This thesis presents three different investigations of materials systems, with possible applications in the area of molecular electronics. Specifically, the binding of glycine, the simplest amino acid, to diamond surfaces, and the quantum transport characteristics of two prototypical metal/semiconductor interfaces are investigated.

With density functional theory based simulations, we have investigated the binding of the amino acid glycine on two of the most prominent diamond surfaces – *i.e.*, C(100) and C(111) (2 × 1) – with a focus on the associated energetics, charge transfer, electronic, and structural characteristics. With regards to the dimerized C(100) surface, interaction is mostly via the amide group of the glycine molecule (both with and without H-atom abstraction) and via a cycloaddition reaction whose activation barrier has been estimated via quantum chemistry methods. In contrast, the C(111) (2 × 1) surface was found to be mostly inert with respect to interactions with the glycine molecule.

Second part of this thesis presents theoretical investigations of electronic transport devices at atomic scale. One such device is a capacitance made of atomic wires, for which we present the results of *ab initio* investigation of the capacitance of Al nanowires. The systems considered include cross sectional areas for the wires: Al(100)(3 × 3), Al(100)(5 × 5), and Al(100)(7 × 7). First principles estimates of capacitance matrix coefficients for the systems are provided.

In the second part of this thesis, we have characterized the fully self-consistent electronic properties of a prototypical metal/nanotube interface using a combined nonequilibrium Greens function and density functional theory based formalism, un-
der different conditions of gate and bias voltages. Both carbon and boron nitride nanotubes between Al electrodes, were considered. The electronic properties of the interface are dominated both by a dipole and by metal induced gap states (MIGS) formed through the transfer of charge between the metal and the nanotube. In addition, first principles estimates – within the local density approximation – of the Schottky barrier heights are given.
Theoretical Investigations for Molecular Electronic Devices: Metal/Semiconductor Interfaces, Capacitance of atomic scale wires, and Organics on Diamond Surfaces

by

Khorgolkhui Odbadrakh

A dissertation submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the Degree of Doctor of Philosophy

Physics

Raleigh, NC

2007

Approved By:

Prof. Christopher Roland
Chair of Advisory Committee

Prof. Lubos Mitas

Prof. Donald W. Brenner

Assoc. Prof. Celeste Sagui
Dedication

To my Mom, Dad and Sister
Biography

I was born in Ulaanbaatar, a capital city of Mongolia, right at the heart of Asia.

My childhood went in a typical way as for many children in the country. My parents were very enlightened people in their times, exposing me to science and technology from my early ages. I still vividly remember my 8th birthday present from my Dad: a small booklet about Freedman’s theory on how life came to existence, in mongolian. Later, meeting my senior high physics teacher was probably the turning point of my life, which brought me where I am today.

After the high school, I started my physics studies at the National University of Mongolia, the only university in the nation which offers science degrees. This university offered us the best the nation could afford, lecturers who obtained their degrees in the former Soviet Union and Eastern Europe.

By the time I was preparing my thesis defence, we were also busy participating in a peaceful democratic revolution, which eventually resulted in a downfall of seventy years of communist rule in the country. So, I was celebrating my graduation in a newborn and only democracy in that part of the world. But we faced brunt of the transition to a free society, though we were one of the active forces behind this revolution. All the contracts that trained young mongolian scientists in former socialist countries have ceased and we have been left with no clue what to do next. So, my long journey started. My first step was to study in Diploma Programme at the Abdus Salam Center for Theoretical Physics, Trieste, Italy. Then I received my Masters degree at the University of New South Wales, Sydney, Australia. Finally I have been accepted to complete my Ph. D. program at North Carolina State University, Raleigh. In between, I had to teach and work for the National University of Mongolia for many years.
Acknowledgements

My first and foremost thanks goes to Professor Christopher Roland and Celeste Sagui for their support and guidance through my studies. Thanks to their support, I completed this program, gaining valuable experience, was able to travel to APS March meetings, and the University of Montreal, and publish my research works. This thanks also extends to Dr. Pawel Pomorski, who taught me how to perform quantum transport simulations. During first part of my project, I enjoyed his generous commitments in overcoming numerous obstacles. During the final leg of my program, I enjoyed close cooperation and help from Dr. Xuan Luo and Jung Goo Lee. So, I extend my gratitude to them as well. My thanks also rightly goes to the members of this group, specially our secretary C. Upchurch, for their support and kindness.

My sincere thanks to Professor M. Paesler and C. Gould, who not only accepted me for this program, also offered their personal support during a difficult period in my life. I faced no difficulties during this program thanks to highly professional commitments from all faculty and staff members of the Department of Physics, especially our devoted secretary Jenny Allen.

My considerably long journey towards this degree was closely supported and watched by those who influenced me in so many positive ways. To my great relief, the time has come to thank Prof. G. Ochirbat, Ts. Gantsog, Kh. Tsookhuu, S. Davaa, Z. Bold, D. Sangaa and Doctor M. Ganbat, all from the National University of Mongolia. Academician Kh. Namsrai and Doctor Z. Omboo from the Academy of Sciences, Mongolia. Professor Gary P. Morriss from the University of New South Wales, Australia and Professor Hilda A. Cerdeira, Abdus Salam Center for Theoretical Physics, Italy.

I would like to especially thank to my high school physics teacher Mrs. B. Tseesuren, for that amazing talk long time ago, in that remote Mongolian classroom. Nobody could have imagined that someone’s life is being profoundly changed during that ordinary physics class.
Many thanks for their personal support and courtship from my friends. Dr. A. Undraa and D. Dashdorj, long time friends, first suggested doing my Ph. D. program here at NCSU. My graduate study years have passed with warm and happy memories thanks to friendship from Eric Adles and his wife Sasha Woods, Matt Walker, Dr. Matt Highland, Joshua Smith and all the fellow graduate students.

My parents are the ones who waited for this moment the longest. Thanks for your love and life long encouragements, thanks for those amazing little stories, and thanks for your patience. Now in their seventies, they are still going strong, asking me if I am publishing any more scientific papers.

My deepest thanks goes to my family for their love and support. My wife D. Otgonchimeg, my sons O. Tuguldur and O. Batu, were the source of my hope and strength through all these years and will remain the same for the rest of my life. Thank you for your love and patience.

# Contents

List of Figures viii

List of Tables x

1 Introduction 1

2 Density Functional Theory 5

2.1 Introduction 5

2.2 The Hohenberg-Kohn formulation of Density Functional Theory 6

2.3 The Self-Consistent Kohn-Sham equations 8

2.4 Local Density Approximation and Beyond 11

2.5 Pseudo-potentials 12

2.5.1 Non-linear Core Corrections 15

2.6 Basis Set 16

2.6.1 Fireball Atomic Orbital Basis 17

3 Theory of Quantum Transport 20

3.1 Landauer-Büttiker Formulation 21

3.2 DFT-NEGF method for Transport problems 24

3.3 Non-equilibrium Green’s function approach to electronic states 27

3.4 Leads as Periodic Systems 28

3.5 Central region of two probe device 30

3.6 Integrating the Green’s Function 33

3.7 Analysis of Real and Complex band structures 34

3.8 Bound States 35

3.9 Computational Scheme 37

4 Theoretical Investigation of the interaction of Glycine with the diamond C(100) and C(111) (2 × 1) surfaces 39

4.1 Abstract 39
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.2</td>
<td>Introduction</td>
<td>40</td>
</tr>
<tr>
<td>4.3</td>
<td>Methodology</td>
<td>42</td>
</tr>
<tr>
<td>4.4</td>
<td>Results and Discussion</td>
<td>44</td>
</tr>
<tr>
<td>4.4.1</td>
<td>Glycine on C(100) (2x1) surface</td>
<td>45</td>
</tr>
<tr>
<td>4.4.2</td>
<td>Glycine on C(111) (2x1) surface</td>
<td>54</td>
</tr>
<tr>
<td>4.5</td>
<td>Conclusions and Summary</td>
<td>58</td>
</tr>
<tr>
<td>5</td>
<td>Capacitance of Al nanowire</td>
<td>60</td>
</tr>
<tr>
<td>5.1</td>
<td>Abstract</td>
<td>60</td>
</tr>
<tr>
<td>5.2</td>
<td>Introduction and Methodology</td>
<td>60</td>
</tr>
<tr>
<td>5.3</td>
<td>Results and Discussion</td>
<td>63</td>
</tr>
<tr>
<td>5.4</td>
<td>Summary</td>
<td>67</td>
</tr>
<tr>
<td>6</td>
<td>Ab initio band bending, metal-induced gap states, and Schottky barriers of a carbon and boron nitride nanotube devices</td>
<td>68</td>
</tr>
<tr>
<td>6.1</td>
<td>Abstract</td>
<td>68</td>
</tr>
<tr>
<td>6.2</td>
<td>Introduction</td>
<td>69</td>
</tr>
<tr>
<td>6.3</td>
<td>Methodology</td>
<td>70</td>
</tr>
<tr>
<td>6.4</td>
<td>Results and discussion</td>
<td>72</td>
</tr>
<tr>
<td>6.5</td>
<td>Summary</td>
<td>78</td>
</tr>
<tr>
<td>7</td>
<td>Summary and future work</td>
<td>80</td>
</tr>
<tr>
<td>Bibliography</td>
<td></td>
<td>84</td>
</tr>
</tbody>
</table>
List of Figures

1.1 Projected growth of Atomic scale Modeling. Credited to Atomistix. 

3.1 A two-probe system is divided into three regions. 

3.2 Contour integral path in complex plane used in integration of Eq.(3.45). 

4.1 Hydrogen terminated clusters used for the quantum chemistry calculations of the activation barriers: (a) $C_{21}H_{20}$ for C(100); (b) $C_{35}H_{30}$ for C(111), respectively. 

4.2 Side and top views of the pristine diamond C(100) $(2 \times 1)$ surface showing HOMO (red) and LUMO (blue) states at the $\Gamma$ point, as well as corresponding $(2x4)$ surface band structure and simulated STM images. Because there is no buckling, the HOMO/LUMO states are distributed uniformly around the surface dimers. 

4.3 From top to bottom, we show configurations with HOMO/LUMO, charge transfer, $(2x4)$ surface band structure, and simulated STM images for intact glycine on C(100) $(2 \times 1)$ surface: (a) amide i.e., N(H$_2$) adsorption; (b) N+O absorption; and (c) C=O cycloaddition reaction. Color scheme: HOMO(LUMO) indicated by red(blue) as in Fig.4.1; yellow (violet) for regions of negative (positive) $\delta\rho$ in charge transfer pictures. 

4.4 Plot of reaction energies for C=O cycloaddition reaction on C(100) $(2 \times 1)$ as obtained with quantum chemistry calculations. The most relevant parts of the cluster are shown for the initial, transition, and final states. Energies are measured with respect to the total energy of the isolated surface and molecule. 

4.5 From left-to-right: side view of configurations with HOMO/LUMO, charge transfer, structural details focusing on bond lengths, $(2x4)$ surface band structure, and simulated STM images for situations where glycine absorption on C(100) $(2 \times 1)$ involves H-atom abstraction: (a) H from N(H$_2$); (b) H from (O-H)$_1$; (c) H from (O-H)$_2$; and (d) H from O+O(H). Color scheme same as in Fig.4.3.
4.6 Side and top views of the pristine C(111) (2×1) surface showing HOMO (red) and LUMO (blue) states at the Γ point, as well as the corresponding (2x1) surface band structure and STM images. .......................... 56

4.7 From left-to-right: side view of configurations with HOMO/LUMO, charge transfer, structural details, and simulated STM images for situations where glycine adsorption on C(111) (2 × 1) involves either (a) cycloaddition with C=O; or H-atom abstraction with (b) H from N(H₂); and (c) H from O(H). Color scheme is same as in Fig.4.3. .... 57

5.1 Schematic display of Al nanowire as a capacitor. ......................... 62

5.2 Charges on Al nanowire capacitors at different bias voltages. .......... 64

5.3 Cross capacitance coefficient C₁₂ as function of separation distance between the capacitors. .......................... 65

5.4 Percent induced charge \( \frac{\Delta Q_{\text{left}}}{\Delta Q_{\text{right}}} \) ........................................ 66

6.1 Typical two-probe (8,0) CNT device (length N+2 units) sandwiched between Al (100) leads (L, R). Note that the central region C also contains a considerable portion of the electrodes. The whole system is surrounded by a series of gates (shown here as one), which set \( V_G \) in different parts of the device. ......................... 71

6.2 (a) Sample I-V curves for CNTs with \( V_G = 0.0 \). Here circles, squares and triangles represent data for tubes of length 3, 8 and 18 units, respectively; (b),(c) and (d) show corresponding \( T(E) \), with evanescent mode contributions shaded. Here (d) also shows \( T(E) \) for an infinite (8,0) CNT (step-like function); (e) complex bandstructure of infinite (8,0) CNT; and (f) (8,0) BNNT. .......................... 74

6.3 LDOS of different parts of the CNT (left panels) and BNNT (right panels) devices (\( V_G = 0 \)). The estimated gap states are shaded and the Fermi energy \( E_F \) of the leads are marked with a vertical line. Data shown is for nanotubes consisting of 18 unit cells. ................. 75

6.4 Ab initio band bending and corresponding excess charge per atom for each of the CNT cells. Here, circles, squares and triangles represent data for tubes with 6, 12, and 18 cells, respectively. For top three panels \( V_b = 0 \); for bottom two, \( V_G = 0 \) ................. 76

6.5 (a) \( \phi_b \) as a function of \( \Delta E_F \) for N=12 CNT devices; (b) Excess charge per atom across the metal/CNT junction (marked by the vertical line) showing the interface dipole. .......................... 78
List of Tables

4.1 Absorption energy (eV) and charge transfer characteristics for the chemisorption of glycine on the C(100) (2 × 1) surface. While the top three items cover adsorption of an intact glycine molecule, the rest involve hydrogen abstraction. .................................................. 48

4.2 Activation energies (eV) for the reactions of glycine on the C(100) (2×1), C(111) (2×1), and Si(100) (2×1) surfaces. Data for the latter is from Ref.[48]. ................................................................. 52

4.3 Adsorption energies (eV) and charge transfer characteristics for glycine on C(111) (2 × 1) surface. Except for the cycloaddition reaction, they all involve abstraction of a H-atom from the glycine molecule. ......... 55
Chapter 1

Introduction

The complexity for minimum component costs has increased at a rate of roughly a factor of two per year ... Certainly over the short term this rate can be expected to continue, if not to increase. Over the longer term, the rate of increase is a bit more uncertain, although there is no reason to believe it will not remain nearly constant for at least 10 years. That means by 1975, the number of components per integrated circuit for minimum cost will be 65,000. I believe that such a large circuit can be built on a single wafer.

So wrote Dr. Gordon Earle Moore in 1965 [1], which was later hailed as Moore’s Law by Carver Mead [2]. This empirical observation still holds true after more than five decades. Most desktop computer processors are fabricated with transistors of 65nm scale today. On January 27, 2007, Intel Co., co-founded by Dr. G.Moore, announced a chip integrates 45 nm transistors. By the year 2014, lateral size of active components in microelectronics is predicted to reach 35 nm [3]. This relentless miniaturization, driven by industry demands to market ever more processing power, will inevitably force the size of the active elements of the integrated chips into the molecular dimensions, by 2020 [4]. This means that much of the functionality of a device will be due to physical processes that occur within a few molecular layers of some interface, or an outer surface exposed to ambient. The utilization of such devices ultimately depends upon the precise understanding of underlying atomic level
processes. Nanoscale technology in one hand, and command over more processing power on the other hand make computer simulations of atomic scale devices a reality. As a result, computer simulations are becoming an indispensable tool for investigating novel molecular electronic devices. In fact, research and Development (R&D) expenditure for atomic scale modeling is expected to grow much faster than, and will equal R&D expenditure for experiment by 2035 (FIG. 1.1)[5]. In this thesis, we employed some of the well established simulation softwares such as VASP, Gaussian, and MDCAL.

It is obvious that manufacturing at the nanometer length scale facilitates not only new technologies, also use of new type of devices. One of the promising new trend in nanoscale technology is organic-semiconductor hybrid devices. The fact that over 95% of all known chemical compounds is organic is quite appealing. Combined with already established semiconductor technology, the hybrid devices promise wide range of applications with highly sensitive optical, electronic and mechanical properties, which can be used in physical, chemical and biological activities [6]. Surface functionalization or organic modification is the process of depositing layers or organic
molecules at semiconductor surface. For example, organic molecules can be deposited on a semiconductor substrate, then the organic layer may be terminated with a variety of end groups that respond to different chemical or biological stimuli. Binding of a specific molecule to the end group triggers a signal within the organic layer. If this signal can be coupled into the semiconductor substrate, then all of the capabilities of semiconductor based micro-electronics, such as signal amplification, processing and storage, become available. This could represent a semiconductor chip based chemical or biological sensor.

In this thesis, we explored possibilities of such system consisting of glycine molecule on diamond surfaces. It is the surface interactions when it comes to fabricating these type of devices. Diamond has unique physical and mechanical properties despite its identical crystal structure with common semiconductors silicon and germanium. Glycine, the simplest, yet featuring key elements of amino acids, is the logical starting point for systematic studies of organic molecules on semiconductor surfaces. With density functional theory based simulations, we have investigated in the first part of this thesis the binding of the amino acid glycine on two of the most prominent diamond surfaces – i.e., C(100) and C(111) (2 × 1) – with a focus on the associated energetics, charge transfer, electronic, and structural characteristics.

Second part of this thesis presents theoretical investigations of electronic transport devices at atomic scale. One such device is a capacitance of atomic scale components. Conductors, placed in close proximity to each other, and connected to separate electron reservoirs, could represent a charge storage device (i.e. capacitor) in classical notions. The interesting questions are, can the atomic scale conductors make an effective capacitor, and how different are their characteristics at this scale? In this thesis, we present fully self-consistent first principles investigation of the capacitance of atomic scale metal wires. The systems investigated include Al(100)(3 × 3), Al(100)(5 × 5), and Al(100)(7 × 7) leads, separated by a small distance comparable with few atomic layers.
The next system investigated is a two probe electronic transport devices based on semiconducting nanotubes. Carbon nanotubes (CNT) are very versatile, and stable material at the nanometer length scale, which could be ideal component in molecular electronic devices. In particular, semiconducting nanotubes show considerable promise as field-effect transistors. Fundamental issue here is to understand electronic structure at the interface between semiconducting CNTs and metal electrodes. The quasi-one-dimensional nature of the nanotubes implies reduced Coulombic screening, so that CNT-based devices are expected to be fundamentally different from conventional devices. In this thesis, we present fully self-consistent first principles investigation, within the context of the local density approximation, of the transport characteristics of a prototypical CNT and boron nitride nanotube (BNNT) based device with Al and metallic CNT electrodes.

This thesis is organized as follows. Chapter 2 is devoted to the introduction to DFT in general. In Chapter 3, we discuss the Non-equilibrium Green’s Function combined with DFT (NEGF-DFT) formalism that we used in quantum transport calculations. In Chapter 4, we present the theoretical investigation of glycine molecule on diamond surfaces. In Chapter 5, capacitance of Al nanowires is investigated. Detailed investigations of electronic structure of a metal/semiconductor interface in nanotube based devices is presented in Chapter 6. And the final chapter is devoted to summary and brief mention of future work.
Chapter 2

Density Functional Theory

2.1 Introduction

Marked by its relative simplicity and efficiency, the density functional theory (DFT) has established itself as one of the core engines for computer simulations of atomic scale systems and solid state physics at large. From a microscopic point of view, matter in the solid state consists of a system of nuclei and electrons interacting through the Coulomb forces. The nuclear degrees of freedom are separated from much lighter electronic ones through Born-Oppenheimer (adiabatic) approximation [7], in which the ions are regarded as fixed charges acting as potential sources. The system is now represented by gas of electrons moving in a static external potential created by the nucleus.

Traditional quantum mechanics treats the gas of interacting particles through the Schrödinger equations, of which the solutions are many-particle wave functions. However, this approach poses two very significant obstacles. Many body wave function, which represents the system and written as Slater determinant, have $N \sim 3 \times 10^{23}$ degrees of freedom, making any attempt to solve the Schrödinger equations for realistic system practically impossible. Also, computing time needed for calculations increase exponentially as the system size increases. Current state of this approach is such
that number of atoms in systems is in order of $\sim 10$. DFT is more approximate, but tractable. It focuses on electron density distribution in real three dimensional space $\rho(r)$, rather than the many body wave-functions. On the other hand, computing time $T$ increases relatively modestly with the number of atoms, currently as $T \sim N^\alpha$, where the $\alpha$ is around 2-3 [8]. Recent progress made in linear scaling is bringing this value down to $\alpha \approx 1$. Currently, DFT based computational methods can handle systems of the order of $N = 10^3$ atoms. More on that, the electronic density $\rho(r)$ is calculated by single particle wave functions in DFT, leaving all many body effects to the external potential in the single particle Schrödinger equations $V_{ext}$.

2.2 The Hohenberg-Kohn formulation of Density Functional Theory

DFT is formulated in terms of the Hohenberg-Kohn theorem:

*The ground state density $\rho(r)$ of a bound system of interacting electrons in some external potential $V_{ext}$ determines this potential uniquely up to an additive constant.*

This theorem reveals two important features in systems of $N$ quantum particles:

- There is one-to-one correspondence between ground state electron density $\rho(r)$ and ground state many body wave functions.

- The ground state electron density $\rho(r)$ minimizes the total electronic energy of the system.

The proof is straightforward. Let $\rho(r)$ be the non-degenerate ground state density of $N$ electrons in the potential $V_{ext}$, corresponding to the ground state $\Psi_1$, end the energy $E_1$. Then,

$$E^1 = (\Psi_1, H_1 \Psi_1) = \int V_{ext}(r)\rho(r)dr + (\Psi_1, (H + U)\Psi_1),$$

(2.1)
where the $H_1$ is the total Hamiltonian corresponding to $V_{1\text{ext}}$, and $T$ and $U$ are the kinetic and potential energy operators. Now assume that there exists a second potential $V_{2\text{ext}} \neq V_{1\text{ext}} + \text{Const.}$, with ground state $\Psi_2$, necessarily $\neq e^{i\theta}\Psi_1$, which also gives the same ground state density $\rho(r)$. Then

$$E_2 = \int V_{2\text{ext}}(r)\rho(r)dr + (\Psi_2,(H+U)\Psi_2). \tag{2.2}$$

Since $\Psi_1$ is non-degenerate, the Rayleigh-Ritz minimal principle for $\Psi_1$ requires

$$E_1 < (\Psi_2,H_1\Psi_2) = \int V_{1\text{ext}}(r)\rho(r)dr + (\Psi_2,(H+U)\Psi_2) = E_2 + \int [V_{1\text{ext}} - V_{2\text{ext}}] \rho(r)dr. \tag{2.3}$$

Similarly,

$$E_2 \leq (\Psi_1,H_2\Psi_1) = E_1 + \int [V_{2\text{ext}} - V_{1\text{ext}}] \rho(r)dr. \tag{2.4}$$

Adding Eqs. 2.3 and 2.4 leads to a contradiction:

$$E_1 + E_2 < E_1 + E_2. \tag{2.5}$$

Therefore, the existence of a second potential $V_{2\text{ext}}$, which is not equal to $V_{1\text{ext}} + \text{Const.}$ and gives the same ground state density $\rho(r)$ is excluded by reductio ad absurdum. Further, $\rho(r)$ determines both $N$ and $V_{\text{ext}}$, which in turn gives us the total Hamiltonian $H$. Therefore, the ground state electronic density distribution $\rho(r)$ determines implicitly all properties derived from the total Hamiltonian through solving the Schrödinger equations: many body eigenfunctions $\Psi^{(k)}(r_i)$, the two-particle Green’s function $G(r_1t_1, r_2t_2)$ etc. [8]. In other words, the ground state energy $E$, more precisely the external potential $V_{\text{ext}}$, is a functional of ground state density $\rho(r)$. Traditional quantum mechanical approaches find the ground state energy either by solving $3N$ Schrödinger equations or the Rayleigh-Ritz minimization, which use $3N$-dimensional normalized trial wave function $\Psi$. As you can see, the DFT
has transformed this formidable task into minimizing the ground state energy with respect to three-dimensional trial function $\rho(r)$.

The original HK theorem is formulated on non-degenerate ground states and can easily be extended to degenerate ground states [9]. At its original form, HK theorem is valid for non-interacting electrons. Further discussions will show that it actually reduces the fully interacting N-electron problem into a problem of finding single-particle ground state density.

This theorem tells only the existence of one-to-one correspondence between $\rho(r)$ and the many body wave function of $N$ electron system. Actual mapping of this correspondence is where the approximation are made.

2.3 The Self-Consistent Kohn-Sham equations

The concept of describing many electron system through a single-particle equations was first formulated by Hartree in 1928. In his formulation, each electron feels an effective potential due to both the nucleus and the other electrons:

$$V_H(r) = -\frac{Z}{r} + \int \frac{\rho(r')}{|r-r'|} dr',$$

where the first term is due to a nucleus of atomic number $Z$ and the second term represents the potential resulted from the average electronic density $\rho(r)$. Then the system can be described by a set of single-electron Schrödinger equations:

$$\left\{ -\frac{1}{2} \nabla^2 + V_H(r) \right\} \varphi_j(r) = \epsilon_j \varphi_j(r),$$

where $j$ denotes spatial or spin quantum numbers, and the mean density is given by

$$\rho(r) = \sum_{j=1}^{N} |\varphi_j(r)|^2.$$

Equations (2.6)-(2.8) constitute a self-consistent approximate description of many electron system and called Hartree equations. To solve this equation, one must start
from a trial $\rho(r)$, then constructs the $V_H(r)$ and then solves the Eq. (2.7) for $\varphi_j(r)$, and then finally recalculates $\rho(r)$ with Eq. (2.8). If this final density is different from the trial one, the cycle is repeated until the system converges to the same value. Now, let us explain this minimization problem from the standpoint of variational principle, generalizing the effective single particle potential $V_H$ into some external potential $V$. Then the variational principle takes form:

$$E[\tilde{\rho}(r)] = \int V(r)\tilde{\rho}(r)dr + T_S[\tilde{\rho}(r)] \geq E_{GS}. \quad (2.9)$$

This equation states that the energy functional $E[\tilde{\rho}(r)]$ is minimal at the ground state density. If $T_S[\tilde{\rho}(r)]$ represents the ground state kinetic energy for a non-interacting electron density, and introducing a Lagrange multiplier $\epsilon$ to ensure particle conservation, the Euler-Lagrange variational equation can be written as:

$$\delta E[\tilde{\rho}(r)] \equiv \int \delta \tilde{\rho}(r) \left\{ V(r) + \frac{\delta T_S[\tilde{\rho}(r)]}{\delta \tilde{\rho}(r)} \bigg|_{\tilde{\rho} = \rho} - \epsilon \right\} dr = 0, \quad (2.10)$$

where $\rho(r)$ is the exact single particle ground state density for the external potential $V(r)$. This stationary condition states that the ground state energy and density are may be obtained by solving this eigenvalue problem for eigenfunctions $\varphi_j(r)$, in:

$$\left(-\frac{1}{2} \nabla^2 + V(r) - \epsilon_j\right) \varphi_j(r) = 0, \quad (2.11)$$

which gives the total ground state energy and density distribution as:

$$E = \sum_{j=1}^{N} \epsilon_j; \quad \rho(r) = \sum_{j=1}^{N} |\varphi_j(r)|^2. \quad (2.12)$$

In his ground breaking work in 1964, Kohn transformed the form for single particle in Eq. (2.11) into a one for interacting particles [10]. To achieve this feat, lets have a closer look at Eq. (2.9). The first term is due to an external potential. The second term is kinetic energy functional of single particle density in Hartree formulation. But kinetic energy is the only energy in a system of non-interacting particles. In fact, the functional $T_S[\tilde{\rho}(r)]$ actually represents a universal energy functional intrinsic to the
system. Therefore, we can generalize it to a much more inclusive energy functional, which also takes into account all interactions between the particles:

$$ F[\tilde{\rho}(r)] \equiv T_S[\hat{\rho}(r)] + \frac{1}{2} \int \frac{\hat{\rho}(r)\hat{\rho}(r')}{|r-r'|} d\text{r}d\text{r}' + E_{XC}[\hat{\rho}(r)]. \quad (2.13) $$

This is still a functional of single-particle density $\hat{\rho}(r)$. However, a new term $E_{XC}[\hat{\rho}(r)]$ is introduced, which represents corrections to both kinetic and potential energies due to electron-electron interactions. By introducing this so called exchange-correlation term, Kohn-Sham formulation essentially approximates the total energy of interacting system as the kinetic energy of noninteracting electrons plus contributions due to interactions. Considering that kinetic energy constitutes a large portion of system’s total energy ($\sim 50\%$), this is a sound starting point, as the many-body corrections constitute only a small portion of total energy [11]. Using the new functional in Eq. (2.13), the variational principle takes the form:

$$ E[\tilde{\rho}(r)] \equiv \int V(r)\tilde{\rho}(r)dr + T_S[\hat{\rho}(r)] + \frac{1}{2} \int \frac{\hat{\rho}(r)\hat{\rho}(r')}{|r-r'|} d\text{r}d\text{r}' + E_{XC}[\hat{\rho}(r)] \geq E_{GS}. \quad (2.14) $$

The Eq. (2.10) can now be rewritten as:

$$ \delta E[\tilde{\rho}(r)] = \int \delta\tilde{\rho}(r) \left\{ V_{\text{eff}}(r) + \frac{\delta T_S[\hat{\rho}(r)]}{\delta \tilde{\rho}(r)} \bigg|_{\tilde{\rho}(r)=\rho(r)} - \epsilon \right\} dr = 0, \quad (2.15) $$

where

$$ V_{\text{eff}}(r) = V(r) + \int \frac{\rho(r')}{|r-r'|} d\text{r}' + V_{XC}(r). \quad (2.16) $$

and

$$ V_{XC}(r) \equiv \frac{\delta E_{XC}[\hat{\rho}(r)]}{\delta \hat{\rho}(r)} \bigg|_{\hat{\rho}(r)=\rho(r)}. \quad (2.17) $$

The form of Eq. (2.15) is identical with the Eq. (2.10) for non-interacting electrons, with the new potential $V_{\text{eff}}(r)$. Therefore, the ground state energy for interacting electron system is now given by solving the single-particle equations

$$ \left( -\frac{1}{2} \nabla^2 + V_{\text{eff}}(r) - \epsilon_j \right) \varphi_j(r) = 0. \quad (2.18) $$
The Equations (2.8), (2.17), and (2.18) constitute a set of self consistent equations for interacting electron system and are called the Kohn-Sham (KS) equations. The ground state energy for interacting system is given by:

$$E = \sum_j \epsilon_j + E_{XC}[\rho(r)] - \int V_{XC} \rho(r) dr - \frac{1}{2} \int \frac{\rho(r)\rho(r')}{|r-r'|} dr.$$  

(2.19)

The electronic density in the KS formulation is still that of a set of non-interacting particles. However, the fictitious single-particle potentials $V_{eff}(r)$ and $V_{XC}(r)$ can always be found, in principle, by equating the single-particle density with the physical density of a real system, if it is known. In practical terms, KS formulation effectively transforms the many-body problem for system of interacting electrons into finding sufficiently accurate, yet reasonably simple functional $E_{XC}[\rho(r)]$. In the next subsections, we will discuss most common approximations used in KS theory.

### 2.4 Local Density Approximation and Beyond

The most common approximation to $E_{XC}$ is the so called Local Density Approximation (LDA). The underlying principle here is to replace the exchange-correlation energy at particular location with the per particle exchange-correlation energy $e_{xc}(\rho(r))$ of a uniform electron gas and integrate over the system’s volume:

$$E_{LDA}^{XC} \equiv \int e_{xc}(\rho(r))\rho(r) dr.$$  

(2.20)

Thus, LDA approximates the characteristics of an inhomogeneous system with that of a more uniform electron gas. It is therefore expected to perform better for systems with slowly varying electron densities. For atomic systems, LDA gives ionization energies of atoms, dissociation energies of molecules, and cohesive energies with accuracy of around 10-20%, but bond lengths with a surprising accuracy of around 1% [8]. LDA typically breaks down when describing strongly correlated electron systems. Overall, experience has shown that LDA gives very useful results in most practical applications. Obviously, LDA has built in limitation in predicting measurable physical
properties of materials. Most notable examples are paramagnetic ground state structure of iron crystal, instead of ferromagnetic, and systematic errors in the dielectric constant in semiconductors by 10% or larger.

As the first step approximation, LDA attempts to reflect the local density of electrons. The next logical step is to feed the KS equation with information on how this density varies in space, that is the gradient of the local density. This gives rise to the so called Gradient Corrected (GC) exchange-correlation functionals. Early attempts have shown that the simple expansions of the gradient violated important sum rules and constraints [12]. Hence, a more generalized form "Generalized Gradient Approximation" (GGA) was later introduced by constructing a functional that satisfies these constraints. Further progress have been made by introducing numerical fitting parameters predetermined from optimization of accuracy in atomization energies of certain well known set of molecules [13].

2.5 Pseudo-potentials

Assuming that the exchange-correlation functional is separately dealt with, the KS equation now depends on how we define the external potential $V_{ext}$. The underlying concept for investigating solids is that the atomic nuclei are fixed in space (Born-Oppenheimer approximation) along with their core electrons, so that their overall effect can be replaced by a well behaved potential (pseudo-potential). This is well reasoned approximation considering the fact that in most solids, the atomic nucleus and their core electrons interact with each other weakly at best, which leaves the pseudo-potential to act on the valence electrons only. One major criteria for pseudo-potential is that it should be identical with all-electron potential outside certain core radius: $r > R_c$.

Benefits from this approximation are that the potentials are now nonsingular at the atomic cores, and that control over the pseudo-potential, so that it can be tailored
to satisfy many constraints and transferability conditions. The core electrons are no longer relevant, and so the eigenfunctions for pseudo-potentials do not have to be normal to the core electron wave functions. This translates into a practical benefit: the pseudo-orbitals can be represented on real space grid rather than atomic orbitals and are smoother.

However, we should keep in our mind that energetics of system is now changed with replacement of atomic nucleus and the core electrons with pseudo-potentials. That means only the change in total energy due to valence electrons is detectable. Again, this is a benefit rather than harm, as long as the ionic cores interact weakly with each other and are insensitive to their environment.

Pseudopotentials are constructed using an \textit{ab initio} procedure. The ‘true’ wavefunctions are calculated for an isolated atom using an all-electron DFT approach. The resulting valence wavefunctions are then modified in the core region to remove the oscillations, while obeying the norm-conservation constraint. The Schrödinger equation is then inverted to find the pseudopotential which will reproduce the pseudowavefunctions. This procedure produces a pseudopotential which may be transferred between widely varying atomic systems. This contrasts with semi-empirical potentials which are constructed to describe a particular atomic environment and may not be simply transferred to different environments.

A pseudo-potential is constructed such that it matches the true potential outside a given radius, designated the core radius. Similarly, each pseudowavefunction must match the corresponding true wave function beyond this distance. In addition, the charge densities obtained outside the core region must be identical to the true charge density. Thus, the integral of the squared amplitudes of the real and pseudowavefunctions over the core region must be identical. This condition is known as norm-conservation [14]. The atomic properties of the element must be preserved, including phase shifts on scattering across the core. These phase shifts will be different for different angular momentum states and so, in general, a pseudopotential
must be non-local, with projectors for different angular momentum components. The pseudopotential is often represented using the form [15]:

\[ V_i(r) = V_i^{\text{loc}}(r) + \sum_{l=0}^{l_{\text{max}}} V_{i,l}^{nl}(r) P_l, \]  

(2.21)

where the \( i \) is atomic index, and the \( P_l \) is a projector which projects the electronic wave functions onto the eigenfunctions of different angular momentum states. In general, great many projectors needed and ways to reduce this number have been explored. The evaluation of the non-local potential in reciprocal space requires a computational time which is proportional to the cube of the system size. The projections may instead be carried out in real-space, using the method of King-Smith et al. [16], which reduces the computational cost to the order of the system size squared.

For pseudo-potential approximations to function properly, they must conform with all angular momentum eigenstates \( l \), and also be a non-local:

\[ V_{NL} = \sum_l |l > V < l|. \]  

(2.22)

In our calculations in quantum transport, we employ an alternative, and very popular formulation by Kleinman and Bylander (KB) [15]. In this formulation, a separable form of non-local pseudo-potential is introduced using spherical harmonics:

\[ V_{KB}(r, r') = \sum_{l,m} |Y_{lm}(\Omega_r)W_l(r) > \epsilon_l < W_l(r')Y_{lm}(\Omega_r)|, \]  

(2.23)

where:

\[ W_l(r) \equiv V_i \Psi_l(r) \]  

(2.24)

\[ \epsilon_l \equiv \frac{< \Psi_l(r)|V_l|\Psi_l(r)>}{< \Psi_l(r)|V_l^2|\Psi_l(r)>}. \]  

(2.25)

This separation allows efficient linear scaling of the calculation with the size of the basis set. KB formulation can lead to unphysical ghost states at energies below or near the physical valence states [17]. These ghost states arise from the KB nonlocality,
in which the nodeless pseudo wavefunctions need not to be the lowest eigenstates. However, this problem is avoided by choosing proper local component and the core cutoff radius of the potentials [18].

2.5.1 Non-linear Core Corrections

Within the pseudopotential approximation, the core energy is assumed to be constant and all interaction between the core and valence electrons is replaced by the pseudo-potential. Essentially, this means that the core and valence electron interaction is linearized, whereas the actual interaction is non-linear. While this is a good approximation when the core and valence electrons are well separated, there would be a major error if the two densities overlap significantly. Especially, spin density calculations introduce additional non-linearity and this necessitates the explicit consideration of the non-linear dependence of the energy on the core charge density. Conventionally, the ionic pseudo-potential is defined as the difference between the Coulomb and exchange-correlation potentials due to the valence charge density, $\rho^v(r)$, and spin polarization,

$$\xi^v(r) = \frac{\rho^v_\uparrow (r) - \rho^v_\downarrow (r)}{\rho^v(r)}$$  \hspace{1cm} (2.26)

from the neutral atomic pseudo-potential. This is a decoupling of core and valence electrons. But in the pseudo-potential approximation, the total exchange-correlation potential is given as:

$$V_{XC}(\rho^v + \rho^c, \xi) = V_{XC}(\rho^v + \rho^c, \xi) - V_{XC}(\rho^v, \xi) + V_{XC}(\rho^c, \xi),$$  \hspace{1cm} (2.27)

where

$$\xi(r) = \frac{\rho^c_\uparrow (r) - \rho^c_\downarrow (r)}{\rho^c(r) + \rho^v(r)}.$$  \hspace{1cm} (2.28)

The first two terms in Eq. (2.27) are included in pseudo-potentials and the last term is calculated in the self consistent KS iterations. Clearly, the $\xi$ in Eq. (2.28) can be
very different than the spin polarization in Eq. (2.26). Because the $V_{XC}$ is a non-linear function of density, the ionic pseudo-potentials depend on the valence density. The solution to this, proposed by Louie et al. [19], is to redefine the ionic pseudo-potential as that given by subtracting the Coulomb potential of the valence charge density and the total exchange-correlation potential $V_{XC}(\rho^v + \rho^c, \xi)$. The core charge is then stored and used to reconstruct the full exchange-correlation potential in the calculation.

### 2.6 Basis Set

The Kohn-Sham orbitals, $\varphi_j(r)$ in Equation 2.18, may be represented in terms of any complete basis set. Many choices are possible including atomic orbitals, Gaussians, Linear Augmented Plane Waves (LAPW), plane waves, and Fireball Atomic orbitals [20, 21, 22, 23]. The use of a plane wave (PW) basis set offers a number of advantages, including the simplicity of the basis functions, which make no preconceptions regarding the form of the solution, the absence of basis set superposition error, and the ability to efficiently calculate the forces on atoms. In general, the representation of an arbitrary orbital in terms of a PW basis set would require a continuous, and hence infinite, basis set. However, the imposition of periodic boundary conditions allows the use of Bloch’s Theorem making the basis set into a discrete one:

$$\varphi_{j,k}(r) = \sum_G c_{j,k}(G)e^{i(k+G)r},$$  \hspace{1cm} (2.29)

where the sum is over reciprocal lattice vectors $G$ and $k$ is a symmetry label which lies within the first Brillouin zone. Thus, the basis set for a given $k$ will be discrete, although in principle it will still be infinite. In practice, the set of plane waves is restricted to a sphere in reciprocal space most conveniently represented in terms of a cut-off energy, $E_{cut}$, such that for all values of $G$ used in the expansion

$$\frac{\hbar^2|\mathbf{k} + \mathbf{G}|^2}{2m_e} \leq E_{cut},$$  \hspace{1cm} (2.30)
Thus, the convergence of the calculation with respect to basis set may be ensured by variation of a single parameter, $E_{\text{cut}}$. This is a significant advantage over many other basis set choices, with which calculated properties often show extreme sensitivity to small changes in basis set and no systematic scheme for convergence is available. The choice of periodic boundary conditions is natural in the case of bulk solids which exhibit perfect translational symmetry. If we wish to model an isolated molecule we must artificially introduce periodic boundary conditions by construction of a super cell. In this scheme the calculations are performed on a periodic array of molecules, separated by large vacuum regions. In the limit of large separation between the periodic images, the results will be those for an isolated molecule.

The principle disadvantage of the use of a PW basis set is the number of basis functions required to accurately represent the Kohn-Sham orbitals. This problem may be reduced by the use of pseudo-potentials but several hundred basis functions per atom must still be used, compared with a few tens of basis function with the use of some atom-centered basis sets.

### 2.6.1 Fireball Atomic Orbital Basis

In the second part of this thesis, we study quantum transport in nanotube based devices. Unlike bulk crystal or an isolated molecule, this system is in nonequilibrium and open. Therefore, can not be treated within conventional DFT with periodic boundary conditions. We employ DFT calculations within Keldysh nonequilibrium Green’s functions (NEGF) formalism. However, the Green’s functions are obtained inverting the Hamiltonian matrix, so that the Hamiltonian matrix can not be very big. To avoid this complication, we first partition the system such that only a specific part of the Hamiltonian is relevant. This method will be discussed in Chapter 3. Then we further optimize by using smaller and more efficient basis set, which can represent the electronic density, and thus the exchange correlation functional with reasonable accuracy. In our calculations of quantum transport, we use linear combi-
nations of atomic orbitals (LCAO) basis set including the $s, p$ pseudo-atomic orbitals. Special case of this basis is so called the Fireball basis set, which is derived from the eigenfunctions of an *ab-initio* pseudopotential.

First, parameters from the norm conserving pseudopotential developed by Hamman, Schüller and Chiang (HSC) [14] are used in solving the $l$-dependent radial KS equations self consistently. This results in a self consistent radial potential $V_{SC}(r)$. Then, the fireball orbitals are obtained by adding a potential barrier at $r = R_{cutoff}$ to the self consistent radial potential of a free atom. The addition of the potential barrier at cut-off radius localizes the orbitals and increases the kinetic energy. Therefore, the fireball orbitals are accurate only when the cut-off radius is sufficiently large. In this work, we used a minimal fireball basis set, which includes only $\{s, p\}$ orbitals. The basis set can be extended to include $\{d\}$ orbitals, since the form of the matrix elements involved are valid in general. In practice, including $\{d\}$ orbitals comes at a cost of computational time.

For a given pseudopotential, the atomic orbital wavefunctions are obtained in the form:

$$\phi_{lm}(r) = R(r)Y_{lm}(\Omega_r).$$

(2.31)

The radial function $R$ is restricted by the cut-off radius $R_{cutoff}$. In many cases of numerical calculations, it is convenient to work in real space Cartesian basis sets. The Cartesian basis set $\{\phi_\mu\}$ with the composite index $\mu = s, p_x, p_y, p_z$ is constructed by the linear combination spherical harmonics:

$$\phi_s(r) = \phi_{00}(r) = \frac{1}{\sqrt{4\pi}}R_s(r),$$

$$\phi_{p_x}(r) = \frac{1}{\sqrt{2}}(\phi_{1,-1}(r) - \phi_{1,1}(r)) = \frac{3}{\sqrt{4\pi}}R_p(r)\frac{x}{r},$$

$$\phi_{p_y}(r) = \frac{i}{\sqrt{2}}(\phi_{1,-1}(r) - \phi_{1,1}(r)) = \frac{3}{\sqrt{4\pi}}R_p(r)\frac{y}{r},$$

$$\phi_{p_z}(r) = \phi_{10}(r) = \frac{3}{\sqrt{4\pi}}R_p(r)\frac{z}{r}. \quad (2.32)$$

Then the solution of the self consistent KS equation are expressed in terms of the
Cartesian set defined above (Eq. 2.32):

$$\Psi^i = \sum_{\nu} c^i_{\nu} \phi_{\nu a}(\mathbf{r} - \mathbf{R}_I), \quad (\nu = s, p, I),$$  \hspace{0.5cm} (2.33)

where $I$ denotes the index of atom I. Then the KS equation becomes a matrix equation:

$$H_{\mu\nu} c^i_{\nu} = E^i S_{\mu\nu} c^i_{\nu}. \hspace{0.5cm} (2.34)$$

The overlap matrix in the above equations needs to be treated with care, since the basis set $\{\phi_{\mu}\}$ consists of nonorthogonal states centered at different atoms. The Hamiltonian matrix elements for KS equation are now defined by:

$$H_{\mu\nu} = \int d\mathbf{r} \phi^*_\mu(\mathbf{r} - \mathbf{R}_I) \left[ -\frac{\nabla^2}{2} + \sum_{K} V^{N A, K}(\mathbf{r} - \mathbf{R}_K) \right] \phi_{\nu}(\mathbf{r} - \mathbf{R}_J)$$

$$+ \int d\mathbf{r} d\mathbf{r}' \phi^*_\mu(\mathbf{r} - \mathbf{R}_I) \sum_{KL} (\mathbf{r} - \mathbf{R}_K, \mathbf{r}' - \mathbf{R}_L) \phi_{\nu}(\mathbf{r}' - \mathbf{R}_J)$$

$$+ \int d\mathbf{r} \phi^*_\mu(\mathbf{r} - \mathbf{R}_I) [-V_{\delta H}(\mathbf{r}) + V_{XC}(\mathbf{r})] \phi_{\nu}(\mathbf{r} - \mathbf{R}_J),$$

$$S_{\mu\nu} = \int d\mathbf{r} \phi^*_\mu(\mathbf{r} - \mathbf{R}_I) \phi_{\nu}(\mathbf{r} - \mathbf{R}_J). \hspace{0.5cm} (2.35)$$

Then the real space electronic density is given by the density matrix:

$$\rho(\mathbf{r}) = \sum_{\mu\nu} \phi^*_\mu(\mathbf{r} - \mathbf{R}_I) \hat{\rho}_{\mu\nu} \phi_{\nu}(\mathbf{r} - \mathbf{R}_J). \hspace{0.5cm} (2.36)$$

Fireball basis set is known to give sufficiently accurate results for a wide range of systems. For detailed description, I redirect readers to more elaborate publications on the Fireball basis set [24, 23, 25, 26, 27].
Chapter 3

Theory of Quantum Transport

Molecular electronic devices are what scientists believe to be the next logical step in the miniaturization of electronic devices. The development of nano-technology has made it possible to engineer materials and devices on a length scale as small as several nanometers (atomic distances are 0.1 nm). The properties of such nanoscale structures cannot be described in terms of macroscopic parameters like mobility and diffusion coefficients and calls for a microscopic or atomistic viewpoint. At this scale, other characteristic length scales become more important in describing the system [28]:

**Fermi wavelength** $\lambda_F$ is the wavelength of electrons at the Fermi surface $\lambda_F \equiv \frac{2\pi}{k_F}$.

When the system size is comparable with the Fermi wavelength $L \sim \lambda_F$, the wave nature of electrons prevails. Fermi wavelength depends strongly on the dimensionality of the system that $\lambda_F$ is larger in quasi-one-dimensional systems.

**Mean free path** $L_m$ is the average distance for which electron travels before changing its momentum in a scattering process. The mean free path is estimated as $L_m = v_F \tau_m$, where $\tau_m$ is the average momentum loss time.

**Phase coherence length** $L_\phi$ is distance for which electron travels before it loses its phase information. Static scatterers with no internal degrees of freedom,
such as impurities, do not contribute to phase relaxation. Main contributors
are electron-electron and electron-phonon scattering. Therefore, the phase re-
relaxation time $\tau_\phi$ is typically longer than the collision time. The phase coherence
length is estimated as $L_\phi \sim v_F \tau_\phi$.

**Thermal diffusion length** $L_T$ represents broadening due to finite electronic tem-
perature. For electrons near the Fermi level, the average phase coherence lives
for approximately $\tau = \frac{\hbar}{k_B T}$ and a distance $L_T \sim \sqrt{\frac{D \hbar}{k_B T}}$, where $D$ is the
diffusion coefficient.

**Screening length** $L_s$ is the distance over which a density perturbation is screened
by an electron gas and proportional to density of states and strongly depends on
the dimensionality of the system. The screening length of quasi-one-dimensional
systems can be much longer than in the bulk.

When the system length $L$ is smaller than the mean free path $L_m$, i.e. $L < L_m$,
the electrons flow without changing their momentum and the transport is ballistic.
The electrons suffer no inelastic scattering, thus retaining their phase during the
transport if the system length is smaller than that phase coherence length $L_\phi$ and
the thermal diffusion length $L_T$. Therefore, the electronic transport at these length
scales is considered to be in the phase coherent ballistic regime. Quantum transport
theory in the phase coherent ballistic regime is described by the Landauer picture,
which will be discussed in the next section.

### 3.1 Landauer-Büttiker Formulation

In a macroscopic system, the electric current is described by the Ohm’s law, and
the resistance is defined in terms the conductor’s length $L$ and cross sectional area $A$:

$$R = \frac{L}{\sigma A},$$

$$G = R^{-1},$$  \hspace{1cm} (3.1)
where $\sigma$ – the conductivity – is the coefficient connecting these two quantities. The inverse of the resistance, $G$, is defined as the conductance. Experimentally, it has been found, that at the nanoscale, conductance is quantized in units of $(2e^2/h)$ [29].

Typical quantum transport device consists of an ideal conductor connected to two electron reservoirs, which are large and in equilibrium with well defined electrochemical potentials. The ideal conductor is made of quasi-one-dimensional unit cells with length $L_c$. It is assumed that electrons are defined by 1D effective potential with no magnetic field. To single out the electronic transport between the two reservoirs through this idealized conductor, it is assumed that there is no back scattering at junctions of conductors and the reservoirs. Therefore, the population of the states in the conductor is determined only by the electrochemical potentials of the reservoirs, and the energies of electrons from the reservoirs. Then the wave function in the conductor will be in a Bloch state:

$$\Psi_{kn}(r) = u_{kn}(r)e^{ikz}, \tag{3.2}$$

where $\mathbf{k}$ is a wavevector in the Brillouin zone, $z$ is the propagating direction, and the $u$ is a periodic function $u_{kn}(x, y, z) = u_{kn}(x, y, z + L_c)$. If one denotes the energy of this state as $E_n(k)$, then the sign and group velocity of this wave is given by $v_{kn} = dE_n(k)/dk$. The number of states (or modes) propagating in $+z$ and $-z$ directions are equal and can be represented by a function $M(E)$, which of course depends on the electron energy. Recalling the assumption made before, the occupation of the propagating modes will be given by distribution of electrons $f^{+/-}(E)$ in the right and left reservoirs. A uniform electron gas with $n$ electrons per unit length moving with velocity $v$ carries a current $I = env$. In our case, electron density for a single $k$-state is $(1/L_c)$. Therefore, we can write the current $I^+$ carried by the modes propagating in $+z$ direction as

$$I^+ = \frac{e}{L_c} \sum_{kn}^+ v_{kn} f^+(E_n(k)) = \frac{e}{L_c} \sum_{kn}^+ \frac{1}{\hbar} \frac{\partial E_n(k)}{\partial k} f^+(E_n(k)), \tag{3.3}$$
where the sum is over Bloch’s states with positive \( v \). Converting the sum over \( k \) into integral

\[
\sum_k \rightarrow 2 \times \frac{L_c}{2\pi} \int dk,
\]  

(3.4)
on one obtains

\[
I^+ = \frac{2e}{h} \int_{-\infty}^{+\infty} f^+(E)M(E)dE,
\]  

(3.5)
where \( M(E) \) is the number of propagating modes at energy \( E \). A current propagating in the \(-z\) direction is obtained in a similar way. If the two reservoirs are at different electrochemical potentials, total current can be obtained by adding \( I^+ \) and \( I^- \):

\[
I = \frac{2e^2}{h} M \frac{\mu_1 - \mu_2}{e} = \frac{2e^2}{h} M (V_1 - V_2),
\]  

(3.6)
where it is assumed that the number of modes \( M \) is a constant in the energy range \( \mu_1 < E < \mu_2 \). Physically, this situation describes a conductor in between two metallic leads at a bias voltage. Therefore, the conductance is obtained as

\[
G = \frac{2e^2}{h} M.
\]  

(3.7)
We note that the conductance is quantized due to integer nature of \( M \). Furthermore, the resistance is now given by:

\[
R = G^{-1} = \frac{(\mu_1 - \mu_2)/e}{I} = \frac{h}{2e^2 M} \approx \frac{12.9k\Omega}{M}.
\]  

(3.8)
Looking at this formula, it is clear, that an ideal conductor has a finite resistance, even though there is no electron scattering. The resistance arises from the contact of the conductor and the reservoirs, since only a finite number of modes that can carry a current within a real real conductor.

If we introduce some time independent scattering imperfections into such an ideal conductor, an electron propagating in mode \( i \) will have probability \( T_i \) of passing through the conductor, and \( R_i = 1 - T_i \) of scattering back into the reservoir of its origin. Then, Eq. (3.7) has to be modified to read:

\[
G = \frac{2e^2}{h} \sum_{i=1}^{M} T_i = \frac{2e^2}{h} MT,
\]  

(3.9)
where $T$ is the average probability of an electron passing through the conductor. This relation is known as the Landauer formula [30]. Extending Landauer’s formula for multiple reservoirs and multiple scatterers was done by Büttiker, giving the famous Landauer-Büttiker formula [31, 32]:

$$G_{\alpha\beta} = \frac{2e^2}{h} \text{Tr}(s_{\alpha\beta}^* s_{\alpha\beta}) = \frac{2e^2}{h} T_{\alpha\beta},$$

(3.10)

where $G_{\alpha\beta}$ and $s_{\alpha\beta}$ are the corresponding conductance and scattering matrices.

For system consisting of two metal leads and conductor, use of the Landauer-Büttiker formula gives the expression for the current:

$$I = \frac{2e^2}{h} \int_{-\infty}^{+\infty} T(E)[f_L(E) - f_R(E)]dE,$$

(3.11)

in terms of $T(E)$—i.e. the total transmission probability through all modes, where population is given by the Fermi distribution in the leads,

$$f_{L/R} = \frac{1}{e^{\beta(E-\mu_{L/R})} + 1},$$

(3.12)

where $\beta = 1/k_B T$, and $\mu_{L/R}$ is the electrochemical potential in left/right lead, respectively.

### 3.2 DFT-NEGF method for Transport problems

Typical molecular electronic devices consists of two leads and the central region, which is usually a cluster of atoms or an individual molecule. Furthermore, the two leads may be of a different design or species. These inhomogeneities lead to scattering of the electrons in such devices, rendering them inherently open. Also, transport under bias voltage is inherently a non-equilibrium process. To treat this kind of open, nonequilibrium systems, the calculation has to be modified to account for these two key properties: one of such scheme was introduced by Taylor and coworkers[33] in 2001.
Consider the surface of a crystal. Near the surface, the Kohn-Sham (KS) potential will be affected by the surface reconstructions and specific charge redistributions. But sufficiently deep inside the crystal, the KS potential will be restored to its bulk value. This important observation is key to understanding the combined DFT and Non Equilibrium Green’s Function (DFT-NEGF) method [25].

A typical device system may be divided into three regions (Fig. 3.1): the left lead, a central region and the right lead. The central region includes a particular molecule with large parts of the left and right leads at its ends. If the parts of the leads in the central region are large enough, effects due to the interface between the molecule and the leads are screened at the boundaries of the central region. Therefore, the boundary condition for the central part can be written as:

\[
V^{\text{eff}}(\mathbf{r}) = \begin{cases} 
V^{\text{eff}}_l(\mathbf{r}), & z \leq z_l \\
V^{\text{eff}}_c(\mathbf{r}), & z_l < z < z_r \\
V^{\text{eff}}_r(\mathbf{r}), & z \geq z_r
\end{cases}
\]  
(3.13)

where the planes \( z = z_l \) and \( z = z_r \) are deep inside the leads so that (3.13) holds. This partitioning makes 2 important computational steps possible:

1. The lead electronic structures can now be calculated separately by conventional DFT method, just as in bulk calculations.

2. Once the lead electronic densities are obtained and stored in the database, they are used to set the boundary conditions for KS the potentials for the central region. That is, the Hartree potentials of the central region at \( z_l \) and \( z_r \) are equal to the Hartree potentials of the leads at the corresponding planes:

\[
V^H_c(\mathbf{r})|_{z_l} = V^H_{l,\text{bulk}}(\mathbf{r})|_{z_l}, \\
V^H_c(\mathbf{r})|_{z_r} = V^H_{r,\text{bulk}}(\mathbf{r})|_{z_r},
\]  
(3.14)

where \( V^H_{l,\text{bulk}}(\mathbf{r})|_{z_l} \) and \( V^H_{l,\text{bulk}}(\mathbf{r})|_{z_r} \) are the Hartree potentials of bulk. It has been observed that the charge densities of the center region and the leads are equal at
Figure 3.1: A two-probe system is divided into three regions.

the boundaries if the Hartree potentials are equal, and therefore the total KS potentials are also equal within the local density approximation [25].

The problem of infinite open system can now be solved via calculating involving the three separate regions in space. The partitioning implicit by (3.13) gives one the freedom to choose the two leads, which may be made of different elements, be held at different electrochemical potentials etc., making it possible to calculate current-voltage (IV) characteristics of devices.

The DFT-NEGF method essentially replaces details of semi-infinite leads by Hartree potentials and electronic densities at the boundaries. Thus, the entire problem is now reduced to solving the non-equilibrium central region self consistently. In this region, we encounter two type of electronic states, propagating states for currents and bound states. Traditional basis sets such as plane wave require huge numbers of basis functions and in may cases inadequate to represent these states. In the next section, we discuss the NEGF approach, in which the electronic density is constructed by the NEGF rather than basis sets.
3.3 Non-equilibrium Green’s function approach to electronic states

The correct density matrix for a typical two probe system can be constructed if we know the left and right scattering states, and the bound states. These states can be used to diagonalize the Hamiltonian matrix of the entire device. The scattering states can be calculated efficiently, but the bound states are difficult to find[25]. The non equilibrium Green’s function is used to construct the density matrix [28, 34, 35] via:

$$\hat{\rho} = -\frac{i}{2\pi} \int dEG^<(E),$$

where the non equilibrium Green’s function $G^<$ is constructed via:

$$G^< = G^R \Sigma^<[f^{k_l^p}, f^{k_r^p}]G^A,$$

where $G^R$ and $G^A$ are the retarded and advanced Green’s functions and $\Sigma^<[f^{k_l^p}, f^{k_r^p}]$ the self energy of the respective electrodes. At equilibrium at zero temperature, the distributions are equal and given by the equilibrium Fermi distribution:

$$f^{k_l^p}(E; \mu_l + \Delta V_l) = f^{c_n}(E; \mu_c) = f^{k_r^p}(E; \mu_r + \Delta V_r) = f^{eq}(E; \mu^*).$$

When there is a non equilibrium condition $\mu_l + \Delta V_l \neq \mu_r + \Delta V_r$, and the distribution functions of the electrodes are approximated as [28]:

$$f^{k_l^p}(E; \mu_l + \Delta V_l) = f^{eq}(E; \mu_l + \Delta V_l),$$

$$f^{k_r^p}(E; \mu_r + \Delta V_r) = f^{eq}(E; \mu_r + \Delta V_r).$$

This approximation is essentially the mathematical formulation of the assumption that the device scattering region has no great influence to the distributions of the electrons in the leads. The distribution function of molecular device region is approximated as:

$$f^{c_n} = f^{eqn}(E; \mu_c)$$
where the electrochemical potential is $\mu_c = \mu_l + \Delta V_l$ or $\mu_c = \mu_r + \Delta V_r$. Given the distribution functions $f_k^n, f_c^n, f_k^n$, the self-consistent solution of KS equations is essentially exact within the DFT approximation, provided that the central part of the system is large enough [25].

### 3.4 Leads as Periodic Systems

In this section, we will discuss how electronic properties of the leads of a typical two probe device. As mentioned in the Section 3.2, leads are calculated separately as periodic system consisting of identical unit cells. As a result, the atomic structure, charge density and potentials are identical in each of the unit cells. Let $\{\zeta_z\}$ be a collection of atomic orbitals in unit cell $z$. Typically, the unit cell is chosen to be large enough that any atomic orbital in unit cell $z$ can only interact with orbitals from its immediate neighbors $z-1$ and $z+1$. The Hamiltonian then has a tridiagonal form:

$$
\begin{pmatrix}
\ddots & & & & \\
 h_{z-2,z-3} & h_{z-2,z-2} & h_{z-2,z-1} & & \\
 & h_{z-1,z-2} & h_{z-1,z-1} & h_{z-1,z} & \\
 & & h_{z,z-1} & h_{z,z} & h_{z,z+1} \\
 & & & h_{z+1,z} & h_{z+1,z+1} & h_{z+1,z+2} \\
 & & & & \ddots 
\end{pmatrix}
$$

Periodicity of the system imposes translational invariance on the Hamiltonian:

$$
\begin{align*}
h_{z-2,z-1} &= h_{z-1,z} = h_{z,z+1} = \ldots \\
h_{z-2,z-2} &= h_{z-1,z-1} = h_{z,z} = \ldots \\
h_{z-1,z-2} &= h_{z,z-1} = h_{z+1,z} = \ldots 
\end{align*}
$$
According to the Bloch’s theorem, electronic wave functions for this periodic system have the following form:

$$\Psi^k(z) = e^{ikz} \phi^k = (\lambda^k)^z \phi^k$$  \hspace{1cm} (3.23)

The Schrödinger’s equation for this system is then written as:

$$[H(z, z') - ES(z, z')]\Psi^k(z') = 0,$$  \hspace{1cm} (3.24)

where $H$ and $S$ represent the Hamiltonian and overlap operators respectively. Using equations 3.20 and 3.23, the Schrödinger equation takes following matrix form:

$$
\begin{pmatrix}
\vdots \\
 h_{z-1,z}^E & h_{z,z}^E & h_{z,z+1}^E \\
 h_{z-1,z}^E & h_{z,z}^E & h_{z,z+1}^E \\
 h_{z-1,z}^E & h_{z,z}^E & h_{z,z+1}^E \\
 \vdots \\
\end{pmatrix}
\begin{pmatrix}
\lambda^k \\
\lambda^k \\
\lambda^k \\
\vdots \\
\end{pmatrix} = 0$$  \hspace{1cm} (3.25)

where $h^E = h - Es$. Then, the final form for the Schrödinger equation becomes:

$$[h_{z,z-1}(\lambda)^{-1} + h_{z,z} + h_{z,z+1} \lambda^k] \phi^k = E [s_{z,z-1}(\lambda)^{-1} + s_{z,z} + s_{z,z+1} \lambda^k] \phi^k.$$  \hspace{1cm} (3.26)

It is clear from Eq. (3.26) that only six matrices are enough to describe this quasi-one-dimensional periodic system: $h_{z,z}, s_{z,z}, h_{z,z-1}, s_{z,z-1}, h_{z,z+1}, s_{z,z+1}$. These matrices represent couplings within the cell $z$ and with its neighbors. This leads to another important point that expectation values of physical observables are calculated using only these six matrices. For example, the electronic density $\rho$ has the following form:

$$\rho =
\begin{pmatrix}
\rho_{z-1,z-1} & \rho_{z-1,z} & 0 \\
\rho_{z-1,z} & \rho_{z,z} & \rho_{z,z+1} \\
0 & \rho_{z,z+1} & \rho_{z+1,z+1} \\
\end{pmatrix},$$  \hspace{1cm} (3.27)

where translational invariance and Hermiticity require that

$$\rho_{z-1,z-1} = \rho_{z,z} = \rho_{z+1,z+1},$$

$$\rho_{z,z-1} = \rho_{z+1,z} = \rho_{z-1,z}^\dagger = \rho_{z+1,z+1}^\dagger.$$  \hspace{1cm} (3.28)
In periodic system, Bloch functions are identical for each unit cell. However, there are infinite number of eigenfunctions corresponding to infinite number of \( k \)-points. The density \( \rho \), must therefore be evaluated by integrating over \( k \). In practice, one samples a suitable set of discrete \( k \)-points and then interpolate the values of \( \Psi^k \). The Equation (3.26) is then solved for each of the sampling \( k \)-points to obtain the Bloch eigenstates \( \phi^k \) and the density matrix is finally calculated by summing over the sampled \( k \)-points:

\[
\rho_{i,j} = \sum_k \Psi_i^k f^{EQ}(E^k; \mu) \Psi_j^k = \sum_k f^{EQ}(E^k; \mu) \phi_i^k e^{ik(j-i)} \phi_i^k,
\]

(3.29)

where \( f^{EQ}(E^k; \mu) \) is a Fermi distribution at the chemical potential \( \mu \). The chemical potential of the leads are determined from the charge neutrality condition

\[
\text{Tr} [\rho f^{EQ}(E^k; \mu)] = \sum_k f^{EQ}(E^k; \mu) = N.
\]

(3.30)

Once the densities for the leads are calculated, effective potential, Hamiltonian and overlap matrices are calculated and used in the KS equation for the central region as boundary conditions.

### 3.5 Central region of two probe device

As discussed in the Section 3.2, the use of NEGF-DFT approach is typical for the treatment of two probe systems. In this approach, the lead characteristics are precalculated, as well as their densities, and effective potentials, the Hamiltonian and overlap matrices are used as boundary conditions for the central region. The central region, consisting of large parts of the leads and the central cluster, represents the open and nonequilibrium character of the entire system. Therefore, nonequilibrium Green’s function is more efficient in constructing the density matrix in the central region. The retarded Green’s function \( G^R \), for a Hamiltonian \( H \) at energy \( E \) is given as:

\[
G^R(E) = ((E + i\eta)I - H)^{-1}, \eta \to 0.
\]

(3.31)
It is clear from the Eq. (3.31) that calculation of the Green’s function for an infinite system through matrix inversion can not be done. However, to calculate the charge distribution of the central region, we only need the components of the density matrix in which the basis states have support. Hamiltonian for the central region has a form:

$$
\begin{pmatrix}
  h_{l,-\infty} & h_{l,l+1} & 0 & 0 & 0 \\
  h_{l,l-1} & h_{l,l} & h_{l,c} & 0 & 0 \\
  0 & h_{c,l} & h_{c,c} & h_{c,r} & 0 \\
  0 & 0 & h_{r,c} & h_{r,r} & h_{r,r+1} \\
  0 & 0 & 0 & h_{r,r-1} & h_{r,\infty}
\end{pmatrix}, \quad (3.32)
$$

where $h_{l,\infty}$ is the coupling within the leads, and indexes $l, c, r$ denote the left lead, central region and the right lead. This matrix is infinite due to the semi-infinite lead Hamiltonians $h_{l,-\infty}$ and $h_{r,\infty}$:

$$
h_{l,-\infty} \equiv \begin{pmatrix}
  \ddots & \ddots & \ddots & 0 \\
  0 & h_{l,l-1} & h_{l,l} & h_{l,l+1} \\
  0 & 0 & h_{l,l-1} & h_{l,l}
\end{pmatrix}, \quad (3.33)
$$

$$
h_{l,\infty} \equiv \begin{pmatrix}
  h_{r,r} & h_{r,r+1} & 0 & 0 \\
  h_{r,r-1} & h_{r,r} & h_{r,r+1} & 0 \\
  0 & \ddots & \ddots & \ddots
\end{pmatrix}. \quad (3.34)
$$

Overlap matrices $s_{l,-\infty}$ and $s_{l,\infty}$ can also be defined similar way. Bias or gate voltages can be applied by shifting the lead potentials:

$$
V_{i/r,bulk}^{\text{eff}}(r) \leftarrow V_{i/r,bulk}^{\text{eff}}(r) + \Delta V_{i/r},
$$

$$
V_{c}^{H}(r)|_{z_{i/r}} \leftarrow V_{c}^{H}(r)|_{z_{i/r}} + \Delta V_{i/r},
$$

$$
h_{i/r,\infty} \leftarrow h_{i/r,\infty} + s_{i/r,\infty}\Delta V_{i/r}. \quad (3.35)
$$

The Hamiltonian matrix (3.32) can be diagonalized using scattering eigenstates. But bound states are rather difficult to calculate [25]. Therefore, the non-equilibrium
Green’s function $G^<$ is used in constructing the density matrix. The non-equilibrium Green’s function is derived from the retarded Green’s function:

$$G^< = G^R \Sigma^< [f^{k_l^p}, f^{k_r^p}] G^A,$$  \hspace{1cm} (3.37)

where $\Sigma^< [f^{k_l^p}, f^{k_r^p}]$ represents injection of charge from the leads and can be written in terms of self energies $\Sigma^L_{l,l}$ and $\Sigma^R_{r,r}$ due to coupling to the left and right leads [28]:

$$\Sigma^< [f^{k_l^p}, f^{k_r^p}] = -2i Im(f^{k_l^p} \Sigma^L_{l,l} + f^{k_r^p} \Sigma^R_{r,r}).$$  \hspace{1cm} (3.38)

Therefore, it is enough to define only the retarded Green’s function in the central region. Expansion of the retarded Green’s function by real space basis set:

$$G^R_{\mu,\nu}(r, r') = |\zeta_{\mu}(r) > G^R_{\mu,\nu} < \zeta_{\nu}(r')|$$ \hspace{1cm} (3.39)

leads to the equation for $G^R_{\mu,\nu}$:

$$\lim_{\eta \to 0} [(E + i\eta) S_{\mu,\nu'} - H_{\mu,\nu'}] G^R_{\nu',\nu} = \delta_{\mu,\nu}.$$ \hspace{1cm} (3.40)

To calculate the charge distribution of the central region, we only need the components of the density matrix in which the basis states have support. That means a submatrix of $G^R_{\mu,\nu}$ corresponding to the central region needs to be calculated. The retarded Green’s function for the central region in LCAO orbital space is given as [28]:

$$g_{CC} = \begin{pmatrix}
    h_{l,l}^E - \Sigma^l_{l,l} & h_{l,c}^E & 0 \\
    h_{c,l}^E & h_{c,c}^E & h_{c,r}^E \\
    0 & h_{r,c}^E & h_{r,r}^E - \Sigma^r_{r,r}
\end{pmatrix}^{-1},$$ \hspace{1cm} (3.41)

where $h_{i,j}^E \equiv (E + i\eta)s_{i,j} - h_{i,j}$. The self energies can be written, in terms of surface Green’s functions:

$$\Sigma^l_{l,l} \equiv h_{l,l-1}^E g_{l,\infty} h_{l,l+1}^E,$n  \hspace{1cm} (3.42)

$$\Sigma^r_{r,r} \equiv h_{r,r-1}^E g_{r,\infty} h_{r,r+1}^E.$$

There are many methods for evaluating the self energy due to coupling to the leads [28, 36]. In method developed by Sanvito et al. [37] is better suited for this type of calculations [25].
3.6 Integrating the Green’s Function

One of the central points in the NEGF-DFT calculation is an accurate and efficient integration of the Green’s functions. The charge density is calculated using the non-equilibrium Green’s function \( G^< \) (Eq. (3.15)). In contrast with the retarded Green’s function \( G^R \), the non-equilibrium Green’s function contains information about the electronic distribution of the leads through the self energy of the leads \( \Sigma^< \). In the equilibrium, the chemical potentials of the leads are equal, and thus, the non-equilibrium and retarded Green’s functions relate as:

\[
\text{Re}[G^<(E)] = 2\text{Im}[G^R(E)]. \tag{3.43}
\]

This relation is true even if the leads have differing chemical potentials as long as the distribution functions satisfy:

\[
f_l(E) = f_r(E). \tag{3.44}
\]

Therefore, the Eq. (3.15) can be expressed as sum of two terms: contributions from electrons satisfying relation Eq.(3.44), and non-equilibrium contribution:

\[
\hat{\rho} = \frac{2}{\pi} \text{Im}\left[ \int_{\mu_{\text{min}}}^{\mu_{\text{max}}} dE G^R(E) \right] + \frac{1}{\pi} \text{Re}\left[ \int_{\mu_{\text{min}}}^{\mu_{\text{max}}} dE G^<(E) \right], \tag{3.45}
\]

where \( \mu_{\text{min}} = \min(\mu_l + V_l, \mu_r + V_r) \) and \( \mu_{\text{max}} = \max(\mu_l + V_l, \mu_r + V_r) \) [25]. In many cases, this integral is non-trivial to perform along the real axis, due to non-analytic functions or divergences of the integrands. This problem is solved using a contour integral technique in complex plane. As an example, let us consider the first term of Eq. (3.45). The retarded Green’s function \( G^R \) can be expressed in terms of the eigenstates \( \{\Psi^n\} \) of the Hamiltonian their eigenvalues \( \{E_n\} \):

\[
G^R(E) = \lim_{\eta \to 0} \sum_n \frac{|\Psi^n \rangle \langle \Psi^n|}{E - E_n + i\eta}. \tag{3.46}
\]

If the \( G^R \) has some of its poles near \( E = E_m - i\eta \), where the \( \{E_m\} \) are the discrete energies of bound states, it is analytic above the real axis [38]. Therefore, the first
term of the Eq. (3.45) can be integrated along a semi-circle in the upper half complex plane (Fig. 3.2). The second term in Eq. (3.45) corresponds to the non-equilibrium contribution, for which $Re(G^<) \neq Im(G^R)$, and expressed by the non-equilibrium Green’s function $G^<$. The $G^<$ is analytic only along the real axis, and therefore the non-equilibrium charge distribution has to be calculated along the real axis. This is sometimes problematic if there are sharp peaks in the density of states due to a bound states.

### 3.7 Analysis of Real and Complex band structures

After obtaining the electronic density and Hamiltonian self consistently using the NEGF-DFT, extensive analysis is needed to understand detailed electronic structures of the two probe system. Band structure calculations are one of the key tools of such analysis.

The Fireball atomic orbitals basis set allows to diagonalize only the relevant sub-
matrix of the Hamiltonian matrix for the entire system, and calculate renormalized band structure for the corresponding part of the system, such as the central region. After obtaining the self consistent Hamiltonian, only the submatrix of H corresponding to interactions within a particular unit cell, and the submatrix corresponding to interactions of the unit cell with its neighbors are taken to solve a Bloch problem for the Eq.(3.26). This equation gives all modes, including the evanescent modes with complex $k$-vectors. The evanescent modes appear as loops in the imaginary half of $k$-space, connecting pair of real bands. Because the complex bands feature loops, imaginary part of their $k$-vector never extend beyond certain finite value. Therefore, decay length of the evanescent modes into the system is also finite.

We also calculate band structures of infinitely long nanotubes. When compared with the band structure of the central region of the two probe system, it is important tool in understanding the electronic structure of the device.

3.8 Bound States

The central scattering region in of a two probe system naturally gives rise to bound states. The bound states are localized in the central region and decay into the leads, while the scattering states correspond to propagating electrons into and out of the central region. While the scattering states in the system can be calculated effectively, the bound states are difficult to calculate numerically. The main reasons are the highly nonlinear dependence of these functions on energy $E$, and the fact that it is impossible to predict the number bound states in a priori [25]. In this section, we discuss nature of the bound states and the reason why they are difficult to treat numerically.

The bound states may arise by following mechanisms: some of the molecular states in the central region have energies below the propagating threshold of the leads [25], similar states may have energies in the band of the leads. There is another mechanism
which gives rise to the bound states due to termination of the system. For example, if the system consists of two semi-infinite leads in close proximity to each other, there may arise bound states with discrete energies in the band gap of the leads.

Let us consider the discrete set of bound states \( \{ \Psi^n \} \), which are localized in the central region. These states can be expressed in terms of evanescent modes \( \{ \varphi \} \), which are eigenstates of the Eq. (3.26) with complex \( \lambda \):

\[
\Psi^n = \begin{cases}
\varphi_{l}^{k_{l}^{m} c_{n}^{m}}, & z \in l, \\
\varphi_{c}^{k_{c}^{m}}, & z \in c, \\
\varphi_{r}^{k_{r}^{m} c_{n}^{m}}, & z \in r,
\end{cases}
\tag{3.47}
\]

where the indices \( k_{l}^{m} \) and \( k_{r}^{m} \) represent sums over only the evanescent modes. Applying the Hamiltonian to this bound state leads to the following homogeneous equation for \( \Psi^n \):

\[
\begin{pmatrix}
A_{l}^{k_{l}^{m}} & h_{l,c}^{E} & 0 \\
A_{c}^{k_{l}^{m}} & h_{c,c}^{E} & A_{c}^{k_{r}^{m}} \\
0 & h_{r,c}^{E} & A_{r}^{k_{r}^{m}}
\end{pmatrix}
\begin{pmatrix}
a_{l}^{k_{l}^{m} c_{n}^{m}} \\
\Psi_{c}^{n} \\
b_{l}^{k_{l}^{m} c_{n}^{m}}
\end{pmatrix}
= \begin{pmatrix}
0 \\
0 \\
0
\end{pmatrix},
\tag{3.48}
\]

where the elements \( A_{l}^{k_{l}^{m}} \) represents the terms:

\[
\begin{align*}
A_{l}^{k_{l}^{m}} & \equiv [h_{l,l}^{E} + h_{l,l-1}^{E} (\lambda_{l}^{k_{l}^{m}})^{-1}] \varphi_{l}^{k_{l}^{m}}, \\
A_{r}^{k_{r}^{m}} & \equiv [h_{r,r}^{E} + h_{r,r+1}^{E} \lambda_{r}^{k_{r}^{m}}] \varphi_{r}^{k_{r}^{m}}, \\
A_{c}^{k_{l}^{m}} & \equiv h_{c,l}^{E} \varphi_{l}^{k_{l}^{m}}, \\
A_{c}^{k_{r}^{m}} & \equiv h_{c,r}^{E} \varphi_{r}^{k_{r}^{m}}.
\end{align*}
\tag{3.49}
\]

The above quantities in Eq. (3.49) depend on the Bloch states and therefore highly nonlinear functions of energy \( E \). As a result, Eq. (3.48) is a highly nonlinear root finding problem for each of the discrete eigenvalues \( E \) of the bound states. Thus, finding the bound states numerically becomes very time consuming [25].
3.9 Computational Scheme

In this final section of the charter, the NEGF-DFT technique is summarized into a general computational scheme.

1. The electronic structures of the leads must be calculated first, using periodic supercells and conventional DFT method. The resulting density is stored.

2. The central device is created with enough part of the electrodes on each side, so that the boundary condition (3.13) is satisfied. In many cases, it is usually taken 3 to 4 atomic depending on the cut off radius of the basis atomic states.

3. Then the KS equation for the central region is solved self consistently with boundary conditions set by data from the lead calculations. Standard norm conserving pseudopotentials are used to represent atomic cores. The electronic wave function is expanded first, using minimal s, and p Fireball atomic orbitals basis set to construct the Hamiltonian matrix. The Green’s functions $G^R$, and $G_A$ can be calculated by a direct matrix inversion, while the self energies are calculated by standard iterative technique. [37]. Then the density is constructed using the NEGF and used in calculating KS potential in the central region.

Once the self consistency is reached, the electric density in the central scattering region is calculated using Eq. (3.50). Then the current is calculated using the Landauer formula (Eq. 3.15.3.11)

$$I = \frac{2e}{h} \int_{\mu_{\min}}^{\mu_{\max}} dE (f_l - f_r) T(E, V_b),$$

(3.50)

where $\mu_{\min} = min(\mu + V_b, \mu)$, $\mu_{\max} = max(\mu + V_b, \mu)$ represent the minimum and maximum electrochemical potentials of the leads. The transmission probability $T(E)$ as a function of electron energy is now calculated using the Green’s functions and the Self energies [28]:

$$T(E, V_b) = 4Tr[Im(\Sigma_d G^R \Sigma_d G^A)].$$

(3.51)
This numerical scheme was first implemented in its entirety in MCDCAL software developed at McGill University, Montreal, Canada. Today, number of softwares are available such as TranSIESTA, developed by Atomistix, Copenhagen, Denmark, and MatDCAL, developed at McGill University, Montreal, Canada.
Chapter 4

Theoretical Investigation of the interaction of Glycine with the diamond C(100) and C(111) (2 × 1) surfaces

4.1 Abstract

With density functional theory based simulations, we have investigated the binding of the amino acid glycine on two of the most prominent diamond surfaces – i.e., C(100) and C(111) (2 × 1) – with a focus on the associated energetics, charge transfer, electronic, and structural characteristics. With regards to the dimerized C(100) surface, interaction is mostly via the amino group of the glycine molecule (both with and without H-atom abstraction), or the hydroxyl group with the loss of an associated H-atom. Barriers for these and other reactions were estimated with quantum
chemistry methods. In contrast, the C(111) (2 × 1) surface was found to be mostly inert with respect to interactions with the glycine molecule.

### 4.2 Introduction

Driven by the advent of nanotechnology, there is considerable interest in bringing together materials with disparate electronic and mechanical properties, so as to form new multifunctional devices[39]. Particularly in the areas of molecular electronics and nanoscale sensors, it is hoped that the combination of organic molecules, and especially biomolecules, with current semiconductor technology will lead to the formation of new synthetic devices that are ultra-sensitive and chemically highly specific[6]. One of the key aspects for the formation of such devices is an understanding of how to combine and interface between the different types of materials. In this work, we report on the results of an extensive density functional theory based investigation of the adsorption of glycine with two of the most prominent diamond surfaces – \textit{i.e.}, the C(100) and C(111) (2 × 1) surfaces.

The motivation for investigating this system is the following. Diamond has unique physical and mechanical properties, despite having the identical crystal structure with common semiconductors such as silicon and germanium. With an indirect band gap of 5.5 eV, diamond offers high thermal conductivity and a broad optical transparency [40, 41]. Thus, as a wide band gap semiconductor, it is expected to outperform its other semiconductor relatives under more extreme conditions[41]. While its surfaces are reconstructed[42, 43, 44], many of them distinguish themselves through their chemical inertness, low surface energy, negative electron affinity, and high over-voltage electrode properties[45]. Moreover, recent progress in chemical vapor deposition (CVD) technology has enabled the fabrication of large diamond surfaces, paving the way for its usage in a variety of novel application devices[46]. Given this progress, it is natural to try to combine diamond with organic molecules – and especially bioac-
tive molecules – for unique functionalization. Most complex biomolecules, however, are made up of simpler organic “subunits”. For example, proteins consist primarily of folded amino acids. Hence in order for the community to gain a systematic understanding of the binding of complex biomolecules to the diamond surface, it is important to benchmark the adsorption of the simpler organic units. These can not only interface with the more complex biomolecules, but may also impart surface characteristics that are important in their own right. Here, we concentrate on the amino acid glycine on the diamond surfaces. Glycine, which is characterized by a carboxylic acid and amide group linked together with a $CH_2$ has no optical isomers, and is therefore the logical starting point for gaining insight into the binding of amino acids to the diamond surfaces.

Although there is a wealth of experimental and theoretical data available on diamond surfaces [47, 48, 49, 50, 51], there is a lack of data on the interaction of organic molecules with the diamond surfaces. To date, most of the work has concentrated on the interaction of relatively simpler inorganic molecules such as hydrogen, and oxygen[48, 52] on diamond surfaces. Here, we present a systematic study of glycine on the diamond (100) and (111) (2 × 1) surfaces. Our results indicate that: (i) on the C(100) (2 × 1) surface, glycine interacts mainly via its amino group (both with and without the loss of an H-atom), and the hydroxyl group if a loss of an H-atom is involved; and (ii) the C(111) (2 × 1) surface is mostly inert to glycine and presumably other nonpolar amino acids. We note that in many respects, this work is a continuation of our previous investigations of glycine on the Si(100) surface[53].

This section is organized as follows. In Subsection 4.3, we summarize the methodologies used in our simulations. Results for glycine on both diamond surfaces are given in Subsection 4.4, while the Subsection 4.5 is reserved for the summary and conclusions.
4.3 Methodology

The adsorption of glycine on the diamond surfaces was investigated with standard density functional theory (DFT) based simulations, using the Vienna Ab Initio Simulation Package (VASP) with gradient corrected GGA-PW91[12] functionals, and non-normconserving ultrasoft pseudopotentials[54] with a plane-wave cutoff of 30 Ry. The systems were modeled as periodic supercells consisting of seven (eight) atomic layers for the (100) surface ((111) surface) respectively, with the equivalent number of vacuum layers. The bottom layer of the structures was saturated with hydrogen atoms, and kept frozen throughout the simulations. Calculations were performed using the optimized surface structures, which for the C(100) surface consisted of the equivalent of four 2x1 surface dimers. An equivalent size was used for the C(111) calculations. The glycine molecule was optimized separately, and then placed above the diamond surfaces. The entire system was then optimized using the conjugate gradient (CG) algorithm[55], with no constraints. The molecule and surface structures were considered to be equilibrated when the Hellmann-Feynmann forces were less than 0.01eV/Å.

The optimized structures were characterized in several standard ways. Perhaps the most important quantity calculated is the adsorption energy given by:

$$E_{ad} = (E_{mol+surf} - E_{mol} - E_{surf})$$  (4.1)

where $E_{mol+surf}$, $E_{mol}$, and $E_{surf}$ represent the total energy of the optimized glycine and diamond surface structure, the isolated molecule, and the pristine diamond surface, respectively. Further information is obtained from the surface band structure, and simulated scanning tunneling microscopy (STM) images. Brillouin zone integration for the surface band structure is performed using sets corresponding to 64 k-points in the full (1x1) surface Brillouin zone. The STM images were simulated using the standard Tersoff and Hamann formula[56].

To gain insight into the binding of glycine to the diamond surface, the charge
transfer characteristics were quantified in the following manner\[57\]. The spatially resolved charge density difference $\Delta \rho(\mathbf{r})$ is given by:

$$
\Delta \rho(\mathbf{r}) = \rho(\mathbf{r})_{\text{mol+surf}} - \rho(\mathbf{r})_{\text{surf}} - \rho(\mathbf{r})_{\text{mol}},
$$

(4.2)

where $\rho(\mathbf{r})_{\text{mol+surf}}, \rho(\mathbf{r})_{\text{surf}},$ and $\rho(\mathbf{r})_{\text{mol}}$ are the charge densities of optimized product, surface substrate, and adsorbate molecule, respectively. From this, the total charge transferred is obtained using:

$$
|Q^\pm| = \int_{\Delta \rho(\mathbf{r}) \leq 0} d\mathbf{r} \Delta \rho(\mathbf{r}).
$$

(4.3)

This is then used to calculate the length of an \textit{idealized} dipole length ($D_z$), projected onto the surface normal, which was chosen to be the $z$-axis:

$$
d_z = \frac{1}{Q^+} \int_{\Delta \rho(z) > 0} dr z \Delta \rho(z) - \frac{1}{Q^-} \int_{\Delta \rho(z) < 0} dr z \Delta \rho(z),
$$

(4.4)

The normal component of this dipole is calculated as:

$$
P_z = |Q^\pm| d_z.
$$

(4.5)

The normal charge redistribution is given by:

$$
|Q_{\|}^\pm| = \int_{\Delta \bar{\rho}(z) \leq 0} d\mathbf{r} \Delta \bar{\rho}(z),
$$

(4.6)

with the normal charge separation length

$$
d_{\|} = \frac{1}{Q_{\|}^+} \int_{\Delta \bar{\rho}(z) > 0} dr z \Delta \bar{\rho}(z) - \frac{1}{Q_{\|}^-} \int_{\Delta \bar{\rho}(z) < 0} dr z \Delta \bar{\rho}(z),
$$

(4.7)

where

$$
\Delta \bar{\rho}(z) = \frac{1}{A} \int_A dx dy \Delta \rho(\mathbf{r}).
$$

(4.8)

All of the above quantities are tabulated and discussed in the next section.

In our investigations, we found that some configurations lead to the abstraction of a H-atom from the glycine molecule which subsequently absorbs at a different surface
site. In addition, we also find evidence for a cycloaddition reaction on the C(100) surface. A characterization of these chemical reactions involves the calculation of the transition state energies and configurations. While the VASP-based simulations allow one to characterize the chemisorption of glycine on the diamond surfaces, some cases require an additional tool for these kinds of transition state calculations. Hence, as in previous studies [57] we resort to quantum chemistry calculations using the GAUSSIAN code [58] to estimate these quantities. For instance, reactions on the C(100) surface were investigated by means of a $C_{21}H_{20}$ cluster shown in Fig.4.1a. Essentially, the cluster consists of three surface dimers, of which only the middle dimer was left intact for glycine adsorption. The other two dimers were hydrogen-terminated, in order to eliminate all the dangling bonds. In terms of the vertical direction, the cluster is essentially five atomic layers thick, with all vacant bonds hydrogen terminated. Similarly, for reactions on the C(111) surface, a $C_{35}H_{30}$ cluster – essentially consisting of four atomic planes – as shown in Fig.4.1b was used. Glycine was then placed on top of these clusters to form appropriate initial configurations, which were then optimized using DFT with a 4-31G basis with the popular B3LYP[39, 59, 60] gradient-corrected functional for the exchange and correlation energies. More accurate estimates of the transition states and energies were then obtained using the 6-31G* basis set. The transition states were determined by the analytical hessian calculation[61], and is justified, since only one negative eigenvalue was obtained along the initial-to-final state path.

4.4 Results and Discussion

In this section, we discuss the adsorption of a glycine molecule on the diamond (100) and (111) surfaces. Roughly speaking, these fall into two categories: the chemisorption of the intact glycine molecule on the surface, and chemisorption involving the abstraction of a single H-atom from the glycine, which then attaches
4.4.1 Glycine on C(100) (2x1) surface

When chemical vapor deposition (CVD) methods are used, the C(100) surface is diamond’s slowest growing crystallographic surface. In order to reduce the number of dangling bonds, the topmost atoms of a clean C(100) surface reconstruct by means of $\pi$-bonds to form dimers, which arrange themselves into parallel rows, as shown in Fig. 4.2. These dimers are also evident in the simulated STM images. Unlike the Si(100) surface, there is no buckling or tilting of these dimers. Each atom of a C(100) dimer therefore exhibits the same electronic properties. Hence, the HOMO/LUMO is distributed evenly between the dimers. This is in contrast to the buckled Si(100) surface, where the HOMO (LUMO) is associated with the upwardly (downwardly) titled dimer atom, respectively. The structural characteristics obtained for the pristine surface are in good agreement with previously published works\cite{52, 41, 62} with
Figure 4.2: Side and top views of the pristine diamond C(100) (2 × 1) surface showing HOMO (red) and LUMO (blue) states at the Γ point, as well as corresponding (2×4) surface band structure and simulated STM images. Because there is no buckling, the HOMO/LUMO states are distributed uniformly around the surface dimers.
a C-C bond distance of 1.38 Å for any given dimer. Since the interdimer interaction creates a gap between the occupied and unoccupied surface states, the C(100) surface is semiconducting with a measured band gap of 1.0 eV which, as expected, is somewhat less than the experimental estimate of 1.3 eV [63].

Glycine \( (NH_2CH_2COOH) \) is the simplest amino acid, and features active carboxylic acid and amino groups. We started from several initial geometries available in the protein database (PDB)[64], and relaxed these separately using the VASP code. The resulting bond lengths and bond angles were found to be in good agreement with previous calculations, differing by less than 1%. For example, calculated bond lengths are: 1.42, 1.62, 1.21, and 1.36 Å for C-N, C-C, C-O(=) – i.e., with O-atom of carboxylic group “double-bonded” to the carbon atom, and C-O(H) – i.e., with O-atom of carboxylic acid group with a H-atom attached, respectively. The resulting glycine structures were then placed on top of the C(100) surface, which was then allowed to relax and find its optimal structure, which was then characterized.

While many possible binding configurations were examined, only three led to any appreciable binding of glycine without hydrogen abstraction. The absorption energies and charge transfer properties for these configurations are summarized in Table I, and illustrated in Fig.4.3. In terms of a single group, only the amino \(- NH_2 -\) group chemisorbs directly with the C(100) (2x1) surface, with a binding energy of -0.87 eV. Other structural characteristics include a breaking of the C-C dimer bond and resulting in a bond length of 1.56 Å; a surface C-N bond length of 1.54 Å, and an approximately 4% lengthening of the N-C bond length to 1.48 Å within the absorbed glycine molecule. In terms of the charge transfer, there is a shift of charge from the molecule to the surface of \( \approx 1.16e \), with the other C-atom of the dimer \( (i.e., \text{not where the N-atom absorbs}) \) receiving some of this charge. This atom is now associated with the HOMO, while the LUMO continues to be localized around the other surface dimers. This feature is also evident on the simulated STM images. This charge transfer is also reflected in the sign of the vertical component of the surface...
Table 4.1: Absorption energy (eV) and charge transfer characteristics for the
chemisorption of glycine on the C(100) (2 × 1) surface. While the top three items
cover adsorption of an intact glycine molecule, the rest involve hydrogen abstraction.

| Species             | $E_{ads}$ | $|Q^z|$ | $|Q^z|_{\parallel}$ | $d_z$[Å] | $d_{\parallel}$[Å] | $P_z$[D] |
|---------------------|-----------|--------|---------------------|----------|-------------------|----------|
| NH$_2$              | -0.87     | 1.16e  | 0.25e               | -0.11    | -0.51             | -0.61    |
| N+O                 | -1.79     | 2.03e  | 0.40e               | 0.02     | 0.11              | 0.21     |
| C=O                 | -1.68     | 1.05e  | 0.31e               | 0.32     | 1.10              | 1.62     |
| H from (O-H)$_1$    | -2.96     | 1.16e  | 0.35e               | 0.34     | 1.13              | 1.89     |
| H from (O-H)$_2$    | -1.61     | 1.43e  | 0.39e               | 0.31     | 1.16              | 2.15     |
| H from O+O(H)       | -2.32     | 1.72e  | 0.57e               | 0.14     | 0.41              | 1.13     |
| H from N(H$_2$)     | -3.21     | 1.15e  | 0.30e               | 0.24     | 0.92              | 1.34     |

dipole, which is negative. Figure 6.3a also shows the surface band structure, which
remains semiconducting albeit with a smaller bandgap. Essentially, the adsorption
of a glycine molecule destroys the original symmetry of the four dimers and causes
the four bonding and antibonding states to move into the gap region. With the
chemisorption of the amide group to a surface dimer, further attachment via the O-
atom of the carboxylic acid group to the surface dimers becomes feasible, lowering the
adsorption energy significantly to -1.79 eV. This configuration involves an interaction
between the glycine molecule and two neighboring dimers. Essentially, the π-bond
associated with the carbonyl “C=O” species is broken to form two σ bonds with two
adjacent C-atoms on different dimers – i.e., they form a bridge between two different
dimers. As expected, the π-bond nature of the dimer bond is changed considerably,
with an increase in the C-C bond length to 1.63 Å. Surface C-N and C-O bond lengths
are 1.51 and 1.42 Å respectively. Within the glycine molecule, the N-C bond length
is increased by about 9% to 1.55 Å. As may be expected, the additional interaction
of the O-atom with the surface almost doubles the charge transfer to about $\approx 2.0e$.
With the HOMO now shifted to the C-N bond, the sign of projected surface dipole is
reversed. While the basics of the surface band structure resembles that of the previous
case (Fig.4.3a), two antibonding states are now removed since these form bonds with
the carboxylic acid group. There is also evidence for a [2+2] cycloaddition reaction.
Figure 4.3: From top to bottom, we show configurations with HOMO/LUMO, charge transfer, (2x4) surface band structure, and simulated STM images for intact glycine on C(100) (2 × 1) surface: (a) amide *i.e.*, N(H₂) adsorption; (b) N+O absorption; and (c) C=O cycloaddition reaction. Color scheme: HOMO(LUMO) indicated by red(blue) as in Fig.4.1; yellow (violet) for regions of negative (positive) $\delta\rho$ in charge transfer pictures.
Figure 4.4: Plot of reaction energies for C=O cycloaddition reaction on C(100) (2 × 1) as obtained with quantum chemistry calculations. The most relevant parts of the cluster are shown for the initial, transition, and final states. Energies are measured with respect to the total energy of the isolated surface and molecule.

Involving both carbonyl ("C=O") atoms of the glycine molecule. By breaking the \( \pi \) bonds to form two \( \sigma \) bonds, a "C-C-O-C" ring is formed with an adsorption energy of -1.68 eV. Measured bond lengths are: 1.57 Å for the surface C-C bond, and 1.58, 1.51, and 1.46 Å for the C-C, C-O, and O-C bonds of the "C-C-O-C" ring. In terms of the charge transfer, the HOMO is now associated with the surface dimer undergoing the cycloaddition, with a slight tilt towards the surface "C-O" bond. There is a relatively large charge separation with respect to the surface normal, and a significant surface dipole. The band structure reveals a gap of about 1.0 eV, indicating that the surface is semiconducting. The activation energies for this cycloaddition reaction was estimated with quantum chemistry methods. Unfortunately, the transition state for this reaction proved to be very unstable and difficult to locate. Initial attempts
based on Cartesian coordinate input failed to produce a suitable transition state structure. Input therefore had to be created using constrained internal coordinates on a simplified model consisting of the “C=O” parts of glycine and an ethylene molecule, i.e., $\text{H}_2\text{C} = \text{CH}_2$. Specifically, the bond distance between “C=O” and one of the double-bonded C-atoms of the molecule was parameterized. All other coordinates were then allowed to relax. Then, at the point of minimum force, the entire structure was relaxed as to give the optimum geometry characterized by one negative eigenvalue of the second derivative of the energy matrix. For this simplified model, the transition state barrier from reactant to product was found to be 2.8 eV. Having identified a suitable barrier for this simplified model, we were able to identify a similar transition state for the cycloaddition reaction on the C(100) surface using analytical second derivatives of the energy. While Fig.4.4 illustrates the reaction pathway identified, Table II summarizes the activation barriers. As expected, the activation barrier for the reaction is lowered on the C(100) surface to about 1.96 eV. However, this is quite a high barrier, especially when competing H-atom abstraction barriers are taken into account. Hence, one cannot expect the cycloaddition reaction to be favored on the C(100) (2x1) surface.

We now turn to configurations involving the abstraction of a H-atom from the absorbed glycine molecule. Perhaps most significant is the loss of a H-atom from an absorbed amino group, which attaches itself to the same dimer. This is a particularly stable configuration, with an adsorption energy of -3.21 eV. As expected, the surface dimer loses all its $\pi$ bonding characteristics, increasing its length to 1.65 Å. Other bond lengths are indicated on Fig.4.5a. The charge transfer characteristics in many cases resemble that of the intact amide group absorption, already discussed. However, the addition of the H-atom to the surface dimer reverses the sign of the surface dimer and the vertical charge separation. The surface band structure indicates a gap of about 1.0 eV, characteristic of semiconducting behavior.

In contrast to the cycloaddition reaction, the quantum chemistry calculation of the
Table 4.2: Activation energies (eV) for the reactions of glycine on the C(100) (2x1), C(111) (2x1), and Si(100) (2x1) surfaces. Data for the latter is from Ref.[48].

<table>
<thead>
<tr>
<th>Species</th>
<th>C(100)</th>
<th>C(111)</th>
<th>Si(100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C=O</td>
<td>1.961</td>
<td>1.043</td>
<td>0.31</td>
</tr>
<tr>
<td>N-H</td>
<td>0.726</td>
<td>2.066</td>
<td>1.15</td>
</tr>
<tr>
<td>O-H</td>
<td>0.779</td>
<td>1.311</td>
<td>0.13</td>
</tr>
</tbody>
</table>

transition state and the associated barrier turned out to be relatively straightforward. Using the cluster model described, the activation barrier was estimated to be about 0.73 eV, which is considerably lower than the cycloaddition barrier. Hence, the H-atom abstraction from an absorbed amide group is likely to be a significant reaction pathway. Configurations involving the loss of an H-atom from the carboxylic acid group were also found to be stable, with results depending somewhat on where the lost H-atom reabsorbs. For example, the H-atom may attach itself to the same dimer (labeled as $(OH)_1$ in Table I and shown in Fig.4.5b), or on a neighboring dimer within the same dimer row (labeled as $(OH)_2$ in Table I, and shown in Fig.4.5c). The energetics of these possibilities are quite similar, with final adsorption energies of about -1.50 eV. As expected, the charge transfer characteristics for these cases are very similar. To calculate the activation energy for this reaction with quantum chemistry methods, glycine was first trapped to the diamond surface via a dative-bonded state utilizing the O(H) atom. The energy of this initial state was -0.31 eV. Dissociation of the H-atom from this configuration involved an activation of 0.78 eV which is competitive with the barrier associated with amide dissociation. Of course, once the H-atom is removed from the hydroxyl group, there is the possibility of both O-atoms bonding to the 2x1 surface dimers (labeled as O+O(H) in Table I and shown in Fig.4.5e). As may be expected, this configuration is quite stable with a binding energy of -2.32 eV. To summarize, glycine adsorption on the C(100) (2x1) surface takes via the amino group, whose subsequent stability may be further increased through the abstraction of a H-atoms from the chemisorbed group. While
Figure 4.5: From left-to-right: side view of configurations with HOMO/LUMO, charge transfer, structural details focusing on bond lengths, (2x4) surface band structure, and simulated STM images for situations where glycine absorption on C(100) (2 x 1) involves H-atom abstraction: (a) H from N(H$_2$); (b) H from (O-H)$_1$; (c) H from (O-H)$_2$; and (d) H from O+O(H). Color scheme same as in Fig.4.3.
configuration involving the abstraction of an H-atom from the hydroxyl group are quite stable, such processes are probably less likely given the weak initial absorption energies. While there is evidence for a cycloaddition reaction, the high activation energies will most likely preclude this reaction pathway for glycine on C(100) (2x1) surface.

4.4.2 Glycine on C(111)(2×1) surface

The diamond (111) surface is the natural cleavage plane of diamond, whose structure has not been without controversy[52, 42]. Although several different surface structures have been reported, such as the (1×1), (√3 × √3)R30° [52, 65], etc the Pandey[66, 67] (2×1) π-chain reconstruction is now largely accepted as being the ground state structure[41, 52]. As shown in Fig. 4.6, under the (2×1) reconstruction, the surface atoms couple through π-bonds rearranging themselves into parallel zigzag chains. Zigzag chains are also characteristic of the subsurface atoms. Each surface atom has a three-fold coordination, sharing a delocalized network of π electrons along the chain. This is attributed to the strong intra-atomic Coulomb repulsion of the carbon atoms, which prevents buckling and dimerization of the surface atoms. Currently, there is some debate over the finer details of this reconstruction and the electronic properties of this surface. Early theoretical studies predicted a (2×1) structure with no buckling and no dimerization of the chains, along with a metallic surface band structure[67]. More recent DFT-based studies reported very small dimerization, which leads to a small-gap in the surface band structure[43, 52]. However, experimental work reported a sizable surface band gap, indicating either a possible tilting or dimerization of the chains[68, 69]. The most notable difference between the theoretical and experimental work is the measurement of a 1.0 eV surface band gap via angle-resolved photoemission spectra[63]. A possible resolution of this issue was recently presented by M. Marsili and coworkers, who found that the DFT-based surface band structure may be corrected via the inclusion of many-body effects, rather than
Table 4.3: Adsorption energies (eV) and charge transfer characteristics for glycine on C(111) (2 × 1) surface. Except for the cycloaddition reaction, they all involve abstraction of a H-atom from the glycine molecule.

| Model          | $E_{ads}$ | $|Q|_{\pi}$ | $|Q|_{\|}$ | $d_{\pi}$[Å] | $d_{\|}$[Å] | $P_z$[D] |
|----------------|-----------|-------------|------------|--------------|--------------|----------|
| C=O            | 0.58      | 1.11e       | 0.28e      | 0.09         | 0.35         | 0.48     |
| H from O(H)    | -0.33     | 1.13e       | 0.30e      | 0.39         | 1.45         | 2.12     |
| H from N(H$_2$)| -0.27     | 1.16e       | 0.29e      | 0.22         | 0.88         | 1.22     |

by a buckling or dimerization of the π-chains[70]. In this work, we have investigated glycine on the C(111) (2 × 1) surface, in the context of GGA-based functionals. In agreement with previous work[62, 52, 49], we find no evidence for either buckling or dimerization within the π-chains, as shown in Fig.4.6. Essentially, we find that the π-chains are highly uniform, with no preferred spot for the interaction with a glycine molecule. In fact, we find that the intact glycine molecule does not “stick” to the surface at all. Even the cycloaddition reaction with the carbonyl group of the glycine is not energetically favored, with the “product” (shown in Fig.4.7a) having a higher energy than the separated glycine and pristine C(111) (2 × 1) surface. Both the VASP and quantum chemistry calculations give qualitatively similar results for this calculation. The only configurations in which glycine chemisorbed on the surface involved the abstraction, and reattachment, of an H-atom from either the hydroxyl or amino groups, similar to what was found for the C(100) (2 × 1) surface. For completeness, structural and electronic details along with simulated STM images for these cases are given in Table III, and shown in Fig.4.7(b,c). In both cases (i.e, N(H$_2$) or O(H) losing a H-atom that is reabsorbed), the binding is quite weak – only about -0.30 eV (Table III). As expected, there is a small lengthening of the C-C surface bonds, where the glycine and the hydrogen absorb. In terms of the bandstructure (not shown), the DFT calculations predict the opening up of a very small gap.

Given these very small binding energies, the activation energies for the H-atom abstraction should be quite small if glycine is to interact at all with this surface.
Figure 4.6: Side and top views of the pristine C(111) (2 × 1) surface showing HOMO (red) and LUMO (blue) states at the Γ point, as well as the corresponding (2x1) surface band structure and STM images.
Figure 4.7: From left-to-right: side view of configurations with HOMO/LUMO, charge transfer, structural details, and simulated STM images for situations where glycine adsorption on C(111) (2×1) involves either (a) cycloaddition with C=O; or H-atom abstraction with (b) H from N(H₂); and (c) H from O(H). Color scheme is same as in Fig.4.3.
However, our quantum chemistry estimates, given in Table II, show that this is not the case. Specifically, activation barriers of 1.31 and 2.07 eV were found for the case of O(H) and NH$_2$, respectively. In summary, we therefore conclude that the C(111) (2 × 1) surface should be quite inert to interactions with glycine, and presumably, other neutral amino acids.

### 4.5 Conclusions and Summary

In summary, we have examined the issue of glycine adsorption on diamond (100) and (111) (2 × 1) surfaces with DFT-based calculations. Because of the anisotropic nature of both the glycine molecule and the surfaces, a large number of configurations were examined. Roughly speaking, the cases for which any significant adsorption takes place may be categorized in terms of either involving or not involving the abstraction of a H-atom from either the carboxylic acid or amino group of the glycine molecule.

With regards to the C(111) surface, we find no evidence of any binding when the intact glycine molecule is involved. The cycloaddition reaction is also precluded on energetic grounds. There is evidence of weak binding of the molecule to the surface when the loss of an H-atom is involved. However, quantum chemistry based estimates of the transition states show that these involve activation energies that are much higher than the binding energies. We can therefore expect that the C(111) (2 × 1) surface to be inert with respect to interactions with glycine.

By contrast, glycine interacts much more strongly with the C(100) surface. On this surface, binding takes place primarily via the adsorption of the amino group. The stability of such configurations is considerably enhanced through the abstraction of an H-atom, with an activation energy of about 0.73 eV. Configurations involving the abstraction of a H-atom from the carboxylic acid group were also found to be quite stable, with a competing activation barriers. However, a much lower initial adsorption energy means that this pathway is less likely than the amide group absorption. There
is also evidence for a cycloaddition reaction on this surface, but is unlikely to be competitive because of the high activation barrier associated with this reaction.

It is interesting to compare these last results to those obtained for glycine adsorption on the Si(100) c(2x4) surface[53]. In contrast to the C(100) surface, the Si(100) surface dimers are buckled, with the HOMO (LUMO) located around the “up” (“down”) atom of the surface dimer, respectively. For the most part, the “up” atom of the Si-dimer is inert to glycine, so that the “down” atom is the most important in the absorption process. In contrast to the C(100) surface, both carboxylic and amino groups may chemisorb on the Si(100) with bonding to the latter group being somewhat stronger (-1.46 versus -0.86 eV, respectively). Once absorbed, either of the two species may loose a H-atom, as already described. However, a low activation energy of 0.13 eV for O(H) versus a barrier of 1.15 eV for N(H$_2$) overwhelmingly favors the former reaction. There is also numerical evidence for a cycloaddition reaction involving the “C=O” species. Since the estimated activation barriers for such a reaction is about 1.22 eV, such reactions are most likely excluded.

Having studied the absorption of glycine on two of the most prominent diamond surfaces, we aim to extend this work by investigating other amino acids – especially those involving polar and charged groups – in the near term future.
Chapter 5

Capacitance of Al nanowire

5.1 Abstract

Classical notions of capacitance need modification at the nanometer length scale in order to account for important quantum effects. In this chapter, we present the results of *ab initio* investigation of the capacitance of a metal nanowire. The systems investigated feature two semi-infinite Al nanowires in close proximity to each other. We considered three different cross sectional areas for the wires: Al(100)(3 × 3), Al(100)(5 × 5), and Al(100)(7 × 7). First principles estimates of capacitance matrix coefficients for the systems are provided.

5.2 Introduction and Methodology

Theory of quantum capacitance is developed over a decade ago [71]. However, the recent progress in molecular electronics has given new urgency towards understanding the capacitance at this extreme small scale. At nanometer scale, concept of capacitance must be reformulated in a similar spirit as the conductance. The classical
capacitance coefficients $C_{\alpha\beta}$ of a set of conductors are defined as:

$$Q_{\alpha} = \sum_{\beta} C_{\alpha\beta} V_{\beta},$$

(5.1)

which gives accumulation of charge $Q_{\alpha}$ on conductor $\alpha$ in response to a electrostatic potential $V_{\beta}$ on conductor $\beta$. Central to the classical notion of capacitance is that of a set of well defined conductors, with zero electric field in their bulk and surfaces with equal potentials. These assumptions break down at the nanometer length scale, when the screening length of the material becomes comparable to the dimensions of the system. Therefore, the notion of capacitance has to be modified to that of the electrochemical capacitance, where each conductor is connected to an electron reservoir with an electrochemical potential $\mu$ [72]. Then, charge variation $dQ_{\alpha}$ on conductor $\alpha$ has to be obtained self consistently with the change of electrochemical potential $d\mu_{\beta}$ of the reservoir connected to the conductor $\beta$. The electrochemical capacitance is defined by following equation [71]:

$$dQ_{\alpha} = \sum_{\beta} C_{\alpha\beta} \frac{d\mu_{\beta}}{e} + \sum_{\beta\gamma} C_{\alpha\beta\gamma} \frac{d\mu_{\beta}}{e} \frac{d\mu_{\gamma}}{e} + \ldots$$

(5.2)

It is obvious from the Eq. (5.2) that the charge accumulation is, in general, nonlinear at a given bias voltage. The $C_{\alpha\beta}$ and $C_{\alpha\beta\gamma}$ represent the first two orders of the nonlinear capacitance coefficients. These coefficients also take into account quantum effects due to the screening, and may differ from their classical counterparts [73, 74]. For example, the linear response formula for parallel plate conductors at microscopic scale is given by Büttiker formula [31]:

$$4\pi C = \frac{A}{d + 2\lambda},$$

(5.3)

where $\lambda^{-1} = 4\pi e^2 (d\sigma/dE)$ is the screening length, and $(d\sigma/dE)$ is the density of states per unit area of the leads.

In this work, we studied capacitance of Al nanowires with various sizes. Illustrated in Fig. 5.1 is the schematics of the system. The system consist of left and right leads
Figure 5.1: Schematic display of Al nanowire as a capacitor.

made of Al nanowires, which are connected further to their electron reservoirs with electrochemical potentials $\mu_L$ and $\mu_R$. The device is surrounded by a metal box. The metal box is an important component in molecular electronic devices, which also allows us to account all the electric field lines in numerical simulations. We used the Dirichlet boundary conditions for the electrostatic potential to simulate the metal box surrounding the entire system. The equivalent circuit of the device consists of three capacitors as shown in Fig. 5.1.

In contrast to a two probe transport device, it is the both electrostatic potential and the self-consistent charge density that are the two most important quantities for the calculation of the quantum capacitance. Calculation of the electrostatic capacitance is carried out by the standard NEGF technique, while the calculation of the charge density is based on the following considerations. The electronic states of the atoms are modeled using the LCAO with $\{\phi_\nu\} = \{\phi_s, \phi_{px}, \phi_{py}, \phi_{pz}\}$, and standard pseudopotentials. The KS eigenstates of the system, with a shifted Hartree potential
at the boundary between the leads and the central region are then expanded in terms of following basis:
\[ \Psi^i = \sum_\nu c^i_\nu \phi_\nu(r - R_I)(\nu = s, p_x, p_y, p_z), \] (5.4)
and \( I \) is an atomic index. The KS equation may then be transformed into a standard matrix problem [28]:
\[ H_{\mu\nu}c^i_\nu = E^i S_{\mu\nu}c^i_\nu, \] (5.5)
with \( H_{\mu\nu} \) and \( S_{\mu\nu} \) representing the Hamiltonian and overlap matrix elements between the orbitals located at different sites. Then the real-space charge density \( \rho(r) \) is calculated:
\[ \rho(r) = \sum_{\mu\nu} \phi^*_\mu(r - R_I)\rho_{\mu\nu} \phi_\mu(r - R_I), \] (5.6)
with the density matrix \( \hat{\rho} \) given by Eq. (3.15). After computing the charge density, it is straightforward to calculate the linear capacitance coefficients using the definition:
\[ C_{\alpha\beta} = e \frac{dQ_\alpha}{d\mu_\beta}, \] (5.7)
by applying a change in voltage \( \Delta V = d\mu/e \) to a given reservoir and then measuring the charge difference \( \Delta Q = Q(V + \Delta V) - Q(V) \).

Since we are dealing with a two conductor system, the main quantities of interest are the self and cross capacitance coefficients:
\[ \Delta Q_1 = C_{11}\Delta V_1 + C_{12}\Delta V_2, \]
\[ \Delta Q_2 = C_{21}\Delta V_1 + C_{22}\Delta V_2, \] (5.8)

5.3 Results and Discussion

Important feature of any capacitor is an ability to store electric charge. How well the capacitor performs as a charge storage device depends on the response of the electrodes to their counterpart. In Fig. 5.2, charge accumulations on atomic planes
Figure 5.2: Charges on Al nanowire capacitors at different bias voltages.

of two opposing leads is shown at separation distance $d = 8.1\,\AA$, and two different bias voltages $V_{\text{bias}} = -0.02\,\text{eV}$ and $V_{\text{bias}} = 0.02\,\text{eV}$ for Al(100)(7x7) leads. The bias voltage is applied on the right lead. The self and cross capacitance coefficients at $V_{\text{bias}} = 0.02\,\text{eV}$ were found to be:

$$
C_{11} = 0.088aF \quad C_{12} = -0.035aF, \\
C_{21} = -0.035aF \quad C_{22} = 0.089aF
$$

The most striking feature of the system is that the response charge at the left lead is considerably different from the charge at the right lead. This is due to the very small size of the leads and the effect of the surrounding metal box as an additional capacitor. Further, we analyze this response with the cross capacitance coefficient
Figure 5.3: Cross capacitance coefficient $C_{12}$ as function of separation distance between the capacitors.

as a function of separation distance (Fig. 5.3). The cross capacitance coefficient $C_{12}$ decreases nonlinearly as the separation distance increase. However, the bigger leads have higher cross coefficients. We checked the linear response formula (Eq. (5.3)) for parallel capacitors in the same manner. The inset in Fig. 5.3 shows $1/C_{12}$ depending on the separation distance nonlinearly. This clearly contradicts with the Büttiker linear response formula, which predicts a linear dependence. The reason for this discrepancy, again, is attributed to the small size of the Al nanowire leads and the effect of the surrounding metal box. Further, we analyze charge ratio on the two leads as function of the separation distance (Fig. (5.4)). This plot shows that the response charge as decreases nonlinearly as the separation distance increase, while the bigger leads have better response.
Figure 5.4: Percent induced charge $\frac{\Delta Q_{left}}{\Delta Q_{right}}$.
5.4 Summary

To summarize the result, nanoscale wires make rather poor capacitors due to their small size. In reality, the system considered here resembles more of tips of two needles, rather than parallel plates. Also, the effects of surrounding metal box comes into play. Circuit elements in actual nanoscale devices will inevitably be near other conducting elements, or surrounded by a metal box. Therefore, effects from these elements as additional capacitors contributes to the overall performance of nanoscale capacitors.
Chapter 6

*Ab initio* band bending, metal-induced gap states, and Schottky barriers of a carbon and boron nitride nanotube devices

6.1 Abstract

We have characterized the fully self-consistent electronic properties of a prototypical metal/nanotube interface using a combined nonequilibrium Greens function and density functional theory based formalism, under different conditions of gate and bias voltages. Both carbon and boron nitride nanotubes between Al electrodes, were considered. The electronic properties of the interface are dominated both by a dipole and by metal induced gap states (MIGS) formed through the transfer of charge between the metal and the nanotube. In addition, first principles estimates – within the local
6.2 Introduction

Depending on their helicity, CNTs may be either metals or semiconductors. These electronic properties, when combined with their exceptional strength and stability, make CNTs an ideal material for exploring quantum transport at the nanometer scale[75]. Already, a variety of nanotube-based devices have been produced in the laboratory and their workings explored. In particular, semiconducting CNTs show considerable promise as field-effect transistors, with better device properties than silicon-based metal-oxide-semiconductor field-effect transistors (MOSFETs)[76, 77]. A fundamental aspect of these, and related devices, is the junction between the metal leads and the semiconducting nanotube. While it is understood that metal-semiconductor junctions are characterized by bandbending, Schottky barriers, and Fermi level pinning[78, 79], the microscopic aspects of these phenomena have remained controversial. In particular, Fermi level pinning – the relative insensitivity of the Schottky barrier heights to changes in the work function – have been ascribed to metal-induced gap states[80] (MIGS) and polarization phenomena[81]. Given the central importance of these issues for nanotube devices, the nature of the metal-semiconductor interface has recently received considerable attention, both theoretically[82, 83, 84, 85] and experimentally[86, 87, 88, 89]. In particular, work has shown that the one-dimensional nature of the nanotubes implies reduced Coulombic screening, so that CNT-based devices are expected to be fundamentally different from conventional devices, while experimentally, several groups have reported on Ohmic-like and Schottky barrier nanotube/metal contacts[86, 87, 88, 89].

In this work, we report on a fully self-consistent first principles investigation, within the context of the local density approximation, of the transport characteristics of a prototypical CNT and boron nitride nanotube (BNNT) based device with Al density approximation – of the Schottky barrier heights are given.
electrodes. In particular, we discuss the bandbending and charge transfer under different conditions of bias and gate voltages. There is evidence of Fermi level pinning based on MIGS, with estimated Schottky barrier heights of 0.60 eV (0.89 eV) for the CNT (BNNT) devices.

6.3 Methodology

To calculate the fully self-consistent transport characteristics of the nanotube devices, use was made of a recently developed real-space, nonequilibrium Green’s function (NEGF) formalism[90, 28] combined with density functional theory (DFT) based simulations, as described in the literature[91, 92] The advantages of this approach include: (i) a proper treatment of the open boundary conditions for a device under a bias voltage $V_b$; (ii) a fully atomistic treatment of the electrodes; (iii) a self-consistent treatment within the DFT framework of the charge density via the NEGF formalism thereby incorporating both the effects of scattering and bound states.

Our investigations focused primarily on a CNT and a BNNT directly coupled to Al-leads, as shown in Fig.6.1 (for couplings to nanotube sides, see Ref.[10]). The typical two-probe device consists of $5 \times 5$ (100) Al slabs with 50 atoms per unit cell in the lead, and 32 atoms per N for the semiconducting (8,0) nanotubes. We were able to investigate systems with N+2 up to 18, so that the largest systems considered was made up of 200 Al and 576 device atoms. While such systems are relatively large for self-consistent DFT transport calculations, it is important that these be large enough so that the center of the nanotube devices has the electronic characteristics of a very long (8,0) nanotube. This has been explicitly verified through calculations of the “renormalized” bandstructures of the system[93], which is also important for the bandbending calculations. If one is not directly adjacent to the leads, the system should be approximately translationally invariant. Hence, the electronic states inside the nanotube device should be similar to the states of an infinitely long nanotube
Figure 6.1: Typical two-probe (8,0) CNT device (length N+2 units) sandwiched between Al (100) leads \((L, R)\). Note that the central region \(C\) also contains a considerable portion of the electrodes. The whole system is surrounded by a series of gates (shown here as one), which set \(V_G\) in different parts of the device.

with a similar charge distribution. Given a self-consistent Hamiltonian matrix of the device, the submatrix corresponding to different unit cells \(N\) of the central nanotube region was carved out, and the bandstructures calculated by standard means. These “renormalized” bandstructures reflect the influence of the leads and the externally applied voltages. Generally speaking, there is excellent agreement, within a constant energy shift, between the bandstructure of an infinite nanotube and that of a unit cell from the two-probe device. However, only when \(N \approx 16\), or larger, are the two entirely coincident \(\text{without any energy shifts}\), showing that true bulk device properties are obtained. The distance between the last plane of the electrode and CNT (BNNT) was chosen to be 1.6 (1.44) Å, in order to achieve a reasonably strongly coupled contact. The electronic states of the atoms were modeled using a linear combination of atomic orbitals (LCAO)[94], with \(\{s,p\}\) orbitals and standard pseudopotentials[95]. As a further feature, the entire system is surrounded by a metal container, which serves not only to terminate any electric field lines, but also allows us to subject the device to a gate voltage \(V_G\)[96].

The device current is calculated with the Landauer formula[28]:

\[
I = \frac{2e}{h} \int_{\mu_{\text{min}}}^{\mu_{\text{max}}} dE (f_L - f_R) T(E, V_b, V_G),
\]

where the transmission probability \(T(E, V_b, V_G)\) as a function of the electron energy \(E\), the bias voltage \(V_b\), and \(V_G\) is given by:

\[
T(E, V_b, V_G) = \frac{\left| \sum_I \langle \psi_I^b | \psi_I^L \rangle \langle \psi_I^L | \psi_I^R \rangle \langle \psi_I^R | \psi_I^G \rangle \right|^2}{\left( \sum_I \langle \psi_I^b | \psi_I^L \rangle \langle \psi_I^L | \psi_I^R \rangle \langle \psi_I^R | \psi_I^G \rangle \right)^2}.
\]
$4Tr[Im(\Sigma^L G^R \Sigma^R G^A)]$, with $\Sigma^{L,R}$ representing the self-energies of the left and right leads, respectively; $G^{R(A)}$ the retarded (advanced) Green’s functions; and $f_{L,R}$ the corresponding electron distribution function of the electron eigenstates in the leads. Moreover, $\mu_{\text{min}} = \min(\mu + V_b, \mu)$ ($\mu_{\text{max}} = \max(\mu + V_b, \mu)$) denotes the minimum (maximum) electrochemical potential $\mu$ of the leads. At equilibrium, the current is proportional to the conductance $G$, which is evaluated at the Fermi level of the device such that $G(\mu) = G_0 T(\mu)$ with $G_0 = (2e^2/h) \approx (12.9 \Omega)^{-1}$. All these quantities were evaluated by standard means[91, 92].

### 6.4 Results and discussion

The MIGS are a central component of the nanotube/electrode interface and are probed here by means of $T(E)$ and the local density of states (LDOS). Figure 6.2a shows the I-V characteristics of three CNT devices of increasing length: short nanotubes up to $N \approx 10$ display metallic behavior, while longer devices acquire semiconductor characteristics. To gain insight into this, Figs.2(b-d) plots $T(E)$ for tubes of different lengths. While this transmission spectra is a complicated function reflecting both the bandstructure of the leads and the nanotube levels that mediate the transmission[93], it is evident that $T(E)$ develops a gap around the Fermi level as $N$ increases. Very long, finite-sized tubes are characterized by sharp resonant peaks (Fig.6.2d), which merge in the infinite-size limit to form the continuous band structure of the nanotubes[97]. Transmission through the short nanotubes is attributed to *evanescent modes* with long decay lengths, which may be understood in terms of the complex bandstructure of the nanotubes[98, 93]. These evanescent modes acquire a charge from the Al-leads thereby forming the MIGS[80], whose presence is signaled by a finite LDOS (Fig.6.3) and nonzero $T(E)$ inside the gap regions of the nanotube (Fig.4.2(b,c)). For CNTs, most of the contribution derives from a single peak near the center of the gap, which reaches deep into the device. In the pres-
ence of a positive (negative) $V_G$, the peak height increases (decreases) somewhat. In contrast, the BNNT LDOS spectra is relatively broad and featureless, showing rapid decay away from the junction. The evanescent modes of the nanotubes are given by the imaginary branch of the complex band structure shown in Fig. 6.2(e,f): each of the subbands forms a loop that “connects” the valence and conduction bands of the real bandstructure. Most of the contribution to the transmission comes from the evanescent modes near the Fermi level, which is associated with the largest imaginary part of the most narrow loops. For the CNT, $k\Delta \approx 0.5 - 0.75$, where $\Delta$ is the length of the unit cell. The decay lengths for these modes may be estimated to be $\lambda \sim 2\pi/k \approx 6 - 10$ cells, long, in agreement with the LDOS data (Fig. 6.3). The exact same qualitative features hold true for BNNT-based devices, with the most substantive difference resulting from the shape of the complex bandstructure (Fig. 6.2f). For BNNT, $Im(k\Delta) \approx 1.0 - 1.5$ over much of the gap region, so that the estimated decay lengths of the evanescent modes are less than half the CNT values, in complete agreement with the BNNT LDOS data. The electrode/nanotube junction is characterized by bandbending and charge transfer. However, the planes immediately adjacent to the junction are dominated by an interface dipole (Fig. 6.5b). While this dipole extends over the entire $N=0$ cell, most of its effects are confined to two layers on either side of the interface. To quantify the bandbending of the devices, the renormalized bandstructure was calculated for each unit cell, as a function of the device length and externally applied voltages. As a measure of the bandbending, we tracked the position of the conduction band (valence band) minima (maxima) across the device, when the bending is upwards (downwards), as shown in Fig. 6.4. The upwards (downwards) bandbending is usually associated with n-doped (p-doped) systems, which here may be controlled by means of the $V_G$. Figure 4 shows that for CNTs with $V_b = 0$, the bandbending is symmetric and extends over the entire nanotube. This is not the case for $V_G = 1.36$ eV, for which the bandbending is characteristic only of the first few cells adjacent to the electrodes; the interior cells show almost no bandbending. In all
Figure 6.2: (a) Sample I-V curves for CNTs with $V_G = 0.0$. Here circles, squares and triangles represent data for tubes of length 3, 8 and 18 units, respectively; (b),(c) and (d) show corresponding $T(E)$, with evanescent mode contributions shaded. Here (d) also shows $T(E)$ for an infinite (8,0) CNT (step-like function); (e) complex bandstructure of infinite (8,0) CNT; and (f) (8,0) BNNT.
Figure 6.3: LDOS of different parts of the CNT (left panels) and BNNT (right panels) devices \((V_G = 0)\). The estimated gap states are shaded and the Fermi energy \(E_F\) of the leads are marked with a vertical line. Data shown is for nanotubes consisting of 18 unit cells.
Figure 6.4: Ab initio band bending and corresponding excess charge per atom for each of the CNT cells. Here, circles, squares and triangles represent data for tubes with 6, 12, and 18 cells, respectively. For top three panels $V_b = 0$; for bottom two, $V_G = 0$. 
cases, the bending of the smaller-length tubes follows the trends of the longer devices, albeit with a reduced bandbending. Asymmetric bandbending is a feature of devices with nonzero bias voltages, as shown in Fig.6.4(d,e). Here, the bias potential of the left-lead was fixed at zero, while that of the right-lead was shifted to $V_b$. Accordingly, under a small positive (negative) $V_b$, the bandbending of the cells closest to the right lead increases (decreases) somewhat. When the $V_b$ is substantial (Fig.6.4(d,e)), the distortion is significant and extends over much of the device. Figure 6.4 also plots the excess charge transferred from the leads to the nanotubes. For the most part, the charge transferred follows the trends set by the bandbending curves and the LDOS. All of these features are also evident in the BNNT devices. However, the bandbending and charge transfer takes place from an even fewer number of cells away from the leads, as may be expected from the complex bandstructure picture.

While there are several definition of the Schottky barrier heights ($\phi_b$), we have opted to use $\phi_b = E_C - E_f$ (for n-type materials), where $E_C$ represents the conduction band edge as extrapolated all the way back to the junction. The $\phi_b$ is essentially independent of $V_G$ and is estimated to be 0.60 eV for the CNT and 0.89 eV for the BNNT. The nanotube devices also display some measure of Fermi level pinning, based on the MIGS model. Fermi level pinning refers to the fact that for many systems, changes in $\phi_b$ are small, even though changes in $E_F$ may be substantial. Experimentally, Fermi level pinning is usually probed by changing the metal used for the electrodes. Here, we have checked this explicitly by means of altering gate voltage around the leads only. This alters $E_F$ in an approximately linear fashion, as in several recent experiments[88]. In agreement with other simulations[84], we find that changes in $\phi_b$ are roughly comparable to the Schottky barrier height (i.e., $\approx 0.3$ eV), even though changes in $E_F$ are much larger. These changes are due to the filling of the MIGS of the device, and was explicitly checked by examining the LDOS. The results of these first principles calculations provide justification and insight, at least qualitatively, for models of the workings of nanotube-based diodes[82, 88]. It is clear
that the electronic properties of the Al/nanotube interface are dominated by MIGS. Here, the CNTs and BNNTs provide an example of contrasting materials. From the complex bandstructure, it is clear that the MIGS associated with the CNT reach deep into the device, with a relatively long decay length. By contrast, the MIGS associated with the BNNTs decay rather rapidly. In that sense, the BNNTs act much more like a conventional “planar” device.

### 6.5 Summary

We have examined the transport properties of a semiconducting nanotube (both carbon and boron nitride tubes were considered) between two metal (Al(100)) electrodes, using the self-consistent NEGF-DFT formalism. Focusing on the metal/nanotube interface, the electronic properties are dominated by charge transfer giving rise to significant MIGS. These MIGS give rise to significant transmission for finite-sized tubes, and Fermi level pinning. In addition, first principles estimates for the bandbending
and Schottky barrier heights of the devices, as a function of gate and bias voltages, are
given. Currently, these investigations are being extended to doped nanotube systems
with other metal electrodes[99].

We would like to thank Hong Guo and Jerry Tersoff for useful discussions. We
gratefully acknowledge financial support from NSF Grant No. 0312105 and U.S.
DOE.
Chapter 7

Summary and future work

This thesis presents the results of studies of three different materials systems, each of which may have possible application towards future molecular electronic devices.

First, we have examined the issue of glycine adsorption on diamond (100) and (111) (2×1) surfaces with DFT-based calculations. Because of the anisotropic nature of both the glycine molecule and the surfaces, a large number of configurations were examined. With regards to the C(111) surface, we find no evidence of any binding when the intact glycine molecule is involved. Quantum chemistry based estimates of the transition states show that these involve activation energies that are much higher than the binding energies. We can therefore expect that the C(111) (2×1) surface to be inert with respect to interactions with glycine. By contrast, glycine interacts much more strongly with the C(100) surface. On this surface, binding takes place primarily via the adsorption of the amino group. The stability of such configurations is considerably enhanced through the abstraction of an H-atom, with an activation energy of about 0.73 eV. Configurations involving the abstraction of a H-atom from the carboxylic acid group were also found to be quite stable, with a competing activation barriers. However, a much lower initial adsorption energy means that this pathway is less likely than the amide group absorption. There is also evidence for a cycloaddition reaction on this surface, but is unlikely to be competitive
because of the high activation barrier associated with this reaction.

Having studied the absorption of glycine on two of the most prominent diamond surfaces, we aim to extend this work by investigating other amino acids – especially those involving polar and charged groups – in the near term future.

Second, capacitance of a semi-infinite metal nanowires is studied using the NEGF formulism. The systems investigated feature two semi-infinite Al nanowires in close proximity to each other at their end. Three different cross sectional areas for the wires are considered: Al(100)(3 × 3), Al(100)(5 × 5), and Al(100)(7 × 7). First principles studies show weak response, and nonlinear dependence on the separation distance of cross capacitance coefficients for the systems. We argue that this is due to the extreme small area of the electrodes and the effects of the metal box surrounding the device as an additional capacitor. As a result, the Al and probably most of the metal nanowires make rather poor capacitors.

Last, we have examined the transport properties of a semiconducting nanotube (both carbon and boron nitride tubes were considered) between two metal (Al(100)) electrodes, using the self-consistent NEGF-DFT formalism. Focusing on the metal/nanotube interface, the electronic properties are dominated by charge transfer giving rise to significant MIGS. These MIGS give rise to significant transmission for finite-sized tubes, and Fermi level pinning. In addition, first principles estimates for the bandbending and Schottky barrier heights of the devices, as a function of gate and bias voltages, are given. Currently, these investigations are being extended to doped nanotube systems with other metal electrodes[99].

This work focused on two type of systems, with possible applications in molecular electronics. In the remaining few paragraphs, we briefly discuss possibilities to extent this work.

The opportunities to make further advances in the area of nanotechnology is vast. One example is the so called molecular computer, in which switches and memory components are constructed by molecules. Today, nanometer scale control in the lateral
direction is possible, using the so called radical-initiated hydrosilylation of alkenes, to form self limiting islands or lines of organic material approximately 10-100 nm in size [100]. Therefore, future investigations should include complex pattern formation of organic molecules on the semiconductor surfaces. In this context, one should consider new species of organic molecules and investigating their transport characteristics. The studies being presented should also be extended to films of molecules on semiconductor surfaces with nanometer scale, i.e. a single molecule thickness.

Another possibility is the molecular recognition and chemical and biological sensing. Organic molecules are the key components in this type of devices. In such devices, the organic layer provides the molecular recognition function, while the semiconductor provides signal processing, data storage, and logic capabilities. Key to such device is the attachment of organic layer on the semiconductor substrate, where the organic layer can be engineered with a variety of end groups designed for molecular recognition. In this context, future works should address surface interactions of organic molecules with compound substrate consisting of semiconductor and organic layers.

In this work, we investigated relatively simple model system for nanoscale capacitance. More complex systems, such as Al electrode and capped CNT, nested nanotubes, and nanotube junctions with conductance gap, have been studied recently [96]. The natural extension of this type of research is the AC response of molecular scale circuits. For instance, the interesting problem is to understand capacitive and inductive response of molecular scale devices.

The investigation of two probe electronic transport devices can be extended to a wide variety of future studies. Key to the performance of semiconductor based devices is the control over the Schottky barrier. As we have shown in this work, this is no different in the case of nanotube based devices. For macroscopic scale interfaces, nature of the Schottky barrier, and its dependence on crystal orientations, and impurity level is still debated. CNTs represent much simpler system, thus, better control over such parameters. Therefore, study of the doped nanotubes is the next logical step. Doping
can be on the surface of the nanotubes, or at the center, for example. Stability and simplicity of CNTs provide possibilities for potential new devices. One can consider, for example, a precise super lattice made of periodically doped nanotubes. Our study focused on the model system which is in vacuum. It is quite obvious that future molecular electronic circuits will be integrated in an ambient media. This will create additional interfaces to interact, and the electronic structure of the system will be affected. Therefore, one needs to address this problem. Furthermore, more complex systems such as multi-walled nanotubes with specific doping configurations. Ultimately, the study should be extended for accounting the electron phonon interaction, i.e. temperature, spin, magnetic field, and time dependence.
Bibliography


[42]


