency is in cm^{-1}.

well with the measured ones. This indicates that this force field is reasonably good near the equilibrium. Due to the metallic nature many IR modes in $Cu-O$ plane are not experimentally observable. Our calculated complete set of IR mode should give some clue to the in-plane vibration IR-mode, and thus help experimentalists to seek and assign other modes.

7.3.2 Phonon Spectrum

Coherent Scattering Properties

Next we use this force field to calculate the phonon spectrum for any point in the Brillouin zone. Results on some of the symmetry directions are shown in Figure 7.4 and 7.5. We compare our calculated phonon spectrum with the measurements of Reichardt, which is made on twinned single crystals. [15] Along c-axis, where experiments give unambiguous symmetry assignments, our calculated spectrum agree well with the measured one. (see Fig. 7.5) For (100) and (010) direction experimental values are a mixture of the two directions so it is hard to do the comparison. Given that we see that most measured points fall on the calculated curves. (see Fig. 7.4.) For long acoustical wave it shouldn't make any difference whether we are dealing with twinned or untwinned sample. Our calculated LA mode agree well with the measured one. One of TA mode agrees well with experiment while the other is a little off. For (110) direction, the calculated TA mode also agrees well with the measured one. For optical modes some agree with measured ones, while some are too soft.

In our model all the ions carry nonzero charges so we expect to see LO-TO splitting (so-called LST effect) at Γ point from the force field calculation. Whether or not one can see LO-TO splitting experimentally is not clear at this point. In order to see that one needs very accurate coherent inelastic-neutron scattering on untwinned single-crystals.

Incoherent Scattering Properties

Using histogram we calculated the one phonon density of state. To compare with incoherent inelastic neutron measurement on powdered sample we need the *eigenvector weighted* phonon density of state. The theoretical neutron scattering intensity is given by the following formula: [16, 17]

$$\left(\frac{d^2\sigma}{d\omega d\Omega}\right)_{incoh} = \frac{k_1}{k_0}\frac{n(\omega)+1}{\omega}\sum_i \frac{\sigma_i}{m_i}e^{-2W_i}\sum_{qj}|\mathbf{Q}\cdot\epsilon(i|q,j)|^2\delta\left(\omega-\omega_j(q)\right) \tag{7.6}$$

where $\epsilon(i|q,j)$ is the orthonormal displacement of i th atom in mode $\omega(q,j)$; σ_i is the total neutron cross section for i th atom, here we take 3.76, 7.78, 3.42 and 7.76 for O, Cu, Ba, and Y nuclei respectively from S.F.Mughabghab [18]; m_i is the mass of ith atom; $n(\omega)$ is the Bose distribution for phonon population. $\mathbf{Q} = \mathbf{k_1} - \mathbf{k_0}$,with k_1 and k_0 being outgoing and incident neutron momentum respectively. $\mathbf{Q}$ is fixed in constant-Q mode of measurements. W_i is the Debye-Waller factor calculated as the following:

$$W_i = \frac{\hbar Q^2}{Nm_i}\sum_{qj}|\epsilon(i|q,j)|^2\frac{n(\omega_j(q))+1}{\omega_j(q)}. \tag{7.7}$$

N is the number of unit cells in the sample. Here we take it as the number of points in Brillouin zone when making our histogram. We take Q as midpoint between G and Z and $T = 300K$ in $n(\omega)$. Now our generalized density of states $G(\omega)$ is defined as follows:

$$G(\omega) = \sum_i \frac{\sigma_i}{m_i}e^{-2W_i}\sum_{qj}|\epsilon(i|q,j)|^2\delta(\omega-\omega_i(q)). \tag{7.8}$$

We take 8000 points in BZ. Further addition of points wouldn't change the feature of the $G(\omega)$. We compare our calculated $G(\omega)$ with that of measured by Renker *et al.* [19] See Figure 7.6 We noticed that most peaks and over all trends agree with the experimental data. Our calculation gives too large a peak at around $670cm^{-1}$ (83meV) which comes from the chain oxygen(O4) vibrations. This indicates the flaw of the part of our force field involving O4. We will discuss this further when we come to the thermal expansion calculation.

Long Wave Behaviors

Using standard lattice dynamical theory for long wavelength limit we calculated the elastic constants and bulk modulus. The elastic constants are defined through elastic energy, see P. Brüesch, Chapter 3. [17]

$$W = \frac{1}{2}\sum_{\alpha\beta}\sum_{\gamma\lambda} C_{\alpha\beta,\gamma\lambda}\epsilon_{\alpha\beta}\epsilon_{\gamma\lambda} \tag{7.9}$$

where the strain

$$\epsilon_{\alpha\beta} = \frac{1}{2}\Big(\Big(\frac{\partial u_\alpha}{\partial x_\beta}\Big)_0 + \Big(\frac{\partial u_\beta}{\partial x_\alpha}\Big)_0\Big) \tag{7.10}$$

are symmetrized deformation tensor, $\mathbf{u}$ are displacement vectors with $\mathbf{u}|_0 = 0$, and $\mathbf{x}$ are position vector in a continuum medium.

Elastic constants are related to Hessians (atomic force-constants)by

$$C_{\alpha\gamma,\beta\lambda} = \tilde{C}_{\alpha\beta,\gamma\lambda} + \tilde{C}_{\gamma\beta,\alpha\lambda} - \tilde{C}_{\gamma\alpha,\lambda\beta} \tag{7.11}$$

$$\tilde{C}_{\alpha\beta,\gamma\lambda} = -\frac{1}{2V_a}\sum_k \phi_{\alpha\beta}(0|k) r_{k\gamma} r_{k\lambda} \tag{7.12}$$

where $\phi_{\alpha\beta}(0|k)$ is from Eq.(7.5). Here we used the short note $k = (L, \kappa)$ to represent the site of atoms. $r_{k\gamma}$ are position vector for lattice site k, V_a is the volume of the unit cell, subscript 0 denotes equilibrium value.

Defining notation 11 → 1, 22 → 2, 33 → 3, 23 → 4, 31 → 5, 12 → 6. e.g. $C_{12,32} = C_{64}$, the bulk modulus can be defined as

$$BM = (C_{11} + 2C_{12})/3. \tag{7.13}$$

	C_{11}	C_{33}	C_{44}	C_{66}	C_{12}	C_{13}	B
Theory	194	131	64	166	141	61	94
Exper[a]	230	150	50	85	100	100	120

[a] is from Ref. [20]

Table 7.6: Elastic stiffness constants (Gpa) of $YBa_2Cu_3O_7$. Experiments were on twinned samples with apparent $D4h$ symmetry. Thus to compare with experiment we set $C_{11} = (C_{11} + C_{22})/2$, $C_{44} = (C_{44} + C_{55})/2$, $C_{14} = (C_{13} + C_{23})/2$. B is the bulk modulus.

The result is listed in Table 7.6. The measurement [15] is made on a twinned sample, so we compare our $(C_{11}+C_{22})/2$ and $(C_{44}+C_{55})/2$ to their measured C_{11} and C_{44} respectively. The agreement is fair except for C_{66}. The deviation of C_{66} might be due to the fact that sheer component is more sensitive to the defect structure. Experimental values are measured on samples containing defects while the calculated values are for perfect crystals. In fact, defects tend to decrease the sheer modulus. The measured C_{66} gives smaller value than the calculated one, as one would expect.

7.4 Thermal Expansion

Using our force field we did Parrinello-Rahman-Nosé (PRN) molecular dynamics [4] to calculate the thermal expansion. Parrinello-Rahman-Nosé dynamics is a constant-stress constant-temperature molecular dynamics.(NPT ensemble) Let the original Hamiltonian be H_0

$$H_0 = \sum_i \mathbf{P}_i^2/2m_i + E(\{\mathbf{R}_i\}) \tag{7.14}$$

The PRN Hamiltonian H is

$$H = \sum_i \mathbf{P}_i^2/2m_i V^{\frac{2}{3}} s^2 + E(\{V^{\frac{2}{3}}\mathbf{R}_i\}) + p_s^2/2Q + gkTlns + p_v^2/2W + P_{ex}V \tag{7.15}$$

where s are the scale factor for time as required by the constant-temperature ensemble; V is the volume variable. Q and W are corresponding mass for the dynamical variable

s and V. p_s and P_{ex} are conjugate momentum of s and V, see Nosé [5] for more details.

It can be proved, within the ergodic assumption, that the average over the trajectories generated by the dynamical equations derived from this extended H equals the statistical NPT ensemble average of the original H_0. [23, 4, 5]

We run the Nosé dynamics using the dynamics code of POLYGRAF of MSI. We use the periodic boundary condition (PBC) with a supercell of 3 × 3 × 3 unit cells with 351 atoms moving independently. Cell mass W was chosen to be 0.0625. At each temperature we give an initial velocity according to the Maxwell distribution of that temperature and run for 10ps(10000 steps). Repeat for several temperatures from 10K to 700K. In all our runs the structure are very stable, a nontrivial feat for simulating excited states of a complex system. For each temperature we average the lattice constants over these 10000 configurations. We then calculated the volume expansion, dV/dT, around the room temperature. Our calculated $0.011\mathring{A}^3/K$ is close to the experimental value of $0.014\mathring{A}^3/K$ [4].

We also noticed that our force field predicts a decrease of lattice constant b upon heating, contradictory to the observed results. [4, 24] This implies that the Cu-O chain prefer bending to stretching and indicates that the $(1/r)^{12}$ term we chose for Pauli repulsion is too stiff, leading to too stiff a tension along Cu-O chain. Considering that $YBa_2Cu_3O_7$ will lose its chain O gradually and undergoes a transition at $500C^o(\sim 800K)$ this strong tension along Cu-O chain is reasonable.

Conclusion

In summary we have arrived at a force field for $YBa_2Cu_3O_7$ which incorporates most updated experimental results on this system. The force field reproduce the observed structure and give stable dynamics at temperature as high as 1000K. It reproduces the observed Raman spectrum to within a few cm^{-1} and gives the correct prediction on the isotope shift of Raman frequencies. The predicted phonon spectrum in the whole Brillouin zone agree well with the available measurements. It gives good elastic constants. Unlike harmonic FF, our FF can predicts the thermal expansion coefficient very close to the measured one.

From our force field we gained insight about the crucial energy terms. For structural stability the Cu1-O4 interaction is very important; For superconductivity the angle term between O2-Cu2-O3 is a key indicator. [14]

Acknowledgment We would like thank Dr. Karasawa for the helps regarding POLYGRAF in this work.

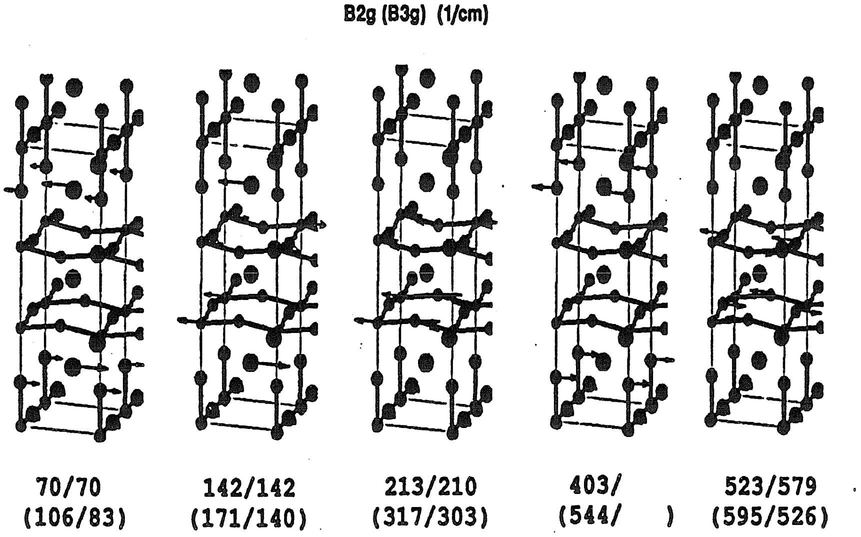


Figure 7.3: Bg Raman modes of $YBa_2Cu_3O_7$ vibrations. The first one quoted is from our calculation and the second ones are from experiments. The frequencies are in unit of cm^{-1}.

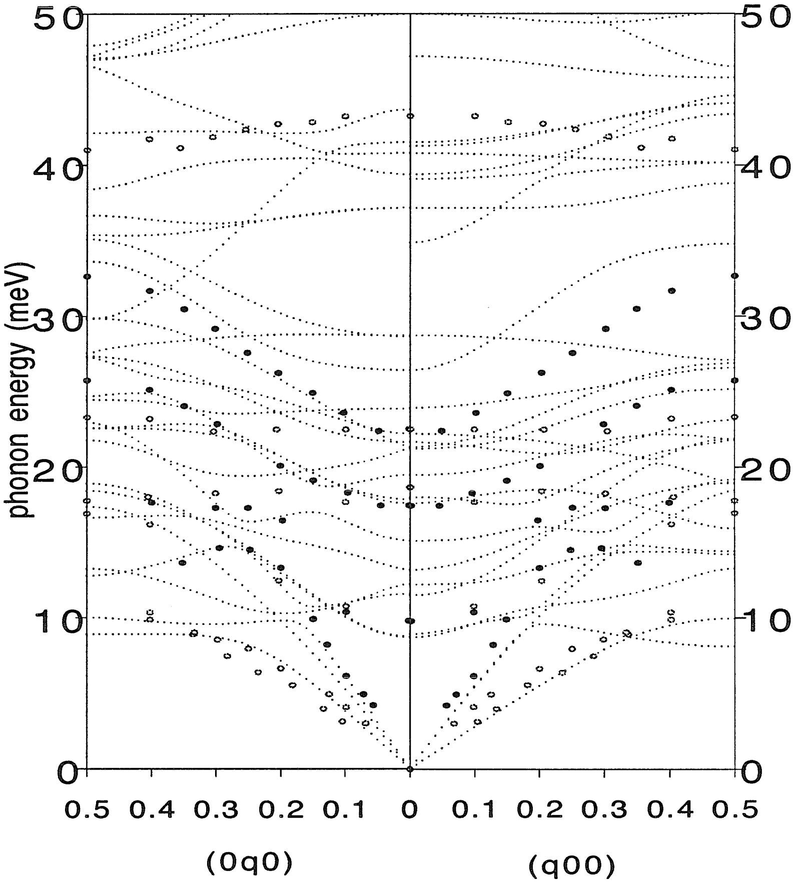


Figure 7.4: Phonon dispersion relation along (100) and (010) direction. The measured ones are in dots and are from the twinned single crystals. Lines are from our calculation.

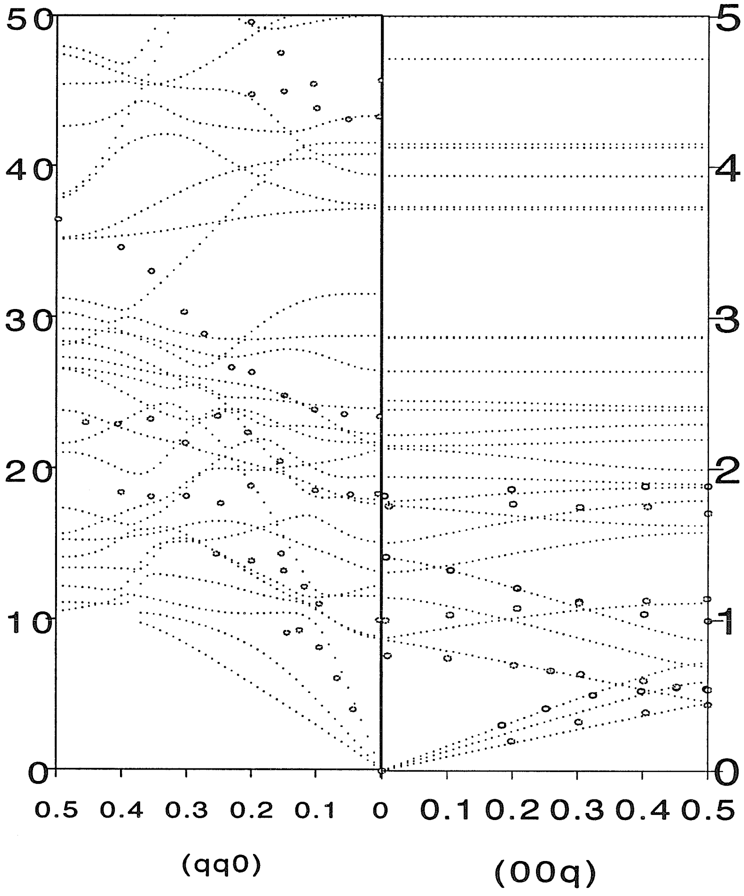


Figure 7.5: Phonon dispersion relation along (110) and (001) direction. The measured ones are in dots and are from the twinned single crystals. Lines are from our calculation.

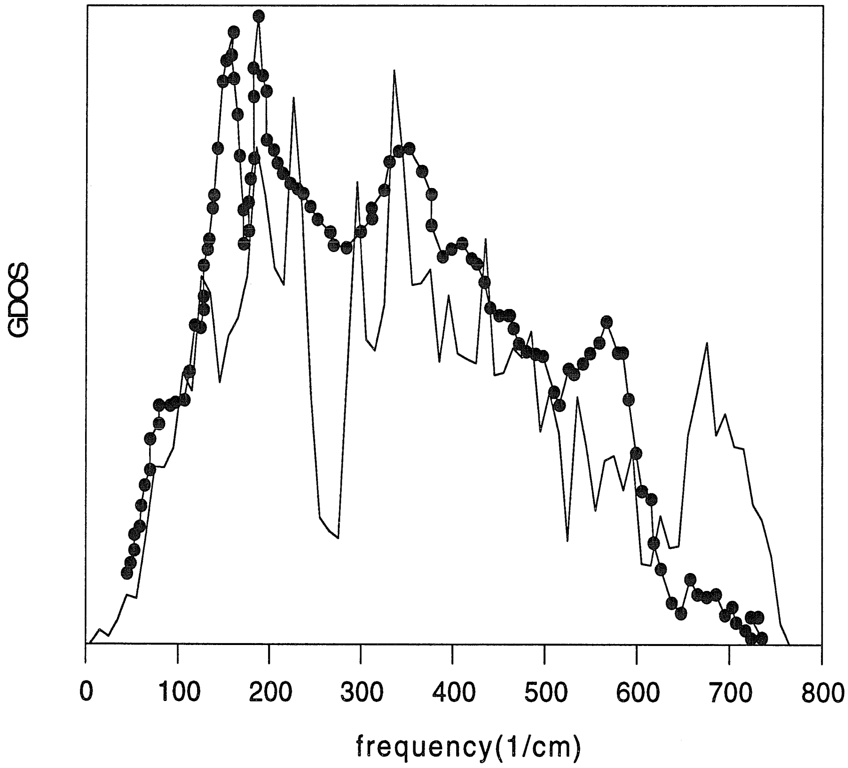


Figure 7.6: Generalized density of state, weighted by the eigenvectors and neutron scattering cross-section. Dots are from incoherent neutron scattering data on powered samples and the thin line is from our molecular dynamics calculation.

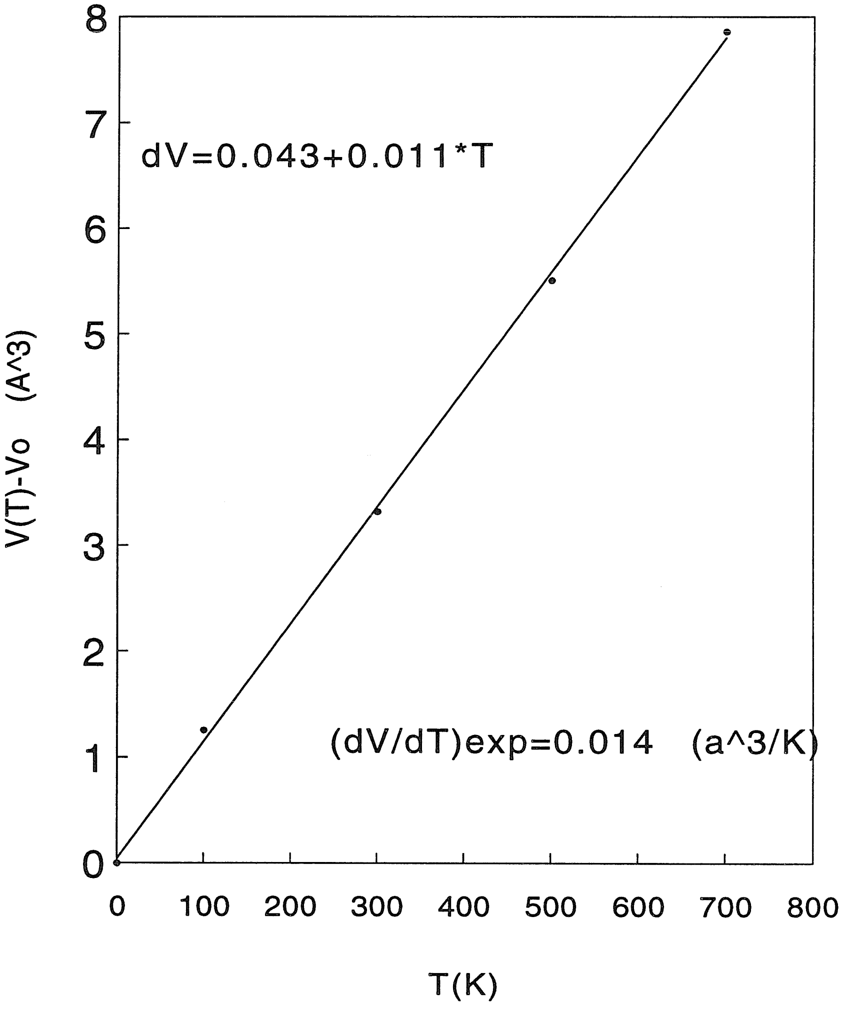


Figure 7.7: $YBa_2Cu_3O_7$ thermal expansion in volume. The line is a fit to the calculated points. Experimental values are quoted for comparison.

Bibliography

[1] R. Feile, Physica C **159** (1989) 1-32

[2] R. E Cohen, W. E. Pickett,and H. Krakauer Phys. Rev. Lett. **62**, 831 (1989) and R. E. Cohen, W. E. Pickett, and H. Krakauer Phys. Rev. Lett. **64**, 2575 (1990)

[3] L. Pauling, Phys. Rev Lett. **59**(2) 225 (1987)

[4] E. D. Specht, C.J. Sparks, A. G. Dhere, J. Brynested. O. B. Cavin, D. M. Kroeger, and H. A. Oye, Phys. Rev. B **37**, 7426 (1988)

[5] F. Beech, S. Miraglia, A. Santoro,and R. S. Roth, Phys Rev B **35** 8778-8781 (1987);

[6] K. F. McCarty, J. Z. Liu, R. N. Shelton, H. B. Radousky, Phys. Rev B **41** 8792 (1990)

[7] C. Thomsen, M. Cardona, R. Liu, B. Gegenheimer, and A. Simon, Physica C **153-155** 1756 (1988)

[8] C. Thomsen, H. Mattausch, M. Bauer, W. Bauhofer, R. Liu, L. Genzel,and M. Cardona, Solid State Comm. Vol. **67**, No. 11, 1069 (1988)

[9] S.L. Chaplot , Phys.Rev.B **37** 7435-7442 (1988)

[10] R. Beyers and T. M. Shaw, *Solid State Physics, Vol. 42*, p142. Academic Press, Inc. (1989)

[11] L.T. Wille, A. Berera, and D.de Fontaine, Phys. Rev. Lett. **60**, 1065 (1988)

[12] Hanae Nozaki and Satoshi Itoh in Phys. Rev. B **48** 7583 (1993)

[13] T.Ruf, C.Thomsen, R.Liu, and M.Cardona Phys Rev. B **38** 11985 (1988)

[14] M. Krantz, H. j. Rosen, R. M. Macfarlane, and V.Y. Lee Phys. Rev. B **38** 4992, (1988)

[15] W. Reichardt, N. Pyka, L. Pintschovius, B. Hennion, G. Collin, Physica C **162-164** 464 (1989)

[16] H. Boutin and S. Yip "Molecular Spectroscopy with Neutrons" Chapter 3. MIT press 1968. or P.Bresch "Phonons: Theory and Experiments II" Chapter 6.3 Springer-Verlag 1986.

[17] P.Brüesch and W. Bührer, Z. für Physik B-Condensed Matter **70**, 1 (1988)

[18] S. F. Mughabghab, M. Divadeenam, N. E. Holden, *Neutron Cross Sections–neutron resonance parameters and thermal cross sectios* Academic Press (1981).

[19] B. Renker, F. Gompf, E. Gering, G. Roth, W. Reichardt, D. Ewert and H. Rietschel, Physica C 153 (1988); J. J. Rhyne, D. A. Neumann, J. A. Gotaas,and F. Beech, L. Toth S. Lawrence, S. Wolf, M. Osofsky, and D. U. Gubser, Phys. Rev. B **36** 2294 (1987)

[20] W. Reichardt, L. Pintschovius, B. Hennion and F Collin, Supercond. Sci. Technol. **1** 173 (1988)

[21] M. Parrinello, A. Rahman, J. Appl. Phys. **52**, 7182 (1981)

[22] Shuichi Nosé, J. Chem. Phys. **81** 511 (1984)

[23] Hans C. Anderson, J. Chem. Phys. **72** 2384 (1980)

[24] K. Usami,N. Kobayashi and T. Doi, Japanese J. of Applied Physics **30**, L96-L98 (1990)

Chapter 8 Appendix

8.1 Second Quantization

In the language of second quantization, omnipotent Hamiltonian is

$$H = \sum_{ij\sigma} f_{ij} a^{+}_{i\sigma} a_{j\sigma} + \sum_{ijkl} \sum_{\sigma\sigma'} V_{ijkl} a^{+}_{i\sigma} a^{+}_{j\sigma'} a_{k\sigma'} a_{l\sigma}. \tag{8.1}$$

HF ground state is

$$|\Psi^{HF}\rangle = \Pi^{occ}_{n\sigma} c^{+}_{n\sigma} |0\rangle \tag{8.2}$$

with

$$c^{+}_{n\sigma} = \sum_{\mu=1}^{L} d_{n\mu} a^{+}_{\mu\sigma}. \tag{8.3}$$

Unlike the exact solution of an approximate Hamiltonian, in HF we approximate the wave function, or reinterpret some of the propagators.

$$H = \sum_{ij\sigma} f_{ij} (a^{+}_{i\sigma} a_{j\sigma} - \langle a^{+}_{i\sigma} a_{j\sigma} \rangle) + E_0 + H_1 \tag{8.4}$$

with

$$\begin{aligned} H_1 &= \frac{1}{2} \sum_{ijkl} \sum_{\sigma\sigma'} V_{ijkl} a^{+}_{i\sigma} a^{+}_{k\sigma'} a_{k\sigma'} a_{l\sigma} && (8.5) \\ & - \sum_{ijkl\sigma} (V_{ijkl} - \frac{1}{2} V_{ilkj}) P_{kl} a^{+}_{i\sigma} a_{j\sigma} + \frac{1}{2} \sum_{ijkl} (V_{ijkl} - \frac{1}{2} V_{ilkj}) P_{ij} P_{kl} && (8.6) \end{aligned}$$

where

$$
\begin{aligned}
P_{ij} &= \sum_{\sigma} \langle a^{+}_{i\sigma} a_{j\sigma} \rangle &(8.7)\\
&= 2\sum_{n}^{occ} d^{*}_{ni} d_{nj}. &(8.8)
\end{aligned}
$$

If the basis is $|\mathbf{r}\rangle$ the above relation is written as

$$
\begin{aligned}
\rho(\mathbf{r},\mathbf{r}') &= \langle \psi^{+}(\mathbf{r})\psi(\mathbf{r}') \rangle &(8.9)\\
&= 2\sum_{n}^{occ} \phi_n(\mathbf{r})\phi_n(\mathbf{r}'). &(8.10)
\end{aligned}
$$

This relation holds even in interactive ground state $|\Psi\rangle$, as can be proved from Green's function.

$$\langle \Psi | \hat{\rho}(\mathbf{r},\mathbf{r}') | \Psi \rangle = \pm i tr G(\mathbf{r}t, \mathbf{r}'t^{+}) \tag{8.11}$$

tr is trace over the spin indices.

8.2 Calculus of Variations

In analogy to the total differential of a function $F(f_1, f_2, ...)$: $dF = \sum_i (\partial F/\partial f_i) df_i$, in the continuum limit one gets the functional difference [1]

$$\delta F = \int \frac{\delta F}{\delta f(x)} \delta f(x) dx \tag{8.12}$$

with

$$\frac{\delta F[f]}{f(x)} = g(x), \tag{8.13}$$

a function of x. Consider the functional [2]

$$F[\rho] = \int f(x, \rho, \rho^{(1)}, \rho^{(2)}, ..., \rho^{(n)}) dx \tag{8.14}$$

where $\rho^{(n)}(x) = d^n\rho(x)/dx^n$, then

$$\frac{\delta F}{\delta\rho(x)} = \frac{\partial f}{\partial\rho} - \frac{d}{dx}\Big(\frac{\partial f}{\partial\rho^{(1)}}\Big) + \frac{d^2}{dx^2}\Big(\frac{\partial f}{\partial\rho^{(2)}}\Big) - ... + (-1)^n\frac{d^n}{dx^n}\Big(\frac{\partial f}{\partial\rho^{(n)}}\Big). \tag{8.15}$$

For three-dimensional generalization, one replaces the operator d/dx with ∇. So for

$$F[\rho] = \int f(\rho, |\nabla\rho|)dx, \tag{8.16}$$

defining $\gamma = |\nabla\rho|$ leads to

$$\frac{\delta F}{\delta\rho(x)} = \frac{\partial f}{\partial\rho} - \nabla\cdot\Big(\frac{1}{\gamma}\frac{\partial f}{\partial\gamma}\nabla\rho\Big). \tag{8.17}$$

Symmetry argument implies that E_x depend only on $|\nabla\rho|$.

8.3 Car-Parrinello Ab initio Molecular Dynamics

Solving electronic Kohn-Sham equations under BO approximation imposes N(N-1)/2 constraints between two sets of otherwise independent variables $\{\phi_i\}$ and $\{\mathbf{R}\}$. These constraints put the system in BO surface. The resulting Lagrangian will then be only a functional of $\{\mathbf{R}_\mu\}, \{\dot{\mathbf{R}}_\mu\}$:

$$L(R, \dot{R}, t) = \sum_\mu \frac{1}{2}M_\mu\dot{\mathbf{R}}_\mu^2 - E(\{\mathbf{R}_\mu\}) \tag{8.18}$$

with

$$E(\{\mathbf{R}_\mu\}) = \sum_i \int d\mathbf{r}\psi_i^*(\mathbf{r})[-(\hbar^2/2m)\nabla^2]\psi_i(\mathbf{r}) + \bar{U}[n(\mathbf{r}), \{\mathbf{R}_\mu\}, \{\alpha_\nu\}] \tag{8.19}$$

where ψ_i is the eigenvector of Kohn-Sham equation:

$$\Big[-\frac{\hbar^2}{2m}\nabla^2 + \frac{\delta\bar{U}}{\delta n(\mathbf{r})}\Big]\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}). \tag{8.20}$$

Now, we loosen up those constraints and make $\{\phi_i\}$ and $\{\mathbf{R}_\mu\}$ independent variables. We will have a finite thickness of BO surface, either by the changes in the occupation number (the statistical description being the Fermi-Dirac distribution), or by changes the orbitals themselves, or both. The electronic variables $\{\phi_i\}$ will not be constrained at the Born-Oppenheimer surface but will deviate from it somewhat. We design the restoring force to be proportional to their deviation from the BO surface. Thus we have the Car-Parrinello Lagrangian, which is a functional of both $(\{\mathbf{R}_\mu\}, \{\dot{\mathbf{R}}_\mu\})$ and $(\{\phi_i\}, \{\dot{\phi}_i\})$.

$$L(\mathbf{R}, \dot{\mathbf{R}}, \{\phi_i\}, \{\dot{\phi}_i\}, t) = \sum \frac{1}{2}\mu \int d\mathbf{r}|\dot{\phi}_i|^2 + \sum \frac{1}{2} M_\mu \dot{R}_\mu^2 + \sum_\nu \frac{1}{2}\mu_\nu \dot{\alpha}_\nu^2 - E(\{\phi_i\}, \{\mathbf{R}_\mu\}\{\alpha_\nu\}) \tag{8.21}$$

$\{\phi_i\}$ and $\{\mathbf{R}_\mu\}$ are now independent variables. The important fact is that this $\bar{U}[\{\mathbf{R}_\mu\}, \{\phi_i\}]$ *is* the variational functional whose Euler-Lagrange equations give the equations of motion for $\{\phi_i\}$, as demonstrated by the success of Kohn-Sham scheme.

Since the Lagrangian doesn't have explicit time dependent it is a constant of time. The Hamiltonian thus defined is a conserved quantity. In the original Car-Parrinello version this quantity doesn't have a physical meaning due to the fictitious kinetic energy, $(1/2)\int \mu|\dot{\phi}_i|^2$. It is introduced for purely algorithmic purposes, But actually, part of the kinetic energy *is* real. As demonstrated by Blöchl and Parrinello, [3] there is a portion related to the orbitals moving with the ions, which they call adiabatic kinetic energy, that does have a physical meaning. They designed a Parrinello-Rahman type of dynamics [4] where the orbital kinetic energy is maintained around this adiabatic kinetic energy via another thermostat, different from the one for ions. This new design has been shown to be very efficient in enlarging the time steps of the ab initio MD. Instead of employing thermostat, which is a statistical way of incorporating our knowledge, can we build a more fundamental dynamics for the orbitals? Here we propose such a dynamics for the cases involving one degree of freedom of ion.

If we keep the gradient terms in using Born-Oppenheimer approximation we will have(see the Appendix Eq.(0.16)):

$$E = \sum_{\mu} \frac{P_{\mu}^{2}}{2M_{\mu}} + \sum_{\mu} \frac{\hbar^{2}}{2M_{\mu}} \sum_{i}^{occ.} \int dr |\nabla_{\mu} \psi_{i}|^{2} + U(\{\mathbf{R}_{\mu}\}). \tag{8.22}$$

Note that the ∇_{μ} is w.r.t. R_{μ}.

For systems where there is only one ionic degree of freedom, we have:

$$\dot{\psi}_{i} = \frac{\partial \psi_{i}}{\partial R} \dot{R} \tag{8.23}$$

$$\begin{aligned} &for \quad \dot{R} \neq 0 \quad \tfrac{\partial \psi_{i}}{\partial R} = \tfrac{\dot{\psi}_{i}}{\dot{R}} \\ &for \quad \dot{R} = 0 \quad \dot{\psi}_{i} = 0. \end{aligned} \tag{8.24}$$

Then the Lagrangian corresponding to Eq.(8.22) is:

$$L(R, \dot{R}, t) = \frac{\hbar^{2}}{2M\dot{R}^{2}} \sum_{i}^{occ.} \int d\mathbf{r} |\dot{\psi}_{i}|^{2} + \frac{1}{2} M \dot{R}^{2} - U[R; \{\psi_{i}\}] \tag{8.25}$$

where ψ_{i} is the solution of Eq.(3.3). Now *assuming* that we can substitute ψ_{i} in Eq.(3.7) by ϕ_{i} and re-interpretation of $(\{\phi_{i}\}, \{\mathbf{R}_{\mu}\})$ as independent variables, we will have:

$$L(\mathbf{R}, \dot{\mathbf{R}}, \{\phi_{i}\}, \{\dot{\phi}_{i}\}, t) = \frac{1}{2} \mu \sum_{i}^{occ.} \int d\mathbf{r} |\dot{\phi}_{i}|^{2} + \frac{1}{2} M \dot{R}^{2} - U[R; \{\phi_{i}\}] \tag{8.26}$$

with

$$\mu = \frac{\hbar^{2}}{M\dot{R}^{2}} \tag{8.27}$$

having the dimension of $[energy][time]^{2}$. The Euler equation derived from this Lagrangian is:

$$\frac{1}{2} \mu \ddot{\phi}_{i} = -\frac{\delta U}{\delta \phi_{i}^{*}} + \sum_{j} \Lambda_{ij} \phi_{j} + \frac{\mu \ddot{R}}{\dot{R}} \dot{\phi}_{i} \tag{8.28}$$

$$\ddot{R} = -\frac{\delta U}{\delta R} + \frac{2\mu}{\dot{R}} \sum_i \int dr \dot{\phi}_i \ddot{\phi}_i - \frac{3\mu \ddot{R}}{\dot{R}^2} \sum_i \int dr |\dot{\phi}_i|^2 \tag{8.29}$$

where Λ_{ij} is the Lagrangian multiplier from orthonormalization of orbitals $\{\phi_i\}$. As we can see, there are coupling terms in Eq.(3.10) between $\dot{R}$ and $\dot{\phi}_i$ that will serve to pump the orbital kinetic energy back to ionic systems.

It'll be interesting to see what dynamics will result from Eqs.(3.9) and (3.10), for example, in diatomic molecules when the electronic states remain in $^1\Sigma$ state. One obvious feature is that this dynamics is reversible. Another way of making a reversible dynamics is to introduce an additional dynamical variable, serving as a thermostat to the original system. The dynamics of the whole system, then, is reversible. This is Blöchl and Parrinello approach. If we make the identification:

$$x_e = ln\left(\frac{1}{\dot{R}^2}\right) \tag{8.30}$$

where x_e is the dynamic variable of thermostat, and make μ a constant, we get the friction term in equation of motion for the orbitals in the Blöchl and Parrinello method. [3, 5]

For general cases of more than one-dimensional, upon assuming that (a)the change of orbitals due to different R_μs are orthogonal and (b)we can applying equipartition principle for the kinetic energy of ions:

$$\frac{1}{2} M_\mu \dot{R}_{\mu l}^2 = \frac{1}{2} k_B T \tag{8.31}$$

the orbital kinetic energy is

$$\frac{\hbar^2}{2k_B T} \sum_i^{occ.} \int dr |\dot{\phi}_i|^2. \tag{8.32}$$

This is just the one that Blöchl and Parrinello used for their adiabatic orbital kinetic energy.

8.4 Solving Kohn-Sham Equations for Atoms

Poisson Equation Solver

Using the logarithmic grid,

$$r = e^t - \delta \tag{8.33}$$

with t being a linear grid and δ a tiny offset, and using the variable $U(r) \equiv rV(r)$, the radial Poisson equation becomes the following,

$$u"(t) - u'(t) - l(l+1)(1+\delta/r)^2 u = -\frac{\rho_{1d}(r(t))(r+\delta)^2}{r(t)}, \tag{8.34}$$

where the one-dimensional density $\rho_{1d}(r) = \int d\Omega |rR_{nl}(r)|^2$. Using 7-point derivative for the second derivative leads to a banded matrix. This banded matrix can easily be solved by calling LINPACK routines.

Roothaan Equation Solver

(0) Get input density, e.g., from grid calculation.

(1) Form the matrix element;

(a)

$$\langle T \rangle = \langle \chi_\mu | -\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} | \chi_\nu \rangle$$

(b) form the nuclear field: $V_{nucl}(r)$

(c) form the Coulomb field: $V_{coul}(r)$

(d) form the exchange field: $V_{xc}(r)$

(2) Symmetrize the matrix;

(3) Call LINPACK routines to diagonalize the banded matrix.

(4) Form the density ρ on χ_μ from $C_{\mu n}$.

(5) Go to (1). SCF loop until converge.

8.5 Matrix Element Evaluation of Nonlocal ECP

Indexing the Basis

Table I. Basis for Fock operator.

$j(< nsh)$	$l(j)$	$lm(j)$	$i(< nfb)$	α_i	$a_i{}^a$	i_s(j)
1	0	S	1	.725000	.559968	1
2	0	S	2	.111200	.137243	2
3	0	S	3	.040400	.064224	3
4	1	X	4	1.240000	-1.865165	4
		Y	5			
		Z	6			
5	1	X	7	.134600	.116211	7
		Y	8			
		Z	9			
6	1	X	10	.042200	.027263	10
		Y	11			
		Z	12			

[a] Contraction coefficient without normalization factor included.

i_s is the starting basis(lm) index for primative Gaussian j.

Table II. All the m components of lm for a contracted Gaussian is called a shell.

n1	l	α	$a_{n\mu}$	n2	n3	n4	n5
1	S	0.7250	-0.206460				
2	S	0.1112	0.594623	1	2	1	0
3	S	0.04042	0.530827	2	1	3	0
4	P	1.2400	-0.036435				
5	P	0.1346	0.494619	3	2	4	1
6	P	0.04225	0.604568	4	1	6	1

n1: ipri(<nprimx), index of primative Gaussians;

n2: icontr(<nconmx), index of contracted Gaussians;

n3: ncontr(icontr), number of primative Gaussians in the i^{th} contracted Gaussian;

n4: kstart(icontr)=kstart(icontr-1)+ncontr(icontr-1), the starting primative index of the i^{th} contracted Gaussian;

n5: $l(kstart(icontr))$, angular momentum of the i^{th} contracted Gaussian.

Table III. Upon diagonalization of the Fock matrix, the canonical orbitals are ordered according to the eigenvalues ϵ_i $(eval(n))$.

n	1	2	3	4
eval(n)	-.160346	-.058169	.076829	.140090
icontr				
1	.437177	.000000	1.813551	.000000
2	.603094	.000000	-1.765324	.000000
3	.000000	.391922	.000000	-1.292476
4	.000000	.693541	.000000	1.158921

Direct Method

A. Identities

$$|\phi_n\rangle = \sum_{\mu} C_{\mu n} |\chi_\mu\rangle \tag{8.35}$$

where $|\phi_n\rangle$ is the orthonormalized eigenvector of an operator $\mathbf{O}$, satisfying $\langle \phi_{n'}|\phi_n\rangle = \delta_{n'n}$; $\{|\chi_\mu\rangle\}$ is basis set with $\langle \chi_\mu|\chi_\nu\rangle = S_{\mu\nu}$. Orthonormalization of $|\phi_n\rangle$ leads to

$$\langle \phi_{n'}|\phi_n\rangle = \sum_{\mu\nu} C_{n'\mu} C_{n\nu} S_{\mu\nu} = \sum_{\mu\nu} C_{n'\mu} S_{\mu\nu} C^{\dagger}_{\nu n} = \delta_{n'n}, \tag{8.36}$$

namely,

$$\mathbf{C}^{\dagger}\mathbf{S}\mathbf{C} = \mathbf{I}. \tag{8.37}$$

Using $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$, we have

$$\mathbf{C}\mathbf{C}^{\dagger} = \mathbf{S}^{-1}. \tag{8.38}$$

In matrix element form, it is

$$\sum_{n} C_{\mu n} C_{\nu n} = S^{-1}_{\mu\nu}. \tag{8.39}$$

B. Some Formuli

Equations for eigenfunctions:

$$\mathbf{O}|\phi_n\rangle = \lambda_n |\phi_n\rangle \tag{8.40}$$

on a finite basis $\{\chi_\mu\}$ becomes the Roothaan type of equation (Secular equation for non-orthogonal basis)

$$\mathbf{OC} = \mathbf{SC}\lambda. \tag{8.41}$$

In calculating the matrix element $\langle\chi_\mu|O|\chi_\nu\rangle$ there are two equivalent ways:

$$\langle\chi_\mu|V|\chi_\nu\rangle = \sum_{iji'j'} \langle\chi_\mu|i'\rangle S^{-1}_{i'i}\langle i|V|j\rangle S^{-1}_{jj'}\langle j'|\chi_\nu\rangle = \sum_{i'j'}\langle\chi_\mu|i'\rangle \tilde{V}_{i'j'}\langle j'|\chi_\nu\rangle \tag{8.42}$$

where

$$\tilde{V}_{i'j'} = \sum_{ij} S^{-1}_{i'i}\langle i|V|j\rangle S^{-1}_{jj'}; \tag{8.43}$$

or

$$\langle\chi_\mu|V|\chi_\nu\rangle = \langle\chi_\mu|\sum_n \lambda_n|\phi_n\rangle\langle\phi_n||\chi_\nu\rangle = \sum_{ij}\langle\chi_\mu|i\rangle D_{ij}\langle j|\chi_\nu\rangle \tag{8.44}$$

where

$$D_{ij} = \sum_n \lambda_n C_{in} C_{jn} \tag{8.45}$$

and

$$|\phi_n\rangle = \sum_i C_{in}|\chi_i\rangle \tag{8.46}$$

is the eigenvector of V. Diagonalization before summation cuts the cost by having one loop instead of two.

VW Method

$$\langle\chi^B_\mu|V^A_{NL}|\chi^C_\nu\rangle = \langle\chi^B_\mu|\sum_{lmn}|U_l R_{nl} lm\rangle_A \frac{1}{\lambda_{nl}}\ {}_A\langle U_l R_{nl} lm|\chi^C_\nu\rangle \tag{8.47}$$

where

$$R_{nl}(r) = \sum_{i=1}^{N_l} C_{nil} \chi_i^l(r). \tag{8.48}$$

Define

$$D_{iljl} = \sum_{n}^{N_l} \frac{1}{\lambda_{nl}} C_{nil} C_{njl} \tag{8.49}$$

$$= \sum_{l}^{N_l} \sum_{ij} D_{iljl} \Big[\sum_{m}^{M_l} \langle \chi_\mu^B | U_l^A | \chi_i^l lm \rangle^A \; {}^A\langle \chi_j^l lm | U_l^A | \chi_\nu^C \rangle \Big] \tag{8.50}$$

$$= \sum_{ij}^{N} \delta_{l(i)l} \delta_{l(j)l} D_{ij} \Big[\sum_{m}^{M_l} \langle \chi_\mu^B | U_l^A | \chi_i lm \rangle^A \; {}^A\langle \chi_j lm | U_l^A | \chi_\nu^C \rangle \Big] \tag{8.51}$$

where

$$D_{ij} = \sum_{n}^{N} \delta_{l(n)l(i)} \frac{1}{\lambda_n} C_{ni} C_{nj}. \tag{8.52}$$

To evaluate the angular momentum inner products we first transform the spherical Gaussians into Cartesian ones, $|\alpha\beta\rangle$, by using

$$|lm\rangle = \sum_{\alpha\beta} \mathcal{D}_{\alpha\beta,lm} |\alpha\beta\rangle. \tag{8.53}$$

Then,

$$\langle \chi_\mu^B | U_l^A | \chi_i lm \rangle^A = \sum_{\alpha\beta} \mathcal{D}_{\alpha\beta,lm} \langle \chi_\mu^B | U_l^A | \chi_i \alpha\beta \rangle^A. \tag{8.54}$$

Now

$$\langle \chi_\mu^B | U_l^A | \chi_i \alpha\beta \rangle^A \tag{8.55}$$

is a two-center integral.

Matrix Elements

Much of our derivation here is based on Melius. [6]

A. Matrix Elements of Local Potentials

Consider three atoms in space, A, B and C. The Gaussian located at center B would be

$$G_B(r_{B_i} r_{B_j} ...) = N_\mu r_{B_i} r_{B_j} ... e^{-\alpha_B r_B^2}. \tag{8.56}$$

Suppose the ECP is at A, we want to evaluate the matrix element

$$\langle G_B(r_{B_i} r_{B_j} ...) | U_{LMAX}(r_A) | G_C(r_{C_k} r_{C_l} ...) \rangle. \tag{8.57}$$

Projecting the B and C onto A

$$\mathbf{r}_B = \mathbf{r}_A - \mathbf{R}_{BA} \tag{8.58}$$

$$r_{B_i} = r_{A_i} - \overline{BA_i} \tag{8.59}$$

$$r_B^2 = r_A^2 + \overline{BA}^2 - 2\overline{BA} r_A \cos\theta_{A'} \tag{8.60}$$

where $\theta_{A'}$ is the angle between $\mathbf{r}_A$ and $\mathbf{R}_{BA}$.

Gaussian at B and at C can be combined into another Gaussian at D

$$e^{-\alpha_B(\mathbf{r}-\mathbf{R}_B)^2} e^{-\alpha_C(\mathbf{r}-\mathbf{R}_C)^2} = e^{\frac{-\alpha_C \alpha_B}{\alpha_C + \alpha_B}} e^{-\alpha_D(\mathbf{r}-\mathbf{R}_D)^2} \tag{8.61}$$

where

$$\alpha_D = \alpha_B + \alpha_C \tag{8.62}$$

and

$$\mathbf{D} = \frac{\alpha_C \mathbf{C} + \alpha_B \mathbf{B}}{\alpha_C + \alpha_B}. \tag{8.63}$$

Then (8.57) becomes

$$\mathcal{A}\langle U_{LMAX}(r_A)|G_D(r_{B_i}r_{B_j}...r_{C_k}r_{C_l}...)\rangle \tag{8.64}$$

$$\mathcal{A} = e^{\frac{-\alpha_C\alpha_B}{\alpha_C+\alpha_B}}. \tag{8.65}$$

Since $U_{LMAX}(r_A)$ is spherically symmetric, only that part of G_D which is spherically symmetric about A will contribute to the integral. Thus, (8.64) can be written as

$$\langle U_{LMAX}|s\rangle\langle s|G_D\rangle. \tag{8.66}$$

This is

$$\sqrt{4\pi}\int_0^\infty r_A^2 dr_A U_{LMAX}(r_A) I_{D,0,0}(r_A) \tag{8.67}$$

where we have used the definition

$$I_{Dlm}(r_A) = \int_{\Omega_A} Z_{lm}(\Omega_A) G_D(r_{B_i}, r_{C_j}, ...) d\Omega_A. \tag{8.68}$$

Thus, (8.57) becomes as follows,

$$\langle G_B(r_{B_i}r_{B_j}...)|U_{LMAX}(r_A)|G_C(r_{C_k}r_{C_l}...)\rangle \tag{8.69}$$

$$= \mathcal{A}\int_0^\infty r_A^2 dr_A \sqrt{4\pi}\, U_{LMAX}(r_A) N_D(r_A) \int_{\Omega_A} Z_{lm}(\Omega_A) r_{B_i} r_{C_j} ... e^{-2\alpha_D R_{DA} r_A \cos\theta_{A'}} d\Omega_A, \tag{8.70}$$

where

$$N_D(r_A) = e^{-\alpha_D r_A^2} e^{-\alpha_D R_{DA}^2}. \tag{8.71}$$

Using $N_B(r_A)N_C(r_A) = \mathcal{A}N_D(r_A)$ leads to

$$\int_0^\infty r_A^2 dr_A \sqrt{4\pi}\, U_{LMAX}(r_A) N_B(r_A) N_C(r_A) \int_{\Omega_A} Z_{lm}(\Omega_A) r_{B_i} r_{C_j} ... e^{-2\alpha_D R_{DA} r_A \cos\theta_{A'}} d\Omega_A. \tag{8.72}$$

Using

$$e^{-2\alpha_D \overline{DA} r_A \cos\theta_{A'}} = \sum_{l=0}^{\infty} (2l+1) M_l(2\alpha_D \overline{DA} r_A) P_l(\cos\theta_{A'}) \tag{8.73}$$

where $M_l(x) = i^l j_l(-ix)$ is the modified spherical Bessel function of the first-kind, after some algebraic manipulations, [7] we have

$$\int_{\Omega_A} Z_{lm}(\Omega_A) r_{B_i} r_{C_j} ... e^{-2\alpha_D R_{DA} r_A \cos\theta_{A'}} d\Omega_A = \sqrt{4\pi}(2l+1) \sum_{n,l'} C_{nl'} r_A^n M_{l'}(2\alpha_D R_{DA} r_A) \tag{8.74}$$

where $C_{nl'}$ are constants. We finally have

$$\langle G_B(r_{B_i} r_{B_j} ...) | U_{LMAX}(r_A) | G_C(r_{C_k} r_{C_l} ...)\rangle \tag{8.75}$$

$$= 4\pi \sum_{nl'} C_{nl'} \int_0^\infty r_A^2 dr_A U_{LMAX}(r_A) N_B(r_A) N_C(r_A) r_A^n M_l'(2\alpha_D R_{DA} r_A). \tag{8.76}$$

B. Two-Center Integrals

The two-center integrals of a nonlocal potential can be converted to and evaluated as matrix-elements of a local potential V_{LMAX}.

We want to calculate

$$\langle G_A(r_{A_i}, r_{A_j}, ...)|V^A|G_B(r_{B_i}, r_{B_j}, ...)\rangle. \tag{8.77}$$

with

$$\mathbf{V}^A = \sum_{l=0}^{\infty} U_l(r) \sum_{m=-l}^{l} |lm\rangle\langle lm|. \tag{8.78}$$

Since G_A and V^A are at the *same* center, only the angular momentum components contained in G_A survive.

For s or p, we have:

$$\langle G_A(s)|U_s|G_B(r_{B_i}, r_{B_j}, ...)\rangle, \tag{8.79}$$

or

$$\langle G_A(r_{A_i})|U_p|G_B(r_{B_i}, r_{B_j}, ...)\rangle. \tag{8.80}$$

These are just matrix elements of local potentials, see (8.76). For Cartesian Gaussians, the d-function have both the s-component and d-component. We have

$$\langle G_A(r_{A_i} r_{A_j})|U_{LMAX}|G_B(r_{B_i}, r_{B_j}, ...)\rangle + \langle G_A(r_{A_i} r_{A_j})| \sum_{l=0}^{LMAX-1} \Delta U_l |lm\rangle\langle lm|G_B(r_{B_i}, r_{B_j}, ...)\rangle$$

$$= \langle G_A(r_{A_i} r_{A_j})|U_d|G_B(r_{B_i}, r_{B_j}, ...)\rangle + \sqrt{4\pi}\frac{\delta_{ij}}{3}\langle G_A(s) r^2|\Delta U_s|G_B\rangle, \tag{8.81}$$

both terms are of the type of (8.76).

Bibliography

[1] R. G. Parr and W. Yang *Density-Functional Theory of Atoms and Molecules* Oxford. (1989)

[2] I. M. Gelfand and S. V. Fomin *Calculus of Variations.* Englewood Cliffs: Prentice-Hall. p42(1963)

[3] P.Blöchl and M. Parrinello, Phys. Rev. B **45** 9413, (1992)

[4] M. Parrinello and A. Rahman, J. Appl. Phys.**52** 7182 (1981)

[5] S. Nosé, J. Chem. Phys. **81** 511 (1984)

[6] C. Melius, *Ph.D. Thesis*, California Institute of Technology. (1973)

[7] For d-function $l = 2$,

$$Z_{lm}(\Omega_A) = \frac{r_{A_i} r_{A_j}}{r_A^2},$$

(a) projected on center A:

$$r_{B_i} = r_{A_i} - R_{A_i};$$

(b) rotate to the primed coordinate:

$$\mathbf{r} = \mathbf{A}\mathbf{r}';$$

(c) carry out the angular integration:

$$\int d\cos\theta_{A'} d\phi_{A'} \frac{r_{A_i}}{r_A} \frac{r_{A_j}}{r_A} P_l(\cos\theta_{A'})$$

where r_{A_i}/r_A is in terms of $\cos\theta_{A'}$, $\sin\theta_{A'}$, $\cos\phi_{A'}$ and $\sin\phi_{A'}$.