Abstract

WANG, SHUCHUN. Atomic Structure, Optical Properties and Electron Transport in Self-Assembled Monolayers on Surfaces. (Under the direction of Dr. Jerry Bernholc and Dr. Wenchang Lu.)

Adsorbate-induced modification of semiconductor or metal surfaces creates a nano-scale quantum structure which offers a rich vein of exotic physical phenomena for investigation. Human desire to harness these properties for technological or scientific purposes has led to extensive experimental and theoretical investigations. This dissertation focuses on the ab initio simulations of atomic, electronic, optical, and transport properties of nano-scale systems. The calculated results for indium nanowires on the Si(111) surface identify their atomic structure and reveal a phase transition at low temperature. Transport simulations on the self-assembled monolayer of ferrocenyl-alkanethiolate on Au(111) surface show negative difference resistance, which is in very good agreement with experimental observations. This opens a new opportunity for applications in nanoscale molecular devices.

The above large-scale simulations have been carried out within density functional theory. Pseudopotentials were used to describe the ion-electron interaction. In our real space multigrid method, the Kohn-Sham equations are solved in real space numerically by using multigrid accelerations. The ultrasoft pseudopotentials are also implemented. A double grid technique is adopted to perform products of two functions with different smoothnesses. Test on diatomic molecules, bulk gold and bulk copper show that converged results are obtained by using energy cutoffs of 25 Ry. The

relaxed bond lengths of O_2, CO and N_2 are in excellent agreement with the experimental results. The calculated ground state properties, i.e., the equilibrium lattice constant and bulk modulus of bulk gold and copper show very good agreements with the experimental measurements and equivalent plane-wave calculations.

For the indium nanowire on Si(111) surface, the atomic structures are identified uniquely from the reflectance anisotropy spectra (RAS) calculations. Our calculated RAS for (4×1) reconstruction shows a very pronounced optical anisotropy around 2 eV. Meanwhile, the (4×2)/(8×2) reconstructed surface, induced by a slight distortion of the indium chains, is shown to result in a splitting of the 2 eV anisotropy peak. These results are in excellent agreement with recent polarized reflectance data acquired during the (4×1)→(4×2)/(8×2) phase transition.

For the self-assembled monolayer of ferrocenyl-alkanethiolate on Au(111) surface, non-equilibrium Green's function calculations of the I-V characteristics are performed self-consistently with the optimized geometry. Our calculated results show a strong NDR feature at a large bias, which is in good agreement with the experimental results. The mechanism of NDR has been identified by analyzing the voltage-dependent transmission, potential profile along the junction, and molecular level alignment under bias. We find that the ferrocene group acts like a quantum dot and that NDR features are due to resonant tunneling between the HOMO state and the density of states of the gold lead.

ATOMIC STRUCTURE, OPTICAL PROPERTIES AND ELECTRON TRANSPORT IN SELF-ASSEMBLED MONOLAYERS ON SURFACES

by

SHUCHUN WANG

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

2006

APPROVED BY:

Dr. Jerry Bernholc Chair of Advisory Committee	Dr. Wenchang Lu
Dr. Marco Buongiorno Nardelli	Dr. Chris Roland
Dr. Zhilin Li	

Biography

Shuchun Wang was born on March 18th, 1977 in Renqiu, Hebei Province, China. He graduated from No. 1 High School in Renqiu in 1995. He received his college education at Nankai University in Tianjin, China, and completed his Bachelor of Sciences degreen in Physics there in 1999. He then was admitted as a graduate student in Physics Department of Nankai University with a honored waiver of entrance test. To pursue higer level academic goals, he decided to come to the United States and joined the Ph.D. program in the Department of physics at North Carolina State University in August 2000. In the next year, He past the Ph.D. written qualification exam and joined Dr. Bernholc's Solid State Theory group.

Acknowledgements

First and foremost, I would like to thank my wonderful supervisor Dr. Jerry Bernholc for his guidance, training, and continuous support during the five-year research work. I also appreciate Dr. Wenchang Lu's daily guidance in computational and programming techiques and help in physics. Working with them has greatly extended my problem solving and creative abilities.

I am grateful to Dr. Marco Buongiorno Nardelli, Dr. Chris Roland, and Dr. Zhilin Li for their helpful suggestions and valuable comments to this work. My special thanks also go to Dr. Wolf Gero Schmidt and Dr. Christopher Gorman for the engaging discussion about my projects.

Thank also to my friends and workmates, Qingzhong Zhao, Liping Yu, Filipe Ribeiro, Miroslav Hodak, Gunn Kim, Vivek Ranjan, and Frisco Rose for their helpful conversation about physics and everyday help.

Finally, I want to thank my wife, my parents and my brothers and sister for their selfless love and support!

Contents

List of Figures **vii**

List of Tables **ix**

1 *Ab Initio* **Total Energy Methodology** **1**

1.1 Born-Oppenheimer Approximation 4

1.2 Density-Functional Theory 6

1.2.1 Hohenberg-Kohn Theorem 8

1.2.2 Kohn-Sham Equations 11

1.2.3 Total-Energy Functional 13

1.2.4 Exchange-Correlation Approximation 14

1.3 Pseudopotential Theory 17

1.3.1 Operator Representation 18

1.3.2 Generalization Criteria 21

1.3.3 Kleinman-Bylander Pseudopotential 22

1.4 Supercell Approximation 24

1.4.1 Bloch's Theorem 24
1.4.2 **k**-point Sampling 26
1.5 Basis Sets 28
1.5.1 Analytical Basis Sets 29
1.5.2 Real-Space Basis Sets 31
1.6 Real-Space Grid-Based Multigrid DFT Method 33
1.6.1 Mehrstellen Discretization of the Kohn-Sham Equations . . . 34
1.6.2 Fourier Filtering of Pseudopotentials 36
1.6.3 Multigrid Techniques 38
1.7 Optimized Nonorthogonal Orbital Order-N DFT Method 40
1.7.1 Basis-Invariant Matrix Formulation 41
1.7.2 Computations with Unoccupied Orbitals 43
1.8 Summary 45

2 Implementation of Ultrasoft pseudopotentials in Real-Space Method 46

2.1 Introduction to Vanderbilt Ultrasoft Pseudopotentials 48
2.1.1 Generalized Kohn-Sham Equations 48
2.1.2 Ionic Forces 51
2.2 Implementation of UPP into our Real-Space Methods 53
2.2.1 Implementation of Two Uniform Global Grids 53
2.2.2 Double-Grid Technique for Nonlocal Projectors 55
2.2.3 Six-Order Mehrstellen Discretization 59

2.3 Test Results 61
2.3.1 Bond Lengths of Diatomic molecules 61
2.3.2 Ground state of bulk gold 62
2.3.3 Ground state and band structure of bulk copper 64
2.4 Summary 66

3 Nanowire-induced optical anisotropy of the Si(111)-In surface 68
3.1 Introduction 69
3.2 Methodology 72
3.3 Results and Discussion 75
3.4 Summary and conclusions 80

4 Resonant Coupling and Negative Differential Resistance in Metal/Ferrocenyl-Alkanethiolate/STM structures 81
4.1 Introduction 82
4.2 Electronic Structure Calculation 84
4.3 Quantum Transport Calculation 86
4.4 Summary and Conclusions 94

Bibliography 96

List of Figures

1.1 Schematic illustration of the pseudopotential concept. 20

1.2 Schematic illustration of the supercell geometry of the Indium-Silicon(111) 4×1 surface. 25

2.1 Plot of the pseudo-wavefunction and the nonlocal projector with $l = 1$ for diamond calculation. 56

2.2 Equation of state for bulk gold in fcc structure. 63

2.3 Bulk copper DFT-LDA band structure calculated with UPP. 64

3.1 Optimized atomic structures of the zigzag-chain model of the Si(111)-In surface. 70

3.2 The experimental observed RAS spectrum of the Si(111)-In 4×1 and 4×2/8×2 phases. 72

3.3 The optimized atomic structure of the π-bonded chain stacking-fault model for Si(111)4×1-In reconstruction. 74

3.4 RAS spectrum calculated for the π-SF model of the (4×1)-In reconstruction 76

3.5 RAS spectra calculated for the zigzag-chain models of the Si(111)-In surface reconstructions. 77

4.1 Atomic structures of C_5S-Au(111) $(\sqrt{3} \times \sqrt{3})R30°$, Fc-$C_5S$-Au(111) $(\sqrt{3} \times \sqrt{3})R30°$, and Fc-$C_5S$-Au(111) $(\sqrt{21} \times \sqrt{7})$ 85

4.2 Schematic views of the Au-SC_5S-Fc-Au system and of isosurfaces of the HOMO and LUMO orbitals at zero bias 87

4.3 Calculated transmission spectra $T(E, V)$ for the Au(111)-SC_5S-Fc-Au(111) junction with $d = 3.39$ Å 89

4.4 Current-voltage characteristics of the Au(111)-SC_5S-Fc-Au(111) junction with $d = 3.39$ Å 90

4.5 The NDR mechanism for the Au(111)-SC_5S-Fc-Au(111) junction at 1.2 V 91

4.6 Current-voltage characteristics of the Au(111)-SC_5S-Fc-Au(111) junction as a function of the STM-molecule distance 93

List of Tables

1.1 Discretization weights for the fourth-order orthorhombic Mehrstellen operators. 35

2.1 Discretization weights for the sixth-order orthorhombic Mehrstellen operators. 60

2.2 Bond lengths of diatomic molecules. 62

2.3 Test on ground-state properties of bulk gold. 63

2.4 Comparison of Cu bandwidths with the values from experiment measurement and the plane-wave method calculation with norm-conserving pseudopotential at high-symmetry points. 65

2.5 Test on ground-state properties of bulk copper. 65

Chapter 1

Ab Initio Total Energy Methodology

Since the theory of modern quantum mechanics was established in the early 1890s, it has been proved correct and very successful to provide fundamental understanding for the phenomena on the atomic scale, such as the energy levels of atoms and the covalent bonds of molecules. Although a lot of philosophers argue about the interpretation and the predictions of modern quantum theory, few question the astounding accuracy with which quantum mechanics describes the world around us. For instance, the calculated gyromagnetic ratio of the electron by relativistic quantum field theory [1] agrees with experimental measurement [2] within one in a million. Scientists have little doubt that most of low-energy physics, chemistry and biology can be explained and fundamentally understood by applying the quantum theory to electrons and ions. However, the most challenging work for contemporary physicists is how to set up and solve the complicated equations of quantum mechanics, which describes the complex processes happening in real materials.

Many different theoretical methods have been proposed and developed over the last century. Among them, the *ab initio* methods or first-principles methods stand out, because these methods require nothing but the Schrödinger equation, the values of the fundamental constants, and the atomic numbers of the atoms present. Unlike the semi-empirical methods, the *ab initio* methods do not depend on experimental data. These methods can be used to design new materials and predict new properties. Of all the methods, one, the total-energy pseudopotential method [3, 4] stands alone due to the fact that many ground state physical properties of molecules and solids are related to total energies or to differences between total energies. With the most efficient numerical algorithms, this method can simulate systems with thousands of atoms and opens a wide range of interesting problems to quantum-mechanical calculations. It not only plays a leading role in the theoretical study of physics, but also benefits chemists, biologists, geophysicists, and material scientists.

This chapter presents a detailed and comprehensive review of the total-energy pseudopotential method. The essential concepts and the major theoretical simplifications and approximations are introduced. In Section 1.1, we describe the Born-Oppenheimer approximation. This "adiabatic principle" decouples the electron and nuclear degrees of freedom and reduces the many-body problem to the solution of the dynamics of the electrons in some frozen-in configuration of the nuclei. Section 1.2 introduces the state-of-the-art density-functional theory, which models the electron-electron interactions. It allows one, in principle, to map exactly the problem of strongly interacting electron gas (in the presence of nuclei) onto that of a sin-

gle particle moving in an effective potential, which include Coulumb and exchange-correlation potential (described by Kohn-Sham equations). The exchange-correlation potential can be approximated by Local Density Approximation or Generalized Gradient Approximation. Section 1.3 contains a detailed description of pseudopotential theory, which models the electron-ion interactions. This theory allows one to replace the strong electron-ion potential with a much weaker potential–a pseudopotential–that describes all the salient features of a valence electron moving through the solid, including relativistic effects. Section 1.4 presents the supercell approximation used to deal with aperiodic configurations of atoms within the framework of Bloch's theorem. One simply constructs a large unit cell containing the configuration in question and enough vaccum space and repeats it periodically. By studying the properties of the system for larger and larger unit cells, one can gauge the importance of the induced periodicity and systematically filter it out. A numerical solution of the Kohn-Sham equations is based on the choice of an appropriate basis to represent the electronic wave functions. An overview of some popular basis sets is given in Section 1.5. In Section 1.6, we provide a brief introduction to our real-space multigrid DFT method in which the Kohn-Sham equation is solved numerically in a real space uniform grid. We use this method to calculate the atomic and electronic structure, and the optical property of surfaces. In Section 1.7 we discuss an O(N) method in a basis of variationally optimized local orbitals. The Last section summarizes the whole chapter.

1.1 Born-Oppenheimer Approximation

In condensed matter physics and chemistry, the electronic structure of atoms, molecules and solids is known as a quantum many-body problem due to the coupling of the electron interactions. The properties of any non-relativistic time-independent quantum system can be determined by solving the Schrödinger equation,

$$\hat{H}\tilde{\Psi}(\{\mathbf{r}_i\},\{\mathbf{R}_\alpha\}) = E\tilde{\Psi}(\{\mathbf{r}_i\},\{\mathbf{R}_\alpha\})\,, \tag{1.1}$$

where $\tilde{\Psi}(\{\mathbf{r}_i\},\{\mathbf{R}_\alpha\})$ is the many-body wavefunction, which is a function of a set of electronic position variable $\{\mathbf{r}_i\}$ and another set of ionic position variable $\{\mathbf{R}_\alpha\}$, E is the system energy, and the Hamiltonian of the system $\hat{H}$ is given by (in atomic units)

$$\begin{aligned}\hat{H} &= -\frac{1}{2}\sum_i \nabla_i^2 - \frac{1}{m_a}\sum_\alpha \nabla_\alpha^2 - \sum_i\sum_\alpha \frac{Z_\alpha}{|\mathbf{r}_i-\mathbf{R}_\alpha|} \\ &\quad +\frac{1}{2}\sum_i\sum_{j\neq i}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|} + \frac{1}{2}\sum_\alpha\sum_{\alpha\neq\beta}\frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha-\mathbf{R}_\beta|}\,,\end{aligned} \tag{1.2}$$

in which m_α is the ionic mass and Z_α is the atomic number. The first two terms on the right-hand side represent the kinetic energies of the electrons and ions respectively. The subsequent terms describe the electron-nucleus, electron-electron, and inter-nucleus Coulomb interactions.

The Born-Oppenheimer approximation [5] is the first of several approximations made when trying to solve Schrödinger equation for complex systems with more than one or two electrons. It is based on the idea that the mass of ions is much larger than the mass of electrons and the fact that the forces on the particles due to their electric charges are of the same order of magnitude. This implies that a velocity

of an electron is much larger than that of an ion. As a consequence, the electrons respond essentially instantaneously to the motion of the ions and rapidly relax to the instantaneous ground-state configuration on the typical time-scale of the ionic motion. Thus the ion can be treated adiabatically, leading to a separation of electronic and ionic coordinates in the many-body wavefunction—known as Born-Oppenheimer approximation:

$$\tilde{\Psi}(\{\mathbf{r}_i\},\{\mathbf{R}_\alpha\}) = \Psi(\{\mathbf{r}_i\})\Phi(\{\mathbf{R}_\alpha\})\,, \tag{1.3}$$

where $\Phi(\{\mathbf{R}_\alpha\})$ is the nuclear wavefunction and $\Psi(\{\mathbf{r}_i\})$ is the electronic wavefunction, which is a wavefunction only of the $\{\mathbf{r}_i\}$ and satisfies Schrödinger equation for the electrons in a static array of nuclei:

$$\left[-\frac{1}{2}\sum_i \nabla_i^2 - \sum_i\sum_\alpha \frac{Z_\alpha}{|\mathbf{r}_i-\mathbf{R}_\alpha|} + \frac{1}{2}\sum_i\sum_{j\neq i}\frac{1}{|\mathbf{r}_i-\mathbf{r}_j|}\right]\Psi(\{\mathbf{r}_i\}) = \varepsilon_e(\{\mathbf{R}_\alpha\})\Psi(\{\mathbf{r}_i\})\,. \tag{1.4}$$

Here the dependence of the eigenvalues ε_e on the ionic positions is acknowledged. Substituting the whole wavefunction (Eq. 1.3) to the Schrödinger equation for the whole system(Eq. 1.1) leads to

$$\begin{aligned}\tilde{H}\tilde{\Psi}(\{\mathbf{r}_i\},\{\mathbf{R}_\alpha\}) &= \tilde{H}\Psi(\{\mathbf{r}_i\})\Phi(\{\mathbf{R}_\alpha\}) = E\Psi(\{\mathbf{r}_i\})\Phi(\{\mathbf{R}_\alpha\}) \\ &= \Psi(\{\mathbf{r}_i\})\Big[-\sum_\beta \tfrac{1}{2m_\beta}\nabla_\beta^2 + \varepsilon_e(\{\mathbf{R}_\alpha\}) + \tfrac{1}{2}\sum_\beta\sum_{\gamma\neq\beta}\tfrac{Z_\beta Z_\gamma}{|\mathbf{R}_\beta-\mathbf{R}_\gamma|}\Big]\Phi(\{\mathbf{R}_\alpha\}) \\ &\quad -O\Big(\textstyle\sum_\beta \frac{1}{m_\beta}\Big)\,,\end{aligned} \tag{1.5}$$

in which the last term comes from the non-adiabatic part. The size of this term is proportional to the ratio of the electron and nuclear masses, typically a factor of the

order of 10^{-4} or 10^{-5}, so the contributions from this term can be neglected. Therefore, above Eq. 1.5 is satisfied if the nuclear wavefunction obeys the Schrödinger equation of the form

$$\left[-\sum_{\beta} \frac{1}{2m_{\beta}} \nabla_{\beta}^{2} + \varepsilon_{e}(\{\mathbf{R}_{\alpha}\}) + \frac{1}{2} \sum_{\beta} \sum_{\beta \neq \gamma} \frac{Z_{\beta} Z_{\gamma}}{|\mathbf{R}_{\beta} - \mathbf{R}_{\gamma}|} \right] \Phi(\{\mathbf{R}_{\alpha}\}) = E\Phi(\{\mathbf{R}_{\alpha}\}) \,. \quad (1.6)$$

The adiabatic principle is crucial, because it allows one to separate the nuclear and electronic motions. For most cases, the motion of the nuclei can be ignored and we only need to treat a system with a fixed nuclear configuration. Thus, the complexity of the full many-body Schrödinger equation reduces to that of an electron Schrödinger equation (Eq. 1.4). The dependence of the electronic energy ε_e and of the electron wavefunction on the nuclear positions $\{\mathbf{R}_{\alpha}\}$ will be suppressed.

1.2 Density-Functional Theory

With the Born-Oppenheimer approximation, the many-body problem has been reduced to that of a gas of interacting electrons moving in a static external potential due to the frozen-in nuclei. Solving the electronic Schrödinger Eq. 1.4 is, however, still a formidable task for most cases since the many-electron wavefunction contains 3N variables, which for a solid containing $N \sim 10^{26}$ electrons is simply an intractable number of degrees of freedom. Devising accurate schemes to approximate the many-electron problem has been an important goal since the founding of quantum mechanics. In 1930, the Hartree-Fock theory [6, 7] was developed, which is based on the single-particle approximation proposed earlier by Hartree [8], but in

addition correctly accounts for the exchange interactions between electrons that are a consequence of the Pauli principle. The many-electron wavefunction is constructed by anti-symmetrizing the single-particle functions $\psi_i(\mathbf{r}_i\sigma_i)$

$$\Psi(\{\mathbf{r}_1\sigma_1\},\{\mathbf{r}_2\sigma_2\}\ldots\{\mathbf{r}_N\sigma_N\}) \approx \frac{1}{\sqrt{N}}\mathcal{A}\left[\psi_1(\mathbf{r}_1\sigma_2)\psi_2(\mathbf{r}_2\sigma_2)\ldots\psi_N(\mathbf{r}_N\sigma_N)\right] ,$$

where $\mathcal{A}$ is an anti-symmetrizing operator. The reduction of the Coulomb energy of the electronic system due to the antisymmetry of the wavefunction is called the exchange energy. The Hartree-Fock method yields very good bond lengths in molecules, but the binding energies are in general not in good agreement with experimentally obtained ones. Moreover, the Hartree-Fock method has problems in describing the band structures of solids, which are of immense importance in solid-state physics. This is because the energy of the system calculated in the Hartree-Fock method is not the correct ground state energy, due to the fact that spatial separation of the electrons with the opposite spins can further reduce the Coulomb energy. The difference between the many-electron ground state energy of an electronic system and the energy of the system calculated in the Hartree-Fock approximation is called the correlation energy.

In order to include correlation effects without using the very costly wavefunction methods, a remarkable theorem of density functional theory (DFT) was developed by Hohenberg and Kohn [9] and Kohn and Sham [10]. They proposed a simple way for describing the effects of exchange and correlation in an electron gas. In 1964 Hohenberg and Kohn proved that the total energy, including exchange and

correlation energy of an electron gas, is a unique functional of the electron density, which is a function of three Cartesian variables, rather than $3N$ variables as the full many-body wavefunction is. The minimum value of the total-energy functional is the ground-state energy of the system, and the density that yields this minimum value is the exact single-particle ground-state density. In the next year, Kohn and Sham then showed how it is possible, formally, to replace the many-electron problem by an exactly equivalent set of self-consistent one-electron equations. After many extremely successful applications in quantum chemistry and condensed matter physics, DFT has become the foundation of most *ab initio* calculations.

1.2.1 Hohenberg-Kohn Theorem

For any system consisting of N electrons moving under the influence of an external potential $V_{ext}(\mathbf{r})$ due to the static nuclei

$$V_{\text{ext}}(\mathbf{r}) = -\sum_{\alpha} \frac{Z_\alpha}{|\mathbf{r} - \mathbf{R}_\alpha|} \,, \tag{1.7}$$

the electronic Hamiltonian in 1.4 can be written as

$$\hat{H} = \hat{T} + \hat{V}_{\text{ee}} + \sum_{i}^{N} V_{\text{ext}}(\mathbf{r}_i) = \hat{F} + \hat{V}_{\text{ext}} \,, \tag{1.8}$$

where $\hat{V}_{\text{ext}}$ is the operator corresponding to the external potential, and $\hat{F}$ is the electronic Hamiltonian consisting of a kinetic energy operator $\hat{T}$ and an electron-electron interaction operator $\hat{V}_{\text{ee}}$. The electronic Hamiltonian $\hat{F}$ is the same for all N-electron systems. The Hamiltonian $\hat{H}$ and the ground state $\Psi_0(\mathbf{r})$ are completely

determined by N and $V_{ext}(\mathbf{r})$. The ground-state $\Psi_0(\mathbf{r})$ for this Hamiltonian gives rise to a ground-state electronic density $n_0(\mathbf{r})$

$$n_0(\mathbf{r}) = \langle \Psi_0 | \hat{n} | \Psi_0 \rangle = N \int |\Psi_0 (\mathbf{r}, \mathbf{r}_2, \mathbf{r}_3 \dots \mathbf{r}_N)|^2 \, \mathrm{d}\mathbf{r}_2 \, \mathrm{d}\mathbf{r}_3 \dots \, \mathrm{d}\mathbf{r}_N \tag{1.9}$$

Thus the ground state $|\Psi_0\rangle$ and the density $n_0(\mathbf{r})$ are both functionals of the number of electrons N and the external potential $V_{\mathrm{ext}}(\mathbf{r})$.

The first Hohenberg-Kohn theorem then states: ***The external potential*** $V_{\mathrm{ext}}(\mathbf{r})$ ***is uniquely determined by the corresponding ground state electronic density, to within an additive constant***.

Proof by *reductio ad absurdum*: Assume that there are two potentials $V_{\mathrm{ext}}(\mathbf{r})$ and $V'_{\mathrm{ext}}(\mathbf{r})$, which differ by more than an additive constant, and further that these two potentials lead to the same density, $n_0(\mathbf{r})$. The associated Hamiltonians, $\hat{H} = \hat{F} + \hat{V}_{\mathrm{ext}}$ and $\hat{H}' = \hat{F} + \hat{V}'_{\mathrm{ext}}$ will therefore have different ground state wavefunctions $\Psi_0(\mathbf{r})$ and $\Psi'_0(\mathbf{r})$. The ground state energies are $E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle$ and $E'_0 = \langle \Psi'_0 | \hat{H}' | \Psi'_0 \rangle$ respectively. Using the variational principle [11], we first take $\Psi'_0(\mathbf{r})$ as a trial wavefunction for $\hat{H}$. It yields

$$\begin{aligned} E_0 < \langle \Psi'_0 | \hat{H} | \Psi'_0 \rangle &= \langle \Psi'_0 | \hat{H}' | \Psi'_0 \rangle + \langle \Psi'_0 | \left(\hat{H} - \hat{H}' \right) | \Psi'_0 \rangle \\ &= E'_0 + \int \mathrm{d}\mathbf{r} \; n_0(\mathbf{r}) \left[V_{\mathrm{ext}}(\mathbf{r}) - V'_{\mathrm{ext}}(\mathbf{r}) \right] . \end{aligned} \tag{1.10}$$

whereas taking Ψ_0 as a trial wave-function for $\hat{H}'$ gives

$$\begin{aligned} E'_0 < \langle \Psi_0 | \hat{H}' | \Psi_0 \rangle &= \langle \Psi_0 | \hat{H} | \Psi_0 \rangle + \langle \Psi_0 | \left(\hat{H}' - \hat{H} \right) | \Psi_0 \rangle \\ &= E_0 - \int \mathrm{d}\mathbf{r} \; n_0(\mathbf{r}) \left[V_{\mathrm{ext}}(\mathbf{r}) - V'_{\mathrm{ext}}(\mathbf{r}) \right] . \end{aligned} \tag{1.11}$$

Addition of the two inequalities 1.10 and 1.11 leads to the result:

$$E_0 + E_0' < E_0 + E_0' \,,$$

which is an obvious contradiction, and as the result the ground state density uniquely determines the external potential $V_{ext}(\mathbf{r})$. It also determines all ground state properties since the $V_{ext}(\mathbf{r})$ and N completely define the Hamiltonian and the ground state wavefunction.

The second Hohenberg-Kohn theorem says: ***The ground state energy may be obtained variationally: the density that minimizes the total energy is the exact ground state density***.

Proof: following Jones and Gunnarsson [12], we define a universal functional of the density $n(\mathbf{r})$ for the operator $\hat{F}$ as:

$$F[n(\mathbf{r})] = \min_{\Psi \to n} \langle \Psi | \hat{F} | \Psi \rangle \,, \tag{1.12}$$

where the minimum is taken over all $\Psi(\mathbf{r})$ that yield the density $n(\mathbf{r})$. Now, if we introduce $\Psi^n_{min}(\mathbf{r})$ for a wavefunction that minimizes Eq.1.12 such that

$$F\left[n(\mathbf{r})\right] = \langle \Psi^n_{min} | \hat{F} | \Psi^n_{min} \rangle \,,$$

the variational principle results in

$$\begin{aligned} E\left[n(\mathbf{r})\right] &= \int V_{ext}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} + F\left[n(\mathbf{r})\right] \\ &= \langle \Psi^n_{min} | V_{ext} + \hat{F} | \Psi^n_{min} \rangle \\ &\geq \langle \Psi_0 | V_{ext} + \hat{F} | \Psi_0 \rangle = E_0\left[n_0(\mathbf{r})\right] \,. \end{aligned}$$

Thus the ground-state density $n_0(\mathbf{r})$ minimizes the functional $E[n(\mathbf{r})]$ and the minimum value is the ground-state electronic energy.

1.2.2 Kohn-Sham Equations

The key idea of the Kohn-Sham method is to map the problem of the system of interacting electrons onto a fictitious system of non-interacting "electrons". Variation of the total energy functional obtained from Eq.1.8 with respect to the electron density, subject to the constraint the number of electrons N is conserved, gives us

$$\delta\left[F[n(\mathbf{r})] + \int \mathrm{d}\mathbf{r}\ V_{\mathrm{ext}} n(\mathbf{r}) - \mu\left(\int \mathrm{d}\mathbf{r}\ n(\mathbf{r}) - N\right)\right] = 0\,, \tag{1.13}$$

where μ is the Lagrange multiplier associated with the constraint of constant N. The corresponding Euler-Lagrange equation is given by

$$\frac{\delta F\,[n(\mathbf{r})]}{\delta n(\mathbf{r})} + V_{\mathrm{ext}}(\mathbf{r}) = \mu\,. \tag{1.14}$$

Kohn and Sham separated $F[n(\mathbf{r})]$ into three parts as

$$F[n(\mathbf{r})] = T_{\mathrm{s}}[n(\mathbf{r})] + E_{\mathrm{H}}[n(\mathbf{r})] + E_{\mathrm{XC}}[n(\mathbf{r})]\,,$$

where $T_s[n(\mathbf{r})]$ is the kinetic energy of a non-interacting electron gas of density $n(\mathbf{r})$, $E_H[n(\mathbf{r})]$ is the classical electrostatic (Hartree) energy of electrons

$$E_H[n(\mathbf{r})] = \frac{1}{2}\int\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\ \mathrm{d}\mathbf{r}\ \mathrm{d}\mathbf{r}'\,, \tag{1.15}$$

and E_{XC} is an implicit definition of the exchange-correlation energy, which contains the difference between the kinetic energies of the interacting and non-interacting

systems and also the non-classical electrostatic interaction energy. The Euler Eq. 1.14 is rewritten as

$$\frac{\delta T_{\rm s}[n]}{\delta n(\mathbf{r})} + V_{\rm KS}(\mathbf{r}) = \mu \, . \tag{1.16}$$

Here $V_{\rm KS}(\mathbf{r})$ is the effective Kohn-Sham potential

$$V_{\rm KS}(\mathbf{r}) = \int {\rm d}\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + V_{\rm XC}(\mathbf{r}) + V_{\rm ext}(\mathbf{r}) \, , \tag{1.17}$$

in which the first term is the Hartree potential $V_{\rm H}$, and the exchange-correlation potential V_{XC} is defined by

$$V_{\rm XC}(\mathbf{r}) = \frac{\delta E_{\rm XC}[n]}{\delta n(\mathbf{r})} \, . \tag{1.18}$$

Now, one can think about the Euler Eq. 1.16 in another way, namely that it describes a non-interacting system of electrons moving in an external potential V_{KS}. To find the ground state density, one can alternatively just solve the **N** one-electron Schrödinger equations,

$$\left[-\tfrac{1}{2}\nabla^2 + V_{\rm KS}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}) \, . \tag{1.19}$$

The density is then constructed from

$$n(\mathbf{r}) = \sum_{i=1}^{N} f_i \, |\psi_i(\mathbf{r})|^2 \, , \tag{1.20}$$

where f_i are occupation numbers, and the non-interacting kinetic energy is given by

$$T_{\rm s}[n(\mathbf{r})] = -\frac{1}{2} \sum_{i=1}^{N} f_i \int \, \psi_i^*(\mathbf{r}) \nabla^2 \psi_i(\mathbf{r}) \, {\rm d}\mathbf{r} \, . \tag{1.21}$$

These equations are known as Kohn-Sham equations.

Since the Kohn-Sham potential $V_{\rm KS}(\mathbf{r})$ depends upon the density $n(\mathbf{r})$, the Kohn-Sham equations must be solved self-consistently. Beginning with an initial guess

for the form of the density, the electronic potentials that are used to construct the Kohn-Sham equations are generated by using this charge density through Eq. 1.17. After that, the Schrödinger equation 1.19 is solved to obtain a set of electronic states $\{\psi_i(\mathbf{r})\}$. From the occupied electronic states a new charge density is then formed for the next iteration (Eq. 1.20). The whole process is repeated until the density gets converged, i.e., the input and output densities are the same. In practice, the charge density is directly converged to the ground-state charge density by using efficient minimization algorithms.

1.2.3 Total-Energy Functional

From the above Kohn-Sham equations, the sum of the one-electron Kohn-Sham eigenvalues can be expressed as

$$\begin{aligned}\sum_{i=1}^{N} \boldsymbol{\varepsilon}_i &= T_{\rm s}[n] + \int \mathrm{d}\mathbf{r}\; n(\mathbf{r}) V_{\rm KS}(\mathbf{r}) \\ &= T_{\rm s}[n] + \int \mathrm{d}\mathbf{r}\; \mathrm{d}\mathbf{r}'\; \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} + \int \mathrm{d}\mathbf{r}\; n(\mathbf{r}) V_{\rm XC}(\mathbf{r}) \\ &\quad + \int \mathrm{d}\mathbf{r}\; n(\mathbf{r}) V_{\rm ext}(\mathbf{r})\,, \end{aligned} \tag{1.22}$$

which does not give the total electronic energy because it double counts the Hartree energy and over estimates the exchange-correlation energy. After the correcting for these two items in the total electronic energy and accounting for the Coulomb energy $E_{\rm ion}$ associated with interactions among the nuclei (or ions), the Kohn-Sham total-

energy functional can be written as

$$E_{\rm tot} = \sum_{i=1}^{N} \varepsilon_i - \frac{1}{2} \int {\rm d}\mathbf{r}\ {\rm d}\mathbf{r}'\ \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} - \int {\rm d}\mathbf{r}\ n(\mathbf{r}) V_{\rm XC}(\mathbf{r}) + E_{\rm XC}[n] + E_{\rm ion}\ , \quad (1.23)$$

with

$$E_{\rm ion} = \frac{1}{2} \sum_{\alpha} \sum_{\beta \neq \alpha} \frac{Z_\alpha Z_\beta}{|\mathbf{R}_\alpha - \mathbf{R}_\beta|}\ .$$

The minimized total energy at the ground-state density is just the ground-state energy of the whole system.

1.2.4 Exchange-Correlation Approximation

So far the Kohn-Sham equations are exact, but the actual form of the exchange-correlation energy functional $E_{\rm XC}[n(\mathbf{r})]$ is not known exactly. We must introduce approximate functionals based upon the electron density to describe this term. Many different approaches have been used to create improved functionals after the Kohn-Sham paper. Here I describe two popular and almost generally used approximations in total-energy pseudopotential calculations: the local density approximation (LDA)[12] and the generalized gradient approximation (GGA) [13].

Local Density Approximation

The local density approximation is the oldest and the simplest approximation. In LDA, the exchange-correlation energy of an electronic system is constructed by assuming that the exchange-correlation energy per electron at a point $\mathbf{r}$ in the electron gas, $\varepsilon_{\rm XC}(\mathbf{r})$, is equal to the exchange-correlation energy per electron in a homogeneous

electron gas that has the same density as the electron gas at point $\mathbf{r}$. Thus

$$E_{XC}^{LDA}\left[n(\mathbf{r})\right] = \int \varepsilon_{\mathrm{XC}}(\mathbf{r}) n(\mathbf{r})\, \mathrm{d}\mathbf{r} \tag{1.24}$$

and the exchange-correlation potential V_{XC} may be written

$$V_{\mathrm{XC}}^{\mathrm{LDA}}(\mathbf{r}) = \frac{\delta E_{\mathrm{XC}}^{\mathrm{LDA}}\left[n(\mathbf{r})\right]}{\delta n(\mathbf{r})} = \frac{\partial\left[n(\mathbf{r})\varepsilon_{\mathrm{XC}}(\mathbf{r})\right]}{\partial n(\mathbf{r})} \tag{1.25}$$

with

$$\varepsilon_{\mathrm{XC}}(\mathbf{r}) = \varepsilon_{\mathrm{XC}}^{\mathrm{hom}}\left[n(\mathbf{r})\right] \ . \tag{1.26}$$

The LDA assumes that the exchange-correlation energy functional is purely local. The most common parametrization in use for $\varepsilon_{\mathrm{XC}}^{\mathrm{hom}}$ is that of Perdew and Zunger [14], which is based upon the quantum Monte Carlo calculations of Ceperley and Alder [15] for homogeneous electron gases at various densities. The parameterizations provide interpolation formulas linking exact results for the exchange-correlation energy of high-density electron gases and calculations of the exchange-correlation energy of inter mediate and low-density electron gases.

The LDA ignores corrections to the exchange-correlation energy due to inhomogeneities in the electron density at a point $\mathbf{r}$. It is surprising that the LDA has been so successful for total-energy calculation, considering its inexact nature of the approximation. Work by Harris and Jones [16], Gunnarsson and Lundqvist [17], and Langreth and Perdew [18] showed that this is due to the fact that the LDA gives the correct sum rule for the exchange-correlation hole.

Generalized Gradient Approximation

Although the LDA works well for solid systems and has been used in solid state calculations for many years, it has severe trouble when dealing with molecules because of its overbinding. Generalized gradient approximation was defined in the early and mid eighties [19, 20, 21], aiming at incorporating the effects of inhomogeneities by including the gradient of the electron density. In GGA the exchange-correlation energy functional can be written as

$$E_{\mathrm{XC}}^{\mathrm{GGA}}[n(\mathbf{r})] = \int n(\mathbf{r})\varepsilon_{\mathrm{XC}}^{\mathrm{hom}}\left[n(\mathrm{r})\right] F_{\mathrm{XC}}[n(\mathbf{r}), \nabla n(\mathbf{r})]\mathrm{d}\mathbf{r}\,, \tag{1.27}$$

where $F_{\mathrm{XC}}[n(\mathbf{r}), \nabla n(\mathbf{r})]$ is known as the enhancement factor. Unlike the LDA, there is no unique form for the GGA, and indeed many different variations are possible as the enhancement factor [13, 22, 23, 24]. Among them, the PW91 exchange and correlation functional was constructed by introducing a real-space cut-off of the spurious long-range part of the density-gradient expansion for the exchange and correlation hole [25, 26]. The cut-off procedure was designed in such a way that as many as possible of the known features of the exact exchange and correlation energy were obeyed. It has later been discovered that there are some unphysical wiggles in the PW91 exchange-correlation potential for small and large reduced-density gradients. There are also quite many parameters in the PW91 functional, and it has been found that other important and exactly known features of the exchange-correlation energy exist, apart frpm those satisfied by the PW91. To remedy the weaknesses of the PW91 functional, a modified form was recently devised by Perdew, Burke and Ernzerhof known as PBE

functional [23, 24], which uses a much simplified exchange enhancement factor of the form:

$$F_{\mathbf{X}}^{\mathbf{PBE}}(s) = 1 + \kappa - \frac{\kappa}{1 + \mu s^2/\kappa} \,,$$

where $s(\mathbf{r})$ is the dimensionless reduced density gradient

$$s(\mathbf{r}) = \frac{|\nabla n(\mathbf{r})|}{2k_{\mathbf{F}}(\mathbf{r})n(\mathbf{r})}$$

with

$$k_{\mathbf{F}}(\mathbf{r}) = [3\pi^2 n(\mathbf{r})]^{1/3} \,,$$

$\mu = 0.21951$ and $\kappa = 0.804$. The PBE form was designed to give a simpler functional form by retaining only the most energetically important conditions satisfied by PW91. Today it has become the most popular GGA functional.

1.3 Pseudopotential Theory

The density functional theory greatly simplifies the electron-electron interaction by introducing the exchange-correlation functional, but it does not touch upon the ion-electron interaction. In Kohn-Sham formalism, the electron-ion interaction is described by the external potential V_{ext} in Eq. 1.7. These Coulomb potentials, due to the nuclei, are an important component in *ab initio* calculations. Because they are singular near the cores, special care must be taken when solving the Kohn-Sham equations, particularly when using plane-wave on grid bases. Additionally, the strong nuclear Coulomb potentials also lead to highly localized core electron wavefunctions

and highly oscillatory valence electron wavefunction near the cores, which are difficult to represent computationally. Fortunately, the pseudopotential theory introduced by Phillips [27] and Cohen and Heine [28] makes it possible to overcome these problems.

The Pseudopotential theory is based upon the fact that the majority of physical properties of solids depend upon the valence electrons. In contrast, the core electrons are almost environment independent. Thus, the chemically inert core electrons are assumed to be frozen with the nuclei. Both of them are considered as a pseudo-ion core and act much like the nuclear core in generating an external Coulomb potential. The total external potential of the *all-electron* atom, including the nuclear core and the core electrons, is then replaced by a smooth, non-singular pseudopotential that only acts on the valence electrons. Correspondingly, the valence wavefunctions are changed to pseudo wavefunctions, where the oscillations in the vicinity of the ions have been removed.

1.3.1 Operator Representation

For better understanding of the pseudopotential concepts, we follow the cancelation theorem derived from the orthogonalized plane wave (OPW) formalism by Phillips and Kleinman [29]. In the OPW approach, considering an atom with Hamiltonian H, core states $\{|\chi_n\rangle\}$ and core energy eigenvalues $\{E_n\}$, the valence state $|\psi\rangle$ with energy eigenvalue E is constructed as the combination of a smooth wavefunction $|\psi^{\mathrm{PS}}\rangle$ and

a sum over all core states $\{|\chi_n\rangle\}$ in the form

$$|\psi\rangle = |\psi^{\mathrm{PS}}\rangle + \sum_n^{\mathrm{core}} a_n |\chi_n\rangle \,. \tag{1.28}$$

The requirement that the valence state $|\psi\rangle$ is orthogonal to the core state $|\chi_m\rangle$ yields

$$0 = \langle\chi_m|\psi\rangle = \langle\chi_m|\psi^{\mathrm{PS}}\rangle + a_m \,, \tag{1.29}$$

which fixes the expansion coefficients $\{a_n\}$. Thus

$$|\psi\rangle = |\psi^{\mathrm{PS}}\rangle - \sum_n^{\mathrm{core}} |\chi_n\rangle\langle\chi_n|\psi^{\mathrm{PS}}\rangle \,. \tag{1.30}$$

We know that the valence electron satisfies below Schrödinger equation:

$$H|\psi\rangle = (T + V_{\mathrm{C}})|\psi\rangle = E|\psi\rangle \,. \tag{1.31}$$

where V_{C} is the attractive core potential. Substituting Eq. 1.30 leads to

$$H|\psi^{\mathrm{PS}}\rangle - \sum_n^{\mathrm{core}} E_n|\chi_n\rangle\langle\chi_n|\psi^{\mathrm{PS}}\rangle = E|\psi^{\mathrm{PS}}\rangle - E\sum_n^{\mathrm{core}} |\chi_n\rangle\langle\chi_n|\psi^{\mathrm{PS}}\rangle \,, \tag{1.32}$$

which can be rearranged in the form

$$\begin{aligned} \left[T + V_{\mathrm{C}} + \sum_n^{\mathrm{core}} (E - E_n)|\chi_n\rangle\langle\chi_n|\right]|\psi^{\mathrm{PS}}\rangle &= (T + V^{\mathrm{PS}})|\psi^{\mathrm{PS}}\rangle \\ &= E|\psi^{\mathrm{PS}}\rangle \,. \end{aligned} \tag{1.33}$$

Therefore, the original Kohn-Sham equation for a valence state can be rewritten as an equation for a smooth pseudo-state $|\psi^{\mathrm{PS}}\rangle$, which obeys a Schrödinger equation with an energy-dependent pseudopotential

$$V^{\mathrm{PS}} = V_{\mathrm{C}} + \sum_n^{\mathrm{core}} (E - E_n)|\chi_n\rangle\langle\chi_n| \,. \tag{1.34}$$

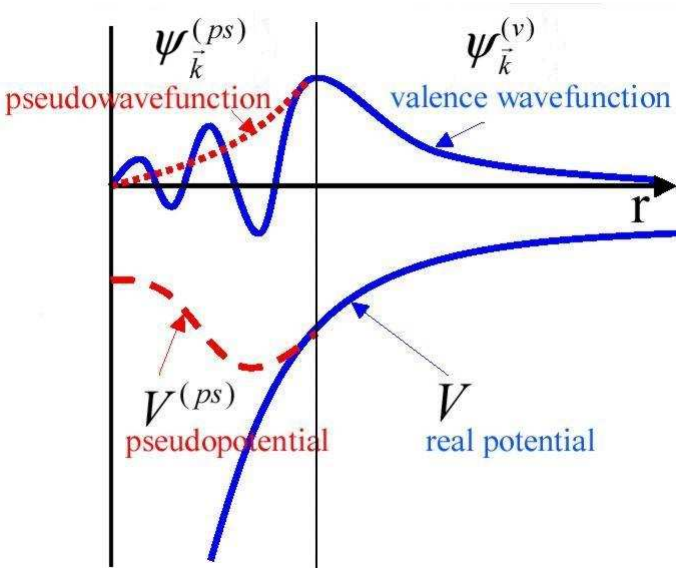

Figure 1.1: **Schematic illustration of the pseudopotential concept. The solid lines represent the all-electron potential and wavefunction, while the dashed lines show the corresponding pseudopotential and pseudo-wavefunction. The radius at which the all-electron and pseudo quantities match is designated r_c.**

It has been shown that the second term of $V^{\rm ps}$ is a strongly repulsive potential, which almost cancels the strongly attractive core potential $V_{\rm C}$, leaving a weak net pseudopotential. It is obvious that the pseudopotential simplifies not only the interactions between the valence electron and the ionic core, but makes it easier to represent the valence wave functions in an efficient manner by the pseudo wavefunctions. Fig.1.1 shows a typical ionic potential, valence wavefunction, the corresponding pseudopotential and the pseudo-wavefunction.

1.3.2 Generalization Criteria

So far, many different schemes to create pseudopotentials [30, 31, 32, 33, 34] exist, but there are some general criteria for the construction of pseudopotentials. They are listed below:

1. The eigenvalues of the pseudo Hamiltonian are equal to those of the *all-electron* Hamiltonian:

$$\epsilon^{\mathrm{PS}} = \epsilon^{\mathrm{AE}} .$$

2. The pseudopotentials outside a cutoff radius r_c agree with the *all-electron* potentials:

$$V^{\mathrm{PS}}(r > r_c) = V^{\mathrm{AE}}(r > r_c) .$$

3. The pseudo wavefunctions are identical to the *all-electron* wavefunctions outside the cutoff radius r_c:

$$\psi^{\mathrm{PS}}(r > r_c) = \psi^{\mathrm{AE}}(r > r_c) .$$

4. The charge inside r_c of a pseudo wavefunction is equal to that of a *all-electron* wavefunction:

$$\int_0^{r_c} \mathrm{d}r r^2 |\psi^{\mathrm{PS}}(r)|^2 = \int_0^{r_c} \mathrm{d}r r^2 |\psi^{\mathrm{AE}}(r)|^2 .$$

The forth constraint is often denoted norm-conservation. This norm-conserving constraint is not satisfied by the so called ultrasoft pseudopotential developed recently by Vanderbilt [35]. In these pseudopotentials an augmentation charge is instead included as a supplement to the pseudo wavefunction. I will discuss this new ultrasoft

pseudopotential in detail in Chapter 3. The last two constraints guarantee that the logarithmic derivatives of the *all-electron* atom's wavefunctions and the pseudoatom's wavefunctions agree to first order beyond the cutoff radius r_c. Since the logarithmic derivative is related to the phase shift of a free electron scattering off a radial potential [11], this ensures that the scattering properties of the pseudoatom match those of the *all-electron* atom. Actually, numerical calculation shows that the scattering properties of the *all-electron* potential and the pseudopotential are identical over a wide range of energies, not just at the eigenenergies. This makes the pseudopotential to have very good transferability, i.e., it can accurately describe the valence electrons in different atomic, molecular and solid-state environments.

1.3.3 Kleinman-Bylander Pseudopotential

The phase shifts produced by the ion core are different for each angular momentum component of the valence wavefunction, and the scattering from the pseudopotential must therefore be angle-dependent, so it is necessary to use the angle-dependent non-local pseudopotential. In general, the non-local pseudopotential is expressed in semi-local [33] form

$$\begin{aligned} V^{\mathrm{PS}} &= V_{\mathrm{loc}} + \sum_{l \neq \mathrm{loc}, m} |lm\rangle (V_l - V_{\mathrm{loc}}) \langle lm| \\ &= V_{\mathrm{loc}} + \sum_{l \neq \mathrm{loc}, m} |lm\rangle \delta V_l \langle lm| \,, \end{aligned} \tag{1.35}$$

in which $|lm\rangle$ denotes the spherical harmonic Y_{lm}. The choice of local pseudopotential V_{loc} is arbitrary. However, this semi-local form is too costly for using in calculations,

since the number of matrix elements that need to be calculated scales as the square of the number of basis states.

Another alternative construction by Kleinman and Bylander [31] (KB) introduces a fully separable form of the pseudopotential, in which the double sum is split into a product of two single sums,

$$V_{\mathrm{KB}} = V_{\mathrm{loc}} + \sum_{l,m} \frac{|\psi_{lm}\delta V_l\rangle\langle\delta V_l\psi_{lm}|}{\langle\psi_{lm}|\delta V_l|\phi_{lm}\rangle}\,, \tag{1.36}$$

where ψ_{lm} are the pseudo-atom wavefunctions. This operator acts on the reference state in an identical manner to the original semi-local one, but the full separation allows the calculation to scale linearly with the size of the basis set.

The Kleinman and Bylander approach has been widely used in pseudopotential constructions because it is computationally efficient. However, in transforming a semi-local pseudopential to the corresponding KB one can lead to unphysical ghost states [36] at energies below or near those of the physical valence states. Such spurious states can occur for specific choices of the underlying semilocal and local pseudopotentials. They are the artifacts of the KB-form nonlocality by which the nodeless pseudo wavefunctions need not be the lowest possible eigenstate, unlike in the semi-local form. In practice, transferable ghost-free KB pseudopotentials are readily obtained by a proper choice of the local component (V_{loc}) and/or the core cutoff radius (r_c^l) of the pseudopotential.

1.4 Supercell Approximation

So far, the density functional and pseudopotential theories substantialy simplify the calculation of solid properties by reducing the many-body interacting problem to the motion of non-interacting electrons in a static effective potential

$$V(\mathbf{r}) = V^{\mathrm{PS}}(\mathbf{r}) + V_{\mathrm{H}}(\mathbf{r}) + V_{\mathrm{XC}}(\mathbf{r}).$$

However, handling a of bulk crystal still remains formidable due to the huge number of electrons and nuclei. Fortunately, most of the systems are periodic, so it is possible to reduce this problem to a finite system by only considering electrons within the unit cell. This marvelous approach is not only suitable for periodic solids, but also turns out to be the easiest way to study nonperiodic systems by imposing some false periodicity, such as molecules, defects and surfaces. The nonperiodic system is contained within a supercell that is replicated periodically throughout the whole space. Fig.1.2 shows the supercell for a silicon indium surface that we studied, which contains a crystal slab and a vacuum region. It is essential that the vacuum region is large enough so that the faces of adjacent crystal slabs, which in reality are isolated, do not signigicantly interact across the vacuum region.

1.4.1 Bloch's Theorem

In a perfect crystal, the nuclei are arranged in a periodic array described by a set of Bravais lattice vectors $\{\mathbf{R}\}$. The effective static potential $V(\mathbf{r})$ is periodic and

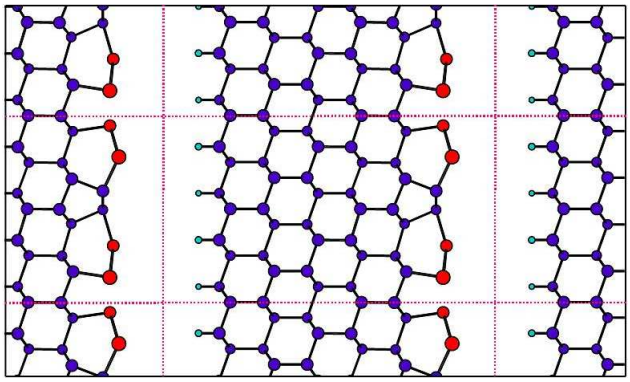

Figure 1.2: **Schematic illustration of the supercell geometry of the Indium-Silicon(111) 4×1 surface.**

invariant under translation by any of the lattice vectors

$$\hat{T}_{\mathrm{R}} V(\mathbf{r}) = V(\mathbf{r} + \mathrm{R}) = V(\mathbf{r}) \,, \tag{1.37}$$

where $\hat{T}_{\mathrm{R}}$ are the translation operators. The Hamiltonian is also periodic $\hat{H}(\mathbf{r}+\mathbf{R}) = \hat{H}(\mathbf{r})$ and commutes with the translation operators:

$$\hat{T}_{\mathbf{R}} \hat{H}(\mathbf{r}) \psi(\mathbf{r}) = \hat{H}(\mathbf{r} + \mathbf{R}) \psi(\mathbf{r} + \mathbf{R}) = \hat{H}(\mathbf{r}) \hat{T}_{\mathbf{R}} \psi(\mathbf{r}) \,. \tag{1.38}$$

Therefore, there must be a set of eigenstates of the Hamiltonian that are simultaneous eigenstates of all the translation operators, marked by a good quantum number corresponding to each lattice vector R.

Bloch's theorem [37] states that the wavefunction of an electron $\psi_{n,\mathbf{k}}$, for a periodic potential, is a product of a cell-periodic part $u_{n,\mathbf{k}}(\mathbf{r})$ and a wavelike part $\mathrm{e}^{i\mathbf{k}\cdot\mathbf{r}}$,

$$\psi_{n,\mathbf{k}}(\mathbf{r}) = u_{n,\mathbf{k}}(\mathbf{r}) \;\; \exp(i\mathbf{k} \cdot \mathbf{r}) \,, \tag{1.39}$$

where the subscript n labels different eigenstates of the Hamiltonian and k is a quantum wavevector, related to the translational symmetry. Since $u_n(\mathbf{r})$ possesses the

same periodicity as the potential ($u_{n,\mathbf{k}}(\mathbf{r}+\mathbf{R}) = u_{n,\mathbf{k}}(\mathbf{r})$), it can always be expanded in terms of a Fourier series i.e.

$$u_{n,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n,\mathbf{k}}(\mathbf{G}) \ \exp(i\mathbf{G}\cdot\mathbf{r}) \,, \tag{1.40}$$

where $\mathbf{G}$ are reciprocal lattice vectors defined by $\mathbf{G}\cdot\mathbf{R} = 2\pi m$, where m is an arbitrary integer, and $c_{n,\mathbf{k}}(\mathbf{G})$ are plane wave expansion coefficients. Plugging Eq. 1.40 into Eq .1.39, we see that the electron wavefunction may be expressed as a linear combination of plane waves:

$$\psi_{n,\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n,\mathbf{k}+\mathbf{G}} \ \exp[i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}] \,. \tag{1.41}$$

The Bloch theorem changes the problem of calculating an infinite number of electronic wavefunctions to one of calculating a finite number of electronic wavefunctions within a single (super) cell at an infinite number of k points. In order to reduce the allowed k points, it is necessary to impose some boundary conditions on the bulk solid.

1.4.2 k-point Sampling

Here, we apply periodic or generalized Born-von Karman boundary condition to the wavefunction. Such choice of boundary condition does not affect the bulk properties of the system and is expressed mathematically as

$$\psi_{n,\mathbf{k}}(\mathbf{r} + N_i\mathbf{a}_i) = \psi_{n,\mathbf{k}}(\mathbf{r}), \qquad i = 1,2,3 \,, \tag{1.42}$$

where N_i is the number of primitive cells along the $\mathbf{a_i}$ direction. Applying Bloch's theorem (Eq.1.39) yields

$$\psi_{n,\mathbf{k}}(\mathbf{r} + N_i\mathbf{a}_i) = \exp(\mathrm{i}N_i\mathbf{k}\cdot\mathbf{a}_i)\psi_{n,\mathbf{l}}(\mathbf{r})\,, \tag{1.43}$$

so that the values of $\mathbf{k}$ are restricted and must satisfy

$$\exp(\mathrm{i}N_i\mathbf{k}\cdot\mathbf{a}_i) = 1, \qquad i = 1, 2, 3\,. \tag{1.44}$$

Such restriction directly results in the general form for the allowed Bloch wavevectors $\mathbf{k}$ such that

$$\mathbf{k} = \sum_{i=1}^{3} \frac{l_i}{N_i}\mathbf{g}_i, \tag{1.45}$$

where the $\{l_i\}$ are integers and the $\{g_i\}$ are the primitive reciprocal lattice vectors satisfying $\mathbf{g}_i \cdot \mathbf{a}_j = 2\pi\delta_{ij}$. We can see that there is still an infinite number of allowed $\mathbf{k}$ points if $N_i \to \infty$, but they are now members of a countably infinite set. Furthermore, we can prove that $\mathbf{k}$-vectors that differ only by a reciprocal lattice vector are in fact equivalent. Considering two such wavevectors related by $\mathbf{k}' = \mathbf{k} + \mathbf{G}$, then the corresponding Bloch states are also related by

$$\begin{aligned} \psi_{n,\mathbf{k}'}(\mathbf{r}) &= \exp(i\mathbf{k}'\cdot\mathbf{r})u_{n,\mathbf{k}'}(\mathbf{r}) = \exp(i\mathbf{k}\cdot\mathbf{r})\left[u_{n\mathbf{k}'}(\mathbf{r})\exp(i\mathbf{G}\cdot\mathbf{r})\right] \\ &= \exp(i\mathbf{k}\cdot\mathbf{r})\tilde{u}(\mathbf{r}) = \psi_{n'\mathbf{k}}(\mathbf{r}). \end{aligned} \tag{1.46}$$

Since the expression in square brackets on the first line is a cell-periodic function, the whole expression is a valid Bloch wavefunction with wavevector $\mathbf{k}$ but different band index n'. Thus we can pay our attention only to those $\mathbf{k}$ points that lie within the first Brillouin zone.

In order to construct the charge density and get an accurate electronic potential, we have to calculate the occupied eigenstates at each allowed $\mathbf{k}$ point within the first Brillouim zone. However, we do not need to consider all of these $\mathbf{k}$ points since the wavefunctions vary smoothly over Brillouin zone [38]. Practically, only a finite set of $\mathbf{k}$ points sampling the Brillouin zone is chosen to determine the total energy of the solid. For very large systems, the Brillouin zone volume becomes very small and only the Γ-point ($\mathbf{k} = \mathbf{0}$) can be choosen without significant loss of accuracy. Many methods for making efficient choices have been developed since 1970 [39, 40, 41, 42, 43, 44]. Using these methods, one can obtain an accurate approximation for the electronic potential and the total energy of an insulator or a semiconductor by calculating the electronic states at a very small number of $\mathbf{k}$ points. For a metallic system, it is more difficult to calculate them because a dense set of $\mathbf{k}$ points is required to define the Fermi surface precisely. From the calculation of the wavefunctions on a much coarser grid of $\mathbf{k}$ points, $\mathbf{k} \cdot \mathbf{p}$ perturbation theory [45, 46] can be used to generate the wavefunctions on the dense set of $\mathbf{k}$ points. This method greatly reduces the computational cost for a metallic system.

1.5 Basis Sets

To solve the Kohn-Sham equations numerically, one has to employ a set of basis functions in order to efficiently represent the electronic wavefunctions. Generally, two categories of basis set are used: analytical base and numerical base. Several

factors are required to consider before choosing a basis for a calculation, such as speed, accuracy, ease of implementation, and the geometry of the studied system.

1.5.1 Analytical Basis Sets

As an example of an analytical basis set, I briefly outline here the plane-wave basis. This basis set has the advantage of being mathematically simple, and is in principle complete, that is, it completely spans the Hilbert space, which makes it more accurate to represent the electronic wavefunctions. In principle, the plane-wave basis used to expand the wavefunctions in Eq.1.41 should be infinite. Fortunately, the plane-waves with small kinetic energy have a more important role than those with a very high kinetic energy, so one can introduce a kinetic energy cutoff energy E_{cut} to truncate the basis set to a finite size. It is defined through

$$E_{cut} = \frac{\hbar^2}{2m} |\mathbf{k} + \mathbf{G}|^2 \ . \tag{1.47}$$

This kinetic energy cut-off will lead to an error in the total energy of the system, but it is possible to make this error arbitrarily small by increasing the size of the basis set by allowing a larger energy cut-off. For these reasons, plane-wave based calculation are common and popular in solid physics.

With the plane-wave basis, the Kohn-Sham equations can be expressed in a particularly simple form. Substitution of Eq.1.41 into the Kohn-Sham equation 1.19,

gives

$$\sum_{\mathbf{G}'} \left[\frac{1}{2} |\mathbf{k}+\mathbf{G}|^2 \delta_{\mathbf{G}\mathbf{G}'} + V_{\text{ext}}(\mathbf{G}-\mathbf{G}') + V_{\text{XC}}(\mathbf{G}-\mathbf{G}') + V_H(\mathbf{G}-\mathbf{G}') \right] c_{n,\mathbf{k}+\mathbf{G}'} = \epsilon_n c_{n,\mathbf{k}+\mathbf{G}}. \tag{1.48}$$

In this form, the reciprocal space presentation of the kinetic energy is diagonal and the various potentials are described in terms of their Fourier components. The usual methods of solving the plane-wave expansion of the Kohn-Sham equations is by diagonalization of the Hamiltonian matrix whose elements $H_{\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}'}$ are given by the terms in the square bracket. The size of the Hamiltonian matrix is determined by the number of plane waves, *i.e.*, the energy cut-off E_{cut}.

Plane-wave based methods have been used to perform electronic structure calculations on a wide range of physical systems over last several decades. The Car-Parrinello [47] iterative method has enabled for systems containing several hundred systems. An excellent review of plane-wave calculations is given by Payne *etal.* [48]. Although highly successful, the traditional plane wave methods encounter considerable difficulties when they are applied to physical system with large length scales, or containing first-row or transition metal atoms. Special techniques have been developed to partially eliminated these difficulties including preconditioned conjugate gradient techniques [49, 50, 51], optimized pseudopotentials [35, 34], augmented-wave methods [52], and plane waves in adaptive coordinates [53, 54, 55]. However, these methods are still constrained by the plane-wave-basis set, which requires periodic boundary conditions for every system and performing fast Fourier transforms (FFT's)

between the real and reciprocal spaces. FFT's involve nonlocal operations and restrict their efficient implementation in massively parallel computer architectures, due to the requirement of long-range communications.

1.5.2 Real-Space Basis Sets

Recently, real space methods have enjoyed considerable popularity because the basis functions are localized in real space and do not lead to a large communication overhead. The construction of the Hamiltonian may thus be achieved in $O(N_{basis})$ operations. Additionally, boundary conditions are not constrained to be periodic, and clusters and surfaces with nonperiodic boundary conditions can be handled efficiently. More importantly, the use of a real-space basis opens up the possibility of using multigrid iterative techniques to obtain solutions of the Kohn-Sham equations. These techniques provide automatic preconditioning on all length scales and thus greatly decrease the number of iterations needed to converge the electronic wavefunctions. Within variety of real-space basis sets, there are two general classes: orbital based and non-orbital based.

In orbital-independent methods, the wavefunction is expanded on a mesh of points in real space. These methods include finite-element [56], grids [57, 58, 59, 60], wavelet [61, 62], and multigrids [63, 64]. The first three methods were developed to study a single atom, diatomics, and small molecular systems. The real-space multigrid method devised by our group is suitable for large-scale calculations (hundreds of atoms). I

will describe this method in detail in the next section. In the orbital-independent methods, just like in the plane-wave method, the size of the basis set is much larger than the number of electrons ($N_{\text{basis}} \gg N_{\text{e}}$). The accuracy of these methods can be increased in a systematic way and one generally expects these methods to be more accurate than the orbital based methods.

The orbital basis is derived from the basic knowledge that the angular momentum eigenstates of a single atom closely approximate the eigenstates of the electron in the presence of several atoms. These methods help naturally understand concepts such as chemical bonds and hybridized orbitals, which are also classified by angular momentum. For orbital-based methods, the number of basis functions is roughly of the order of the number of electron ($N_{\text{basis}} \sim N_{\text{e}}$). In general, the systems they can handle are much bigger than those the orbital-independent methods do, due to the fact that the critical parts of computational work scales linearly with the number of atoms [65]. However, the bottleneck in large calculations is diagonalization, which scales as $\text{O}(N^3)$, unless special techniques are employed. The choices of basis set include Gaussian, optimized nonorthogonal atomic orbitals [66], the Fireball orbital [67, 68], and many semi-empirical basis sets. Comparing with the non-orbital-based methods; their accuracy cannot generally be increased in a systematic way, due to the incompleteness of the basis sets. However this does not constrain their ability to describe a wide range of phenomena.

In Chapter 4, when we investigate the transport properties of the molecular devices, it is necessary to construct the charge density for an open system from the

Green's function and not from the wavefunctions. At present, it is impractical to calculate the Green's function for any of the non-orbital based methods, because of enormous time and memory requirements. We employ the optimized nonorthogonal atomic orbital method to accommplish this task. It will be discussed in detail in section 1.7.

1.6 Real-Space Grid-Based Multigrid DFT Method

The real-space grid-based multigrid DFT method, which is developed by Briggs, Sullivan, and Bernholc [64], uses a real-space grid as a basis and the multigrid technique to provide effective convergence acceleration and preconditioning on all length scales. It is not only suitable for large-scale ab initio calculations, but it is also capable to conduct efficient calculations for ill-conditioned systems with long length scales or high-energy cutoffs.

In this method, the wavefunctions, the density, and all potentials are directly represented on a uniform three-dimensional real-space grid with linear spacing h_{grid} and number of mesh points N_{grid}. The physical coordinates of each point are denoted by

$$\mathbf{r}(i,j,k) = (ih_{\text{grid}}, jh_{\text{grid}}, kh_{\text{grid}}) \,, \tag{1.49}$$

$$i = 1, \ldots, N_x, j = 1, \ldots, N_y, k = 1, \ldots, N_z \,.$$

Since the set of grid points to map the above quantities is discrete, it can introduce spurious dependence of the Kohn-Sham eigenvalues, the total energy, and the ionic

forces on the positions of ions with respect to the real-space grid. A set of techniques has been developed in order to overcome these difficulties and accurately calculate the properties of large physical systems.

1.6.1 Mehrstellen Discretization of the Kohn-Sham Equations

The first task to solve the Kohn-Sham equation in real space is to discretize it on the grid structure described by Eq. 1.49. Usually, the standard central finite-difference method is used to approximate the differential operator in the Hamiltonian. it is found that a higher-order central finite-difference expression (typically $O(h_{\text{grid}}^8)$) is indispensable in order to achieve high accuracy [69]. As a result, it requires a large communication overhead in parallel computation, which makes it very costly for a large-scale calculations.

Here, a generalized eigenvalue formulation is introduced to discretize the Kohn-Sham equations. It is expressed by

$$\mathbf{H}_{\text{Mehr}}[\psi_n] = \frac{1}{2}\mathbf{A}_{\text{Mehr}}[\psi_n] + \mathbf{B}_{\text{Mehr}}[V_{\text{eff}}\psi_n] = \epsilon_n \mathbf{B}_{\text{Mehr}}[\psi_n]. \tag{1.50}$$

where $\mathbf{A}_{\text{Mehr}}$ and $\mathbf{B}_{\text{Mehr}}$ are components of the Mehrstellen discretization [70], which is based on Hermite's generalization of Taylor's theorem. The fourth-order Mehrstellen

Table 1.1: **Discretization weights for the fourth-order orthorhombic Mehrstellen operators for the central, nearest-neighbor, and next nearest-neighbor grid points. The cubic-grid operator corresponds to** $h_i = h_{\text{grid}}$

	a	b_n	$c_{n,m}$
$\mathbf{A}_{\text{Mehr}}$	$\sum_i \frac{4}{3h_i^2}$	$-\frac{5}{6h_n^2} + \sum_i \frac{1}{6h_i^2}$	$-\frac{1}{12h_n^2} - \frac{1}{12h_m^2}$
$\mathbf{B}_{\text{Mehr}}$	$\frac{1}{2}$	$\frac{1}{12}$	0

discretization samples the Hamiltonian and the wavefunction at 19 points such that:

$$\begin{aligned}
\mathbf{A}_{\text{Mehr}} &= af(\mathbf{x}) + \sum_{n=1}^{3} b_n f(\mathbf{x} \pm h_n \hat{\mathbf{x}}_n) \\
&+ \sum_{n<m} c_{n,m} f(\mathbf{x} \pm h_n \hat{\mathbf{x}}_n \pm h_m \hat{\mathbf{x}}_m) \,, \\
\mathbf{B}_{\text{Mehr}} &= a' f(\mathbf{x}) + \sum_{n=1}^{3} b'_n f(\mathbf{x} \pm h_n \hat{\mathbf{x}}_n) \,.
\end{aligned}$$

It can be shown that the small h expansions of the $\mathbf{A}_{\text{Mehr}}$ and $\mathbf{B}_{\text{Mehr}}$ satisfy $\mathbf{A}_{\text{Mehr}} = \mathbf{B}_{\text{Mehr}}(-\nabla^2) + O(h^4)$. Mehrstellen discretization differs from the central finite-difference discretization in that it uses more local information (next-nearest and second-next neighbors) and a weighted sum of the wavefunction and potential values to improve the accuracy of discretization of the entire differential equation, not just the kinetic-energy operator. Those improvements make the accuracy of Mehrstellen discretization one order higher than the corresponding central finite-difference one, but this accuracy is fully realized only at convergence. Table 1.1 lists the fourth-order Mehrstellen operators specified by their discretization weights, which pertain to both cubic and orthorhombic grids.

This representation of the Kohn-Sham Hamiltonian is short ranged in real space

in the sense that the operator can be applied to any orbital in $O(N_{\text{grid}})$ operations. Specifically, the application of the $\mathbf{A}_{\text{Mehr}}$ operator at a point involves a sum over 19 orbital values, while the application of the $\mathbf{B}_{\text{Mehr}}$ operator requires a sum over seven points. The local potential multiplies the orbital pointwise, and the short-ranged nonlocal projectors require one integration over a fixed volume around each ion and a pointwise multiplication. This sparseness permits the use of iterative diagonalization techniques, and the short-ranged representation of the Hamiltonian leads to an efficient implementation on massively parallel computers.

1.6.2 Fourier Filtering of Pseudopotentials

The calculation of total energy (Eq.1.23) involves a lot of integral operations. All of them are performed by using the three-dimensional trapezoidal rule

$$\int \mathrm{d}\mathrm{r} f(\mathrm{r}) \doteq h_{\text{grid}}^3 \sum_{ijk} f(\mathrm{r}(i,j,k)) \,. \tag{1.51}$$

For high accuracy, it is essential that the integrand $f(\mathrm{r})$ should be band limited, that is, its Fourier transform should have minimal magnitude in the frequency range $G > G_{\text{max}} \equiv \pi/h_{\text{grid}}$. Otherwise, the high-frequency components can manifest themselves on the grid and introduce unphysical variation in the total energy or the electron charge density, especially when the ions, and hence their pseudopotentials, shift relative to the grid points. This is explicit in the plane-wave method since the basis is truncated by the energy cutoff E_{cut}. The discrete real-space grid also defines a kinetic energy cutoff of approximately $G_{\text{max}}^2/2$. The variation of the total energy can

be decreased by explicitly eliminating the high-frequency components beyond $G_{\rm cut}$ in the pseudopotentials by Fourier filtering.

Initially, the unfiltered potentials or projectors, after their Coulomb tails have been subtracted out, are defined on a real-space radial grid and short-ranged, being confined to a "core" radius around each atom. They are then transformed to momentum space, where the high-frequency components are filtered out through

$$V_{l,\rm filtered}(G) = F_{\rm filter}(G/G_{\rm cut}) \int V_l(r) j_l(Gr) r^2 dr, \tag{1.52}$$

where the cutoff function $F_{\rm filter}(G/G_{\rm cut})$ is defined by

$$F_{\rm filter}(G/G_{\rm cut}) = \begin{cases} 1 & \text{if } G < G_{\rm cut} \\ e^{-\beta_1 (G/G_{\rm cut}-1)^2} & \text{if } G \geq G_{\rm cut} \end{cases}, \tag{1.53}$$

which smoothly attenuates the radial Fourier components beyond $G_{\rm cut}$, the cutoff wave vector determined by the grid spacing: $G_{\rm cut} = \alpha\pi/h_{\rm grid}$. After momentum-space filtering, the backtransformed potentials and projectors will extend beyond the original core radius. In order to keep the short-range of the original nonlocal pseudopotentials, a second real-space filtering is applied to reduce the large-radius oscillations beyond a cutoff radius $r_{\rm cut}$:

$$V_{l,\rm filtered}(r) = f_{\rm filter}(r/r_{\rm cut}) \int V_{l,\rm filtered}(G) j_l(Gr) G^2 dG. \tag{1.54}$$

The second filtering function $f_{\rm filter}(r/r_{\rm cut})$ is unity below $r_{\rm cut}$ and equals $e^{-\beta_2 (r/r_{\rm cut}-1)^2}$ above it. The parameters α, β_1, β_2, and $r_{\rm cut}$ are empirical and depend on the atomic species.

1.6.3 Multigrid Techniques

When solving Kohn-Sham equations and Poisson's equation $-\nabla^2 V_{\mathrm{H}} = 4\pi\rho$ on real space grids, the convergence rates of general iterative relaxation schemes such as the Jacobi, Gauss-Seidel, successive overrelaxation, or the alternating direction implicit method [71] are too low for the grid sizes required in electronic structure calculation, because these methods only considerably reduce the high-frequency components of the error (the ones with wavelengths comparable to the grid spacing), but do not affect the lowest-frequency components (the ones with wavelength large relative to the grid spacing) by much. The multigrid iteration technique can accelerate convergence by employing a sequence of grids of varying resolutions. The essence of the multigrid approach is the fact that the individual frequency components of the error are best reduced on a grid where the resolution is of the same order of magnitude as the wavelength of the error component. It thus treats the lower-frequency components on a sequence of auxiliary grids with progressively larger grid spacings. This technique provides excellent preconditioning for all length scales and leads to very rapid convergence rates.

In order to illustrate this method, I describe the multigrid algorithm to solve the Poisson's equation with two uniform grids; a fine grid of spacing h and a coarser auxiliary one of spacing H. More generally, the Poisson's equation can be represented by a set of linear algebraic equations

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{1.55}$$

where $\mathbf{x}$ and $\mathbf{b}$ represent the discretized forms of V_{H} and $4\pi\rho$, respectively, and $\mathbf{A}$ is the finite-difference representation of $-\nabla^2$. In the Jacobi iterative method, the residual $\mathbf{r}$ of the current approximate solution $\mathbf{x}$ is expressed by

$$\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}. \tag{1.56}$$

An improved $\mathbf{x}$ is generated as

$$\mathbf{x}^{\mathrm{new}} = \mathbf{x} + \Delta t K \mathbf{r}, \tag{1.57}$$

where Δt is a pseudo-time-step and K is the inverse of the diagonal component of $\mathbf{A}$. First, the high-frequency components of the solution error with wavelength $\approx h$ are reduced by one or two Jacobi iterations. Then the fine grid residual $\mathbf{r}_h$ is calculated by Eq. 1.56, and is transferred to the coarse grid by volume-weight restriction. Next, the Poisson's equation on the coarse grid with the residual as a source term is solved by using the same iterative procedure as in Eq. 1.57. Similarly, the Jacobi iteration on this level removes error components with wavelength $\approx H$. Finally, the coarse-grid solution is transferred back to the fine grid by simple trilinear interpolation [72] and is added to the fine-grid solution. As we know, the final accuracy of the solution is determined only by the finite-difference representation of $-\nabla^2$ on the fine grid level. Thus the Mehrstellen operator is used on the fine grids, while a seven-point central finite-difference operator ($O(H^4)$) is used on the coarse grids, due to stability problems.

An extension of multigrid implementation to the solution of the Kohn-Sham equation introduces several complications since the nonlocal pseudopotential is used and

the Kohn-Sham equations are not linear any more. Instead of Eq. 1.56, the residual to the eigenvalue problem is generalized as

$$\mathbf{r}_h = \epsilon_n \mathbf{B}_{\text{Mehr}}[\psi_n] - \mathbf{H}_{\text{Mehr}}[\psi_n], \tag{1.58}$$

where ϵ_n is an estimate of the eigenvalue calculated from the Rayleigh quotient of the generalized eigenvalue Eq. 1.50,

$$\epsilon_n = \frac{<\psi_n|\mathbf{H}_{\text{Mehr}}[\psi_n]>}{<\psi_n|\mathbf{B}_{\text{Mehr}}[\psi_n]>}. \tag{1.59}$$

To obtain high accuracy and rapid convergence rate, additional techniques have been developed to deal with the orthonormality constraints, subspace diagonalization and so on. For more details, one can refer to the literature [64].

1.7 Optimized Nonorthogonal Orbital Order-N DFT Method

This real-space optimized nonorthogonal orbital DFT method was proposed by Fattebert and Bernholc [66]. The localized nonorthogonal atomic obitals are optimized variationally. In our implementation, it uses a multigrid preconditioner and allows for unoccupied or partially occupied orbitals to be included in the basis set, which substantially accelerates convergence. The critical computational operations achieve roughly linear scaling with the total number of atoms, allowing for *ab initio* calculations well beyond a thousand atoms. The local orbital description is also very useful and helpful in quantum transport calculations [73, 74].

1.7.1 Basis-Invariant Matrix Formulation

Here we follow the basis-invariant matrix formulation derived by Fattebert and Bernholc [66]. Begining with a trial basis of normalized functions $\{\phi_1, \cdots, \phi_N\}$, the basis will be optimized by iterative updates, and it will accurately describe the true Kohn-Sham ground state of the system at convergence. In matrix notation, a matrix Φ is written as

$$\Phi = (\phi_1, \cdots, \phi_N).$$

The eigenfunctions of the system can be expanded as

$$\Psi = (\psi_1, \cdots, \psi_N) = \Phi C, \tag{1.60}$$

where C is an $N \times N$ transformation matrix. This matrix satisfies

$$CC^T = S^{-1}$$

where $S = \Phi^T \Phi$ is the overlap matrix.

In the following, for an operator A we will use the notation

$$A^{(\Psi)} = \Psi^T A \Psi, A^{(\Phi)} = \Phi^T A \Phi.$$

we then have the relation $A^{(\Psi)} = C^T A^{(\Phi)} C$.

The transformation matrix C can be obtained by solving the generalized symmetric eigenvalue problem

$$H^{(\Phi)} C = SC\Lambda, \tag{1.61}$$

where $\Lambda_{ij} = \varepsilon_i \delta_{ij}$ with ε_i being the eigenvalues.

In the basis Ψ, the steepest descent (SD) directions, along which the energy functional decreases at the fastest rate, are easy to compute. They are given by the negative residuals of the Kohn-Sham equations, with the current potential kept fixed, and can be expressed as the $N \times M$ matrix

$$D^{(\Psi)} = \Psi\Lambda - H\Psi, \tag{1.62}$$

which satisfies the relation $\Psi^T D^{(\Psi)} = 0$. In the nonorthogonal orbitals Φ, the SD directions can differ substantially from the derivative with respect to $\phi_j(\mathbf{r})$, especially when the basis Φ is highly nonorthogonal. The simplest way to obtain it is to use the matrix $D^{(\Psi)}$ and C through

$$\begin{aligned} D^{(\Phi)} = D^{(\Psi)}C^{-1} &= (\Psi\Lambda - H\Psi)C^{-1} \\ &= \Phi\Theta - H\Phi, \end{aligned} \tag{1.63}$$

where

$$\Theta = S^{-1}H^{(\phi)}. \tag{1.64}$$

The preconditioned steepest descent (PSD) direction in the nonorthogonal basis is thus

$$\begin{aligned} \delta\Phi &= KD^{(\Psi)}C^{-1} \\ &= K(\Phi\Theta - H\Phi), \end{aligned} \tag{1.65}$$

and the basis Φ is updated by Pulay mixing [75].

In actual calculations, the basis functions Φ are corrected at each iteration using the PSD directions (Eq. 1.65). A new electronic density $\rho(\mathbf{r})$ is then generated

according to

$$\rho(\mathbf{r}) = 2 \sum_{j,k=1}^{N} (S^{-1})_{jk} \Phi_j(\mathbf{r}) \Phi_k(\mathbf{r}).$$

Hartree and exchange-correlation potentials are then updated with a linear mixing.

1.7.2 Computations with Unoccupied Orbitals

The inclusion of unoccupied or partially occupied orbitals can substantially enhance the convergence rate. The PSD algorithm (Eq. 1.65) can be used to improve the trial subspace of N computed orbitals, regardless whether they are occupied or empty. In the computation of the electron density and the total energy, however, one needs to account for the occupations. Thus, new density-matrix formalism in the nonorthogonal basis needs to be introduced.

For a chemical potential μ, define the $N \times N$ matrix Υ by its matrix elements

$$\Upsilon_{ij} = \delta_{ij} f[(\epsilon_i - \mu)/k_B T],$$

where f is the Fermi-Dirac distribution. The density operator $\hat{\rho}$ is then given by

$$\hat{\rho} = \Psi \Upsilon \Psi^T = \Phi C \Upsilon C^T \Phi^T.$$

In our notation, one can represent $\hat{\rho}$ in the basis Φ

$$\rho^{(\Phi)} = \Phi^T \hat{\rho} \Phi = C^{-T} \Upsilon C^T.$$

We introduce the matrix $\bar{\rho}^{(\Phi)}$

$$\bar{\rho}^{(\Phi)} = S^{-1} \rho^{(\Phi)} S^{-1} = C \Upsilon C^T. \tag{1.66}$$

This matrix can be used to obtain the expectation value $\bar{A}$ of an operator A represented in the basis Φ

$$\bar{A} = 2\mathrm{Tr}(\Upsilon A^{(\Phi)}) = 2\mathrm{Tr}(\bar{\rho}^{(\Phi)} A^{(\Phi)}).$$

In particular, the total number of electrons in the system is given by

$$N_e = 2\mathrm{Tr}(\bar{\rho}^{(\Phi)} S).$$

In a practical implementation, the solution C is generalized from the symmetric eigenvalue problem (Eq. 1.61). Next, one can compute the matrix $\bar{\rho}^{(\Phi)}$ using Eq. 1.66. The electronic density is then given by

$$\rho(\mathbf{r}) = 2 \sum_{j,k=1}^{N} (\bar{\rho}^{(\Phi)})_{jk} \phi_j(\mathbf{r}) \phi_k(\mathbf{r}). \tag{1.67}$$

The band-structure energy can be rewriten as

$$\begin{aligned} 2\sum_{j=1}^{N} \epsilon_j f[(\epsilon_j - \mu)/k_B T] &= 2\mathrm{Tr}(\Lambda \rho^{(\Psi)}) \\ &= 2\mathrm{Tr}(H^{(\Phi)} \bar{\rho}^{(\Phi)}). \end{aligned} \tag{1.68}$$

This method also uses Mehrstellen finite difference discretization to improve the accuracy and multrigrid techniques to enhance the convergence rate. The details are in Ref. [66]. In Chapter. 4, we use this new method to carry out quantum transport calculations. This Order-N method can handle a very large system, includeing transition metal atoms. It does show nearly linear behavior.

1.8 Summary

The set of approximations introduced in this chapter form the core of the *ab initio* total energy calculation method, which is widely used by condensed matter physicists. Detailed explanations were presented for the Born-Oppenheimer, DFT, LDA and GGA, and pseudopential concepts: Born-Oppenheimer approximation separates the nuclear and electronic motion (Sec. 1.1); DFT reduces the many body system to a single-partical system with an effective nonlocal potential, and the LDA and the GGA provide remarkable approximations to their potential (Sec.1.2); Finally the pseudopotential theory allows one to separate valence and core electrons, replacing the strong electron-ion potential with a much weaker pseudopotential (Sec. 1.3). The above approximations allow the total-energy calculations to be performed accurately and efficiently. In order to extend this method to the study of nonperiodic systems, the periodic supercell approximation was introduced in Sec. 1.4. Popular and commonly used basis sets, like plane-wave basis and real-space basis were reviewed (Sec. 1.5). Finally, two real-space methods extensively involved in this thesis: the multigrid DFT method and the optimized nonorthogonal orbital order-N DFT method, were discussed in Sec.1.6 and Sec.1.7, respectively.

Chapter 2

Implementation of Ultrasoft pseudopotentials in Real-Space Method

There has been a rapid growth in the calculations of solid-state properties within density functional theory since the development of first-principles norm-conserving pseudopotentials [30, 31, 33] in the last three decades. As we mentioned in Chapter.1.3.2, the condition of norm-conservation (fourth constraint) ensures that the charge inside a cutoff radius r_c of the pseudo wavefunction is equal to that of the *all-electron* wavefunction. However, this condition, in certain cases, e.g., O $2p$ or Ni $3d$ orbitals [35], becomes an obstacle to construct a pseudo wavefunction significantly smoother than the *all-electron* wavefunction. The high localization of the valence p orbitals

of electronegative first row atoms such as F [76], and the valence d orbitals of the transition-metal atoms, causes the same problems. In the early 1990s, a new method to generate the ultrasoft pseudopotentials (UPPs) was introduced by Vanderbilt [35]. In his scheme, the norm-conserving condition is relaxed, and the pseudo wavefunctions are allowed to be as soft as possible within the core region, yielding a dramatic reduction of the cutoff energy required to describe them. The UPPs have already been used in plane-wave based methods and their successful applications open the door to electronic structure calculations for systems containing first-row elements [77, 78, 79] or transition metals [80]. However, these pseudopotentials have not yet been implemented in the real-space methods, which limits the capability of the real-space method to deal with such systems. I implemented the Vanderbilt UPPs in our real-space multigrid method and in the nonorthogonal orbital order-N method. This implementation enables efficient calculations of the atomic structure and electron transport properties of self-assembled monolayers on Au(111) surface, described in the following chapter.

This chapter is organized as follows. In Section 2, we present a detailed introduction to the Vanderbilt UPP, showing its general properties and describing the differences with respect to the conventional norm-conserving schemes. In Section 3, we implement this new pseudopotential in our real-space code; we identify problems that arise and illustrate how these problems are overcome. Section 4 shows our test results for diatomics and transition atoms.

2.1 Introduction to Vanderbilt Ultrasoft Pseudopotentials

2.1.1 Generalized Kohn-Sham Equations

The Vanderbilt UPP is a generalization of the Kleinman-Bylander pseudopotentials. In the Vanderbilt scheme, the norm-conservation constraint is relaxed, i.e.,

$$\int_0^{\mathbf{r}_c} \psi_n^*(\mathbf{r})\psi_m(\mathbf{r}) - \int_0^{\mathbf{r}_c} \phi_n^*(\mathbf{r})\phi_m(\mathbf{r}) \neq 0 \,,$$

where $\psi_n(\mathbf{r})$ are the *all-electron* wavefunctions and $\phi_m(\mathbf{r})$ are the constructed ultrasoft pseudo wavefunctions. Such relaxation offers substantial flexibility in the construction of the pseudo wavefunctions that are smooth inside the core region (defined by $r < r_c$) and modifies the conventional approach significantly. Consequently, new augmentation functions $Q^I_{n,m}(\mathbf{r})$ are defined as

$$Q^I_{nm}(\mathbf{r}) = Q_{nm}(\mathbf{r} - \mathbf{R}_I) = \psi_n^*(\mathbf{r})\psi_m(\mathbf{r}) - \phi_n^*(\mathbf{r})\phi_m(\mathbf{r}) \,,$$

which describe the difference between true and pseudo valence orbitals. It is apparent that $Q^I_{n,m}(\mathbf{r})$ are strictly localized in the core region since they coincide beyond r_c. In order to recover the full electronic charge, the electron density given by the squared moduli of the wavefunctions is augmented in the core region as

$$n(\mathbf{r}) = \sum_i \left[|\phi_i(\mathbf{r})|^2 + \sum_{nm} Q^I_{nm}(\mathbf{r}) < \phi_i|\beta^I_n >< \beta^I_m|\phi_i > \right] . \tag{2.1}$$

Here $\beta^I_n = \beta_n(\mathbf{r} - \mathbf{R}_I)$ are projector functions that are centered on the ionic positions and vanish outside the core region. For UPP, often two projectors are required for

each angular momentum channel. Therefore the total number of projectors N_β is generally twice as large as that for a corresponding Kleinman-Bylander norm-conserving pseudopotential.

In addition, the orthonormality condition for the wavefunctions takes on a generalized form due to the relaxation of the norm-conserving condition,

$$\langle\phi_i|S(\{\mathbf{R}_I\})|\phi_j\rangle = \delta_{ij}\,, \tag{2.2}$$

where the Hermitian overlap operator $S(\mathbf{R}_I)$ depends on the ionic positions through the $|\beta_n^I>$ and is defined by

$$S = 1 + \sum_{nm,I} q_{nm}|\beta_n^I\rangle\langle\beta_m^I|\,, \tag{2.3}$$

with

$$q_{nm} = \int Q_{nm}(\mathbf{r})\,d\mathbf{r}\,. \tag{2.4}$$

In Vanderbilt's UPP scheme, the total energy functional is given by

$$\begin{aligned} E_{\rm tot}[\{\phi_i\},\{\mathbf{R}_I\}] &= \sum_i <\phi_i| -\frac{1}{2}\nabla^2 + V_{\rm NL}|\phi_i> + \frac{1}{2}\int\int {\rm d}\mathbf{r}{\rm d}\mathbf{r}'\frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \\ &\quad + E_{\rm xc}[n] + \int {\rm d}\mathbf{r} V_{\rm loc}^{\rm ion} n(\mathbf{r}) + E_{\rm ion}(\{\mathbf{R}_I\})\,, \end{aligned} \tag{2.5}$$

with

$$V_{\rm NL} = \sum_{nm,I} D_{nm}^{(0)}|\beta_n^I><\beta_m^I|\,, \tag{2.6}$$

where $D_{nm}^{(0)}$ are just parameters which characterize the UPP. Vanderbilt has proved that the minimization of this total energy functional subject to the new constraint (Eq. 2.2) directly leads to a generalized eigenvalue problem

$$H|\phi_i> = \epsilon_i S|\phi_i>\,, \tag{2.7}$$

where

$$H = -\frac{1}{2}\nabla^2 + V_{\text{eff}} + \sum_{nm,I} D^I_{nm} |\beta^I_n >< \beta^I_m| \,. \tag{2.8}$$

Here $V_{\text{eff}}(\mathbf{r})$ is the screened effective local potential

$$V_{\text{eff}}(\mathbf{r}) = V^{\text{ion}}_{\text{loc}}(\mathbf{r}) + V_{\text{H}} + V_{\text{XC}} \,, \tag{2.9}$$

and the coefficients D^I_{nm} of the non-local part of UPP are given by

$$D^I_{nm} = D^{(0)}_{nm} + \int V_{\text{eff}}(\mathbf{r}) Q^I_{nm}(\mathbf{r}) \, \mathrm{d}\mathbf{r} \,. \tag{2.10}$$

As we can see, the UPPs also have a fully separable form of the non-local pseudopotential, while the coefficients depend on the wavefunctions through V_{eff}, Eq. 2.10, and have to be updated at each self-consistent step. Due to this involvement of the non-local pseudopotential in the self-consistent screening process, UPPs improve the transferability with respect to changes in charge configuration. However, the total energy E_{tot} still can be obtained from Eq. 1.23 after the eigenvalues are computed from the generalized eigenvalue Eq. 2.7 .

The ultrasoft pseudopotential is thus fully determined by the quantities $V^{\text{ion}}_{\text{loc}}$, $D^{(0)}_{nm}$, $Q_n m(\mathbf{r})$, and $\beta_n(\mathbf{r})$. The detailed algorithm used to generate these quantities is described in Ref.[35]. Comparing with the KB pseudopotential, the relaxation of the norm-conserving condition results in the introduction of the S overlap operator, the generalization of eigenvalue equation, the wavefunction dependence of the coefficients D^I_{nm}, and the increase of the number N_β of β^I_n functions, but it makes the wavefunctions much smoother inside the core region and thence allows for lower cutoff

energies to represent them. Typically, the cutoff energy is about half that of conventional norm-conserving pseudopotentials when using ultrasoft pseudopotentials. Such advantage puts the UPPs on the list of the most widely used pseudopotentials in the modern condensed matter calculations.

2.1.2 Ionic Forces

Following the Car-Parrinello method [47], the electronic wavefunctions and the ionic coordinates evolve according to a classical Lagrangian

$$\mathcal{L} = \mu \sum_i \int \mathrm{d}\mathbf{r} |\dot{\phi}_i(\mathbf{r})|^2 + \frac{1}{2} \sum_i M_I \dot{\mathbf{R}}_I^2 - E_{\mathrm{tot}}(\{\phi_i, \mathbf{R}_I\}) \,,$$

subject to a set of constrains

$$\mathcal{N}_{ij}(\{\phi_i, \mathbf{R}_I\}) = < \phi_i | S(\{\mathbf{R}_I\}) | \phi_j > - \delta_{ij} = 0 \,. \tag{2.11}$$

Here μ is a fictitious mass parameter for the electronic degrees of freedom and M_I is the mass of the atoms. Thus, the Euler equations of motion are derived as

$$\mu \ddot{\phi}_i = -\frac{\delta E_{\mathrm{tot}}}{\delta \phi_i^*} + \sum_j \Lambda_{ij} S \phi_j \,, \tag{2.12}$$

$$\mathbf{F}_I = M_I \ddot{\mathbf{R}}_I = -\frac{\partial E_{\mathrm{tot}}}{\partial \mathbf{R}_I} + \sum_{ij} \Lambda_{ij} < \phi_i | \frac{\partial S}{\partial \mathbf{R}_I} | \phi_j > \,. \tag{2.13}$$

Here Λ_{ij} are the Lagrange multipliers associated with the constraints (Eq. 2.11), and are given by

$$\Lambda_{ij} = < \phi_i | H | \phi_j > = \epsilon_i \delta_{ij} \,, \tag{2.14}$$

where the screened H is given in Eq. 2.8. It is explicit that the second term on the right-hand side of Eq. 2.13 is new with respect to the conventional norm-conserving

schemes. It corresponds to the contribution from the change in the orthonormality constrains due to the R_I dependence of the overlap operator $S(\{R_I\}$. Furthermore, in the first term on the right-hand side of Eq. 2.13, the electron density also depends on R_I through the Q^I_{nm} and β^I_n. Defining the following two quantities

$$\rho^I_{nm} = \sum_i < \phi_i|\beta^I_n >< \beta^I_m|\phi_i > , \tag{2.15}$$

$$\omega^I_{nm} = \sum_{ij} \Lambda_{ij} < \phi_j|\beta^I_n >< \beta^I_m|\phi_i > , \tag{2.16}$$

their derivatives with respect to R_I

$$\frac{\partial \rho^I_{nm}}{\partial \mathbf{R}_I} = \sum_i \left[< \phi_i|\frac{\partial \beta^I_n}{\partial \mathbf{R}_I} >< \beta^I_m|\phi_i > + < \phi_i|\beta^I_n >< \frac{\partial \beta^I_m}{\partial \mathbf{R}_I}|\phi_i > \right] , \tag{2.17}$$

$$\frac{\partial \omega^I_{nm}}{\partial \mathbf{R}_I} = \sum_{ij} \Lambda_{ij} \left[< \phi_j|\frac{\partial \beta^I_n}{\partial \mathbf{R}_I} >< \beta^I_m|\phi_i > + < \phi_j|\beta^I_n >< \frac{\partial \beta^I_m}{\partial \mathbf{R}_I}|\phi_i > \right] , \tag{2.18}$$

and noting that

$$\frac{\partial n(\mathbf{r})}{\partial \mathbf{R}_I} = \sum_{nm} \left[Q^I_{nm}(\mathbf{r}) \frac{\partial \rho^I_{nm}}{\partial \mathbf{R}_I} + \frac{dQ^I_{nm}(\mathbf{r})}{d\mathbf{R}_I} \rho^I_{nm} \right] , \tag{2.19}$$

the forces on the atoms can be expressed by

$$\begin{aligned} F_I = & -\frac{dE_{ion}}{d\mathbf{R}_I} - \int \mathrm{d}\mathbf{r} \frac{dV^{\mathrm{ion}}_{\mathrm{loc}}}{d\mathbf{R}_I} n(\mathbf{r}) - \int \mathrm{d}\mathbf{r} V_{\mathrm{eff}}(\mathbf{r}) \sum_{nm} \frac{dQ^I_{nm}(\mathbf{r})}{d\mathbf{R}_I} \rho^I_{nm} \\ & - \sum_{nm} D^I_{nm} \frac{\partial \rho^I_{nm}}{\partial \mathbf{R}_I} + \sum_{nm} \frac{\partial \omega^I_{nm}}{\partial \mathbf{R}_I} , \end{aligned} \tag{2.20}$$

where D^I_{nm} and V_{eff} have been defined in Eqs. 2.9 and 2.10.

2.2 Implementation of UPP into our Real-Space Methods

2.2.1 Implementation of Two Uniform Global Grids

No doubt, the UPP provides optimally smooth pseudo wavefunctions, but at the same time it also causes the augmentation functions $Q_{nm}(\mathbf{r})$ to be hard in the core regions, typically for the tightly bound d-electron states of transition metals. As a result, the total valence electron density $n(\mathbf{r})$ expressed in Eq. 2.1 is partitioned into a soft delocalized contribution given by the squared moduli of the wave functions, and a hard contribution localized at the cores. In our real-space methods with KB pseudopotential, the wavefunction and electron density are mapped onto the same uniform grid. This partition does not work efficiently for the case of UPP anymore. The soft part of the density can be accurately expanded on the grid for the wavefunctions, which is generally coarser than the corresponding one used for the conventional pseudopotentials, because UPP allows for optimally soft wavefunctions. However, a much dense grid is required to represent the augmentation electron density (the second part on the right-hand side of Eq. 2.1). In order to apply the UPP in our real space method efficiently, two different uniform global grids are used. One is the coarse grid for wavefunctions on which the generalized Kohn-Sham Eq. 2.7 is solved. The other is the fine grid for the augmentation charge density. Typically, the fine grid is twice denser than the coarse one.

In the actual calculation of $n(\mathbf{r})$, the augmentation charge related to $Q_{nm}(\mathbf{r})$ is evaluated on the fine grid. The coarse grid is used to calculate the soft part of the electron density. Then, it is transformed to the fine grid by performing proper interpolation from the values on the coarse-grid points and added to the augmentation charge to get the whole electron density $n(\mathbf{r})$. In our case the cubic interpolation is employed. Here we provide the interpolation formula in the one-dimensional case,

$$\begin{aligned} f_{nI+s} &= -\frac{(x_{nI+s}-X_I)(x_{nI+s}-X_{I+1})(x_{nI+s}-X_{I+2})}{6h_{\rm c}^3}\mathcal{F}_{I-1} \\ &\quad +\frac{(x_{nI+s}-X_{I-1})(x_{nI+s}-X_{I+1})(x_{nI+s}-X_{I+2})}{2h_{\rm c}^3}\mathcal{F}_I \\ &\quad -\frac{(x_{nI+s}-X_{I-1})(x_{nI+s}-X_I)(x_{nI+s}-X_{I+2})}{2h_{\rm c}^3}\mathcal{F}_{I+1} \\ &\quad +\frac{(x_{nI+s}-X_{I-1})(x_{nI+s}-X_I)(x_{nI+s}-X_{I+1})}{6h_{\rm c}^3}\mathcal{F}_{I+2}\,, \end{aligned} \tag{2.21}$$

with

$$n = h_{\rm c}/h_{\rm d} \qquad \text{and} \qquad s = 1, 2, \cdots, n-1\,,$$

where $h_{\rm d}$ and $h_{\rm c}$ are the grid spacings of dense-grid points and coarse-grid points respectively, f_{nI+s} are the values of the function on dense-grid points x_{nI+s}, and $\mathcal{F}_I$ is the values on coarse-grid points X_I. The Hartree potential is obtained by solving the Poisson's equation $-\nabla^2 V_{\rm H} = 4\pi n(\mathbf{r})$. The exchange-correlation potential $V_{\rm XC}$ is also evaluated on the fine grid with LDA or GGA. The number of operations needed for the above procedures does not depend on the number of states in the system, but only on the grid size. This cost is relatively small when compared to that of the eigenfunctions, especially for large systems consisting of hundreds of atoms.

The potential $V_{\rm eff}$, Eq. 2.9, is now known on the fine grid and used to calculate all

quantities, including D^I_{nm} and the third part of the ionic force on the right side of Eq. 2.13. However, both of the integrals are localized in the core region due to Q^I_{nm} and the overall cost of these operations is negligible. The $V_{\rm eff}$ is also needed on the coarse grid, to calculate $V_{\rm eff}\phi_i$ when solving the generalized Kohn-Sham equations, Eq. 2.7. Therefore, $V_{\rm eff}$ needs to be transferred to the coarse grid. The simplest way is just picking up the values corresponding to the coarse-grid points from its expansion on the dense grid. We find that this extracting method works well in our real-space total-energy calculation.

2.2.2 Double-Grid Technique for Nonlocal Projectors

In Vanderbilt's UPP scheme, all the computational procedures, including solving the Kohn-Sham equations, orthonormalizing the wavefunctions, constructing the electron density and finding the ionic forces, require performing inner products between the wavefunctions and the nonlocal projectors, or their derivatives with the ionic positions. The accuracy of these inner products significantly affects the accuracy and the stability of the method. The UPPs offer much smoother wavefunctions, which can be represented well on a coarse grid. However, the nonlocal projectors inside the core regions vary rapidly. This means that the grid size suitable for the wavefunction expansion is not fine enough to map the nonlocal part of pseudopotentials correctly. In order to perform these inner products efficiently, a double grid technique [81] can be used, because of the following facts. First, the projectors need to be on a fine grid.

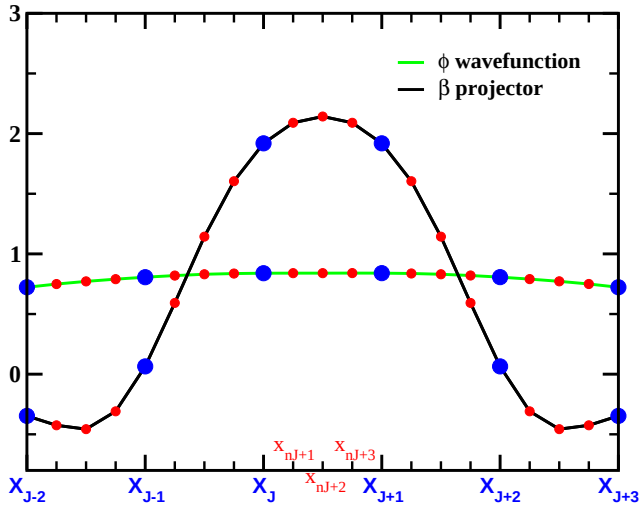

Figure 2.1: **Presentation of pseudo-wavefunction (black line) and the nonlocal projector (green line) with $l = 1$ for diamond. $X_J(x_j)$ represent coarse (dense)-grid points with $j = nJ + s(0 \leq s < n)$, and so $X_J = x_{nj}$.**

Second, they do not change during SCF interation. Third, the wavefunctions on the fine grid can be obtained by interpolating from the values on the coarse grid. We will discuss the double grid technique in the next paragraphs.

The double grid employed to handle this problem consists of two types of uniform and equi-distant grids. The coarse grid is just the global grid for the wavefunctions, and the dense grid used to represent the non-local projectors and is localized in a cubic box (with side L) in the core region of an atom. The procedure is best illustrated with help of Fig.2.1. The black and green solid lines show a wavefunction and a nonlocal projector for diamond calculations. Their corresponding values on coarse- and dense-grid points are marked with big blue and small red solid circles, respectively. We see that the wavefunction is much smoother than the projector and

it could be easily approximated with interpolation. The projector, on the other hand, should be calculated using the fine grid points.

Assuming that the inner product can be accurately approximated by a discrete sum over dense-grid points,

$$< \beta^I|\phi >= \oint_{\text{cubic box}} \beta^I(\mathbf{x})\phi(\mathbf{x})\mathrm{d}\mathbf{x} \approx \sum_{j=0}^{n^3 N_{\text{core}}} \beta_j^I \phi_j \triangle V_{\text{dens}} \,, \qquad n = \frac{h_{\text{coar}}}{h_{\text{dens}}} \,, \tag{2.22}$$

where $h_{\text{dens}}(h_{\text{coar}})$ is the grid spacings of dense (coarse) grid points, $\triangle V_{\text{dens}}$ ($\triangle V_{\text{coar}}$) is the dense (coarse) grid volume element, and $n^3 N_{\text{core}}$ (N_{core}) is the number of dense (coarse) grid points in the core region. We will denote the values of wavefunction on the dense grid by ϕ_j and on the coarse grid with Φ_J. Now we proceed as follows

$$\begin{aligned} \sum_{j=0}^{n^3 N_{\text{core}}} \beta_j^I \phi_j \triangle V_{\text{dens}} &= \sum_{j=0}^{n^3 N_{\text{core}}} \left(\sum_{J=0}^{N_{\text{core}}} F_{jJ} \Phi_J \right) \beta_j^I \triangle V_{\text{dens}} \\ &= \sum_{J=0}^{N_{\text{core}}} \Phi_J \sum_{j=0}^{n^3 N_{\text{core}}} F_{jJ} \beta_j^I \triangle V_{\text{dens}} \,, \end{aligned} \tag{2.23}$$

where F_{jJ} is the interpolation function. As a first step, we have approximated the wavefunction ϕ_j on the dense grid by its known coarse grid values Φ_J. There are several interpolation methods for this task. We apply the Fourier interpolation, i.e.,

$$\phi_j = \sum_{k=0}^{N_{\text{core}}} \tilde{\Phi}(\mathbf{G}_k) e^{i\mathbf{G}_k \cdot \mathbf{x}_j} \,, \tag{2.24}$$

with

$$\tilde{\Phi}(\mathbf{G}_k) = \frac{1}{L^3} \sum_{J=0}^{N_{\text{core}}} \Phi_J e^{-i\mathbf{G}_k \cdot \mathbf{X}_J} \triangle V_{\text{coar}} \,. \tag{2.25}$$

The Fourier interpolation function F_{jJ} can be expressed by

$$F_{jJ} = F_J(\mathbf{x}_j) = \frac{1}{L^3} \sum_{k=0}^{N_{\text{core}}} e^{i\mathbf{G}_k \cdot (\mathbf{x}_j - \mathbf{X}_J)} \triangle V_{\text{coar}} \,. \tag{2.26}$$

We still refine Eq. 2.23 as

$$\sum_{J=0}^{N_{\rm core}} \Phi_J \sum_{j=0}^{n^3 N_{\rm core}} F_{jJ}\beta_j^I \triangle V_{\rm dens} = \sum_{J=0}^{N_{\rm core}} \Phi_J \left(\sum_{j=0}^{n^3 N_{\rm core}} F_{jJ}\beta_j^I \frac{\triangle V_{\rm dens}}{\triangle V_{\rm coar}} \right) \triangle V_{\rm coar}$$
$$= \sum_{J=0}^{N_{\rm core}} \Phi_J \omega_J^I \triangle V_{\rm coar} \, . \qquad (2.27)$$

We can observe that the inner products (Eq. 2.22) still can be approximated with the summation over coarse-grid points inside the core region, but the original nonlocal projectors β_J^I are simply substituted by the following weight factors:

$$\omega_J^I = \sum_{j=0}^{n^3 N_{\rm core}} F_{jJ}\beta_j^I \left(\frac{\triangle V_{\rm dens}}{\triangle V_{\rm coar}} \right) \, . \qquad (2.28)$$

In our case of Fourier interpolation, the weightfactors are

$$\omega_J^I = \sum_{k=0}^{N_{\rm core}} \tilde{\omega}_k e^{i\mathbf{G}_k \cdot \mathbf{X}_J} e^{-i\mathbf{G}_k \cdot \mathbf{R}_I} \, , \qquad (2.29)$$

with

$$\tilde{\omega}_k = \frac{1}{L^3} \sum_{j=0}^{n^3 N_{\rm core}} \beta_j e^{-i\mathbf{G}_k \cdot \mathbf{x}_j} \triangle V_{\rm dens} \, . \qquad (2.30)$$

Here $\beta_j = \beta(\mathbf{x}_j)$, assuming that the atom is at the origin in Cartesian coordinates. An advantage of the Fourier interpolation is that the term representing the position of the atom can be factorized in the expression of ω_J^I as the structural phase factor $\exp[-i\mathbf{G}_k \cdot \mathbf{R}_I]$. Thus the calculation of making a table of $\tilde{\omega}_k$ in Eq. 2.30 for each atomic species has to be carried out only once at the entire job.

It is explicit that the weight factors are independent of the wavefunctions and depend only on the known values of nonlocal projectors on dense-grid points. Typically, the ratio of $h_{\rm coar}$ to $h_{\rm dens}$ is fixed at 3 or 4. The calculation of ω_J^I has to be repeated

each time the ions are moved. The gradients of the projectors needed in the force term of Eq. 2.13 are similarly replaced by the interpolation weights. However, we do not have to recalculate them through the self-consistent iteration steps.

2.2.3 Six-Order Mehrstellen Discretization

Generally, the cutoff energy is about 30 Ry or less when the UPP is used. The corresponding coarse grid spacing is as large as 0.21 Å, which is much bigger than that of KB pseudopotentials. When we discretize the generalized Kohn-Sham equations on such coarse grid, i.e.,

$$\mathbf{H}_{\text{Mehr}}[\phi_n] = \frac{1}{2}\mathbf{A}_{\text{Mehr}}[\phi_n] + \mathbf{B}_{\text{Mehr}}[V_{\text{eff}}\phi_n] = \epsilon_n \mathbf{B}_{\text{Mehr}}[S\phi_n]\,, \tag{2.31}$$

the fourth-order Mehrstellen discretization ($O(h^4)$) method is not accurate enough to approximate the differential operator in the Hamiltonian. Here we use a sixth-order Mehrstellen discretization, which uses one more neighbor than the fourth-order one.

The sixth-order Mehrstellen discretization samples the Hamiltonian and the wave-

Table 2.1: **Discretization weights for the sixth-order orthorhombic Mehrstellen operators for the central, first, second, third, fourth, and fifth nearest-neighbor grid points. The cubic-grid operator corresponds to** $h_i = h_{\rm grid}$

	a	b_n	$c_{n,m}$	$d_{n,m,l}$	e_n	$f_{n,m}$
$\mathbf{A}_{\rm Mehr}$	$\sum_i \frac{116}{90h_i^2}$	$-\frac{49}{60h_n^2} + \sum_i \frac{31}{80h_i^2}$	$-\frac{1}{10h_n^2} - \frac{1}{10h_m^2} + \sum_i \frac{1}{72h_i^2}$	$-\sum_i \frac{1}{140h_i^2}$	$\frac{1}{120h_n^2} - \sum_i \frac{1}{120h_i^2}$	$\frac{1}{240h_n^2}$
$\mathbf{B}_{\rm Mehr}$	$\frac{61}{120}$	$\frac{13}{180}$	$\frac{1}{144}$	0	$-\frac{1}{240}$	0

functions at 57 points for both orthorhombic and cubic grids:

$$
\begin{aligned}
\mathbf{A}_{\rm Mehr} &= af(\mathbf{x}) + \sum_{n=1}^{3} b_n f(\mathbf{x} \pm h_n \hat{\mathbf{x}}_n) + \sum_{n<m} c_{n,m} f(\mathbf{x} \pm h_n \hat{\mathbf{x}}_n \pm h_m \hat{\mathbf{x}}_m) \\
&\quad + \sum_{n<m<l} d_{n,m,l} f(\mathbf{x} \pm h_n \hat{\mathbf{x}}_n \pm h_m \hat{\mathbf{x}}_m \pm h_l \hat{\mathbf{x}}_l) + \sum_{n=1}^{3} e_n f(\mathbf{x} \pm 2h_n \hat{\mathbf{x}}_n) \\
&\quad + \sum_{n<m} f_{n,m} f(\mathbf{x} \pm h_n \hat{\mathbf{x}}_n \pm 2h_m \hat{\mathbf{x}}_m) \,, \\
\mathbf{B}_{\rm Mehr} &= a' f(\mathbf{x}) + \sum_{n=1}^{3} b'_n f(\mathbf{x} \pm h_n \hat{\mathbf{x}}_n) + \sum_{n<m} c'_{n,m} f(\mathbf{x} \pm h_n \hat{\mathbf{x}}_n \pm h_m \hat{\mathbf{x}}_m) \\
&\quad + \sum_{n=1}^{3} e'_n f(\mathbf{x} \pm 2h_n \hat{\mathbf{x}}_n) \,.
\end{aligned}
$$

Such expansions of $\mathbf{A}_{\rm Mehr}$ and $\mathbf{B}_{\rm Mehr}$ show $-\nabla^2 = \mathbf{B}_{\rm Mehr}^{-1}\mathbf{A}_{\rm Mehr} + O(h^6)$. Table.2.1 lists the corresponding weights of the sixth-order Mehrstellen operators used in our present code.

Here we want to point out that the fourth-order Mehrstellen operators are still used to discretize the Poisson equation, since both Hartree potential and electronic charge density are represented on the dense grid.

2.3 Test Results

In this section we examine the performance of UPPs in our real-space codes through calculations of several properties of molecules and bulk crystals. To facilitate comparisons with plane-wave calculations, an equivalent coarse-grid energy cutoff $E_{\rm cut}^{\rm coar} = \pi^2/2h^2$ is defined, which is equal to that of a plane-wave calculation that uses a FFT grid with the same grid spacing. For these tests, the size of the global dense grid for the potentials and the electronic charge density is eight times as big as that of the coarse grid, and for the calculation of the weight factors of nonlocal projectors the ratio $n = 4$ is fixed.

2.3.1 Bond Lengths of Diatomic molecules

The first test application is the bond length of diatomic molecules. The simulation supercell is a cube of length 15 a.u. and the grid spacing is 0.421 a.u., corresponding to an energy cutoff of 25 Ry. The k-space sampling was restricted to the Γ point. After the converging the electronic system, the ions are relaxed using the same relaxation scheme as before. The relaxed bond lengths of N_2, O_2, and CO are listed in Table.2.2. For comparison, it also lists the experimental values and the plane-wave calculated values. Our calculated bond lengths agree with the experimental data and the plane-wave calculations very well.

Table 2.2: **Bond lengths of diatomic molecules. The experimental data are from [82]. The theoretical results are from our present calculation and from plane-wave calculation [79, 83] with ultrasoft pseudopotentials.**

	$\mathbf{N_2}$	$\mathbf{O_2}$	**CO**
Experiment	2.07	2.28	2.13
This work	2.08	2.32	2.16
Plane-wave calculation	–	2.33	2.132

2.3.2 Ground state of bulk gold

To illustrate the ability of our codes to handle the transition metals, test calculations are performed on a 16-atom Au tetragonal supercell. 64 special k-points are included in k-space sampling. The ultrasoft pseudopotential for Au uses GGA [10] in the PBE form [23] for the exchange and correlation term. Partial core correction is added in order to account for nonlinear effects [32].

We calculated the total energy as a function of volume and then fitted the Murnaghan equation of state [84]

$$E(V) = \frac{B_0 V}{B_0'(B_0'-1)}\left[B_0'\left(1-\frac{V_0}{V}\right)+\left(\frac{V_0}{V}\right)^{B_0'}-1\right]+E(V_0) \qquad (2.32)$$

to our calculated points to extract the equilibrium volume V_0, the equilibrium bulk modulus B_0, and the derivative of the bulk modulus with pressure B_0'. The equilibrium lattice constant can be easily obtained from V_0, since the bulk gold has the fcc structure. In Fig.2.2, we display the calculated energy versus volume equation of state with an energy cutoff of 20 Ry. The fit to the Murnaghan equation of state is excellent. Table.2.3 lists the extracted lattice constant and the bulk modulus with

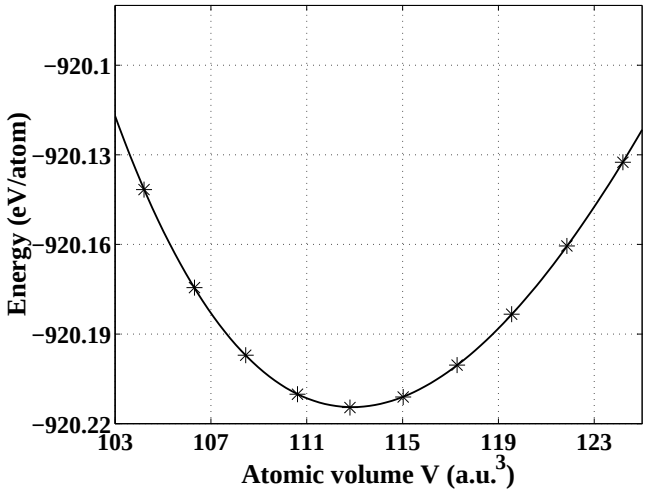

Figure 2.2: **Equation of state for bulk gold in fcc structure. The continuous curve is the Murnaghan equation of state; the points are the calculated energies at the indicated volume.**

different energy cutoffs. Our results are in good agreement with those of experiments and the plane-wave methods.

Table 2.3: **Ground-state properties of bulk gold from a fit to the Murnaghan equation of state. The experimental data [85] and the results from a plane-wave calculation within density-functional perturbation theory (DFPT) [83] are also listed for comparison.**

	Energy cutoff (Ry)	Lattice constant (a.u)	Bulk modulus (Kbar)
Experiment	–	7.67	1720
Present work	20	7.66	1919
Present work	30	7.67	1797
Plane-wave method	20	7.66	1823

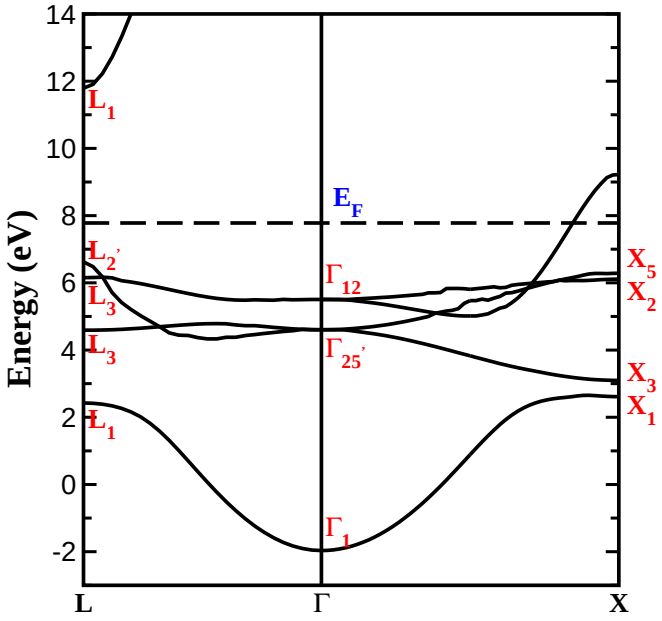

Figure 2.3: **Bulk copper DFT-LDA band structure calculated with ultrasoft pseudopotential.**

2.3.3 Ground state and band structure of bulk copper

The second test application for transition metals is the band structure of bulk copper. Generally, for a norm-conserving pseudopotential, an energy cutoff of 300 Ry [86] is necessary to handle copper due to its highly localized $3d$ wavefunctions. Recently, several smooth pseudopotentials have been generated for copper, which allows one to use energy cutoffs ranging from 60 Ry to 72 Ry. Within the ultrasoft scheme, the energy cutoff can be reduced to 25 Ry in our real-space calculation.

First, a band structure calculation is performed within the local-density approximation (LDA) for the exchange and correlation energy. The calculated band structures along $\Gamma-$L and $\Gamma-$X are plotted in Fig. 2.3. In Table. 2.4, we list our calculated

Table 2.4: **Comparison of Cu bandwidths at high-symmetry points with the values from experiment measurement [87] and the plane-wave method calculation [88] with norm-conserving pseudopotential (NCPP). All energies are in eV.**

		Present work (with UPP)	Plane-wave method (with NCPP)	Experiment
Widths of d bands	$\Gamma_{12} - \Gamma_{25'}$	0.901	0.91	0.81
	$X_5 - X_3$	3.187	3.23	2.79
	$X_5 - X_1$	3.672	3.70	3.17
	$L_3 - L_3$	1.562	1.58	1.37
	$L_3 - L1$	3.728	3.72	2.91
L-gap	$L_1 - L_{2'}$	5.168	4.21	5.20

bandwidths at high symmetry points. All of them are in very good agreement with recent plane-wave calculations using norm-conserving pseudopotential with energy cutoffs of over 60 Ry. The experimental values are listed in the right most column for reference.

In Table. 2.5, we report the fcc lattice constant and bulk modulus of bulk copper

Table 2.5: **Calculated lattice constant and bulk modulus of bulk copper from a fit to the Murnaghan equation of state. The experimental data [89] and the results from the plane-wave calculation within DFPT [83] are also listed for comparison.**

	Energy cutoff (Ry)	Lattice constant (a.u)	Bulk modulus (Kbar)
Experiment	–	6.81	1380
Present work	25	6.76	1572
Plane-wave method	20	6.72	1659

obtained from a fit with the Murnaghan equation of state for the total energy as a function of volume. The same supercell and k-points sampling as for bulk gold are used. They are in good agreement with plane-wave calculations and experimental data.

2.4 Summary

In this chapter, the new and novel ultrasoft pseudopotential scheme proposed by Vanderbilt was briefly introduced. The UPPs provide much smoother pseudo-wavefunctions through relaxing the norm-conserving condition. With respect to the conventional pseudopotentials, the major difference resides in the presence of the overlap operator S, the augmented electronic charge density $n(\mathbf{r})$, the wavefunction dependence of the D^I_{nm}, and the substantially modified ionic forces F_I. These differences require additional care and effort when implementing the UPPs in our real-space methods. First, in order to accurately represent the augmentation functions, an additional global dense grid is introduced, besides the one for the wavefunctions. Second, a double-grid technique is employed to perform the inner products between the rapidly varying nonlocal projectors of the pseudopotentials and the relatively smooth wavefunctions. The memory requirements and the computational cost are drastically reduced. Finally, sixth-order Mehrstellen operators are used to discretize the generalized Kohn-Sham equations. Our test applications on properties of diatomic molecules, bulk gold and bulk copper show very good agreements with the experimental data and

the plane-wave calculations. The advantage of a low energy cutoff allows us to perform calculations for large systems containing transition metal atoms, such as large supercells containing organic molecules adsorbed on a Au(111) surface.

Chapter 3

Nanowire-induced optical anisotropy of the Si(111)-In surface

Ab initio calculations of the reflectance anisotropy of Si(111)-In surfaces are presented. A very pronounced optical anisotropy around 2 eV is predicted for the structural model proposed by Bunk *et al.* [Phys. Rev. B 59, 12228 (1999)] for the (4×1) reconstructed surface. The (4×2)/(8×2) reconstructed surface, induced by a slight distortion of the indium chains, is shown to result in a splitting of the 2 eV peak. The calculated results are in excellent agreement with recent polarized reflectance data acquired during the (4×1) $\longrightarrow$ (4×2)/(8×2) phase transition.

3.1 Introduction

Indium may self-assemble on Si(111) in a (4×1) pattern that is formed by long rows of In atoms. The study of this reconstruction goes back to the sixties [90], but it has intensified recently. On one hand, the chains formed by the In atoms may be the smallest known atomic wires in existence [91, 92], which makes them interesting in the technological context of downsizing the microelectronics into the nanoscale regime. On the other hand, one-dimensional electronic systems are expected to show interesting electronic features, such as spin-charge separation in a Luttinger liquid or Peierls instabilities (see, e.g., Ref. [93]). Indeed, the metallic (4×1) phase of the In chains was found to undergo a reversible Peierls-like transition below room-temperature, where the period along the chain is doubled, forming (4×2) and finally (8×2) reconstructions [94, 95].

On the basis of surface x-ray diffraction (SXRD) data [96] a detailed structural model for the Si(111)(4×1)-In surface has been developed, which also accounts for previous experimental findings. Its main features are zigzag chains of Si atoms alternating with zigzag rows of In atoms on top of an essentially bulk-like Si lattice (see Fig. 3.1 (a) and (b)). The Si chain is thus similar to the π-bonded chain of the Si(111)(2×1) reconstruction. Strong support for this structural model comes from a series of *first-principles* calculations [97, 98, 99]. The calculations not only confirmed the experimentally determined surface geometry, but also reproduced the dispersion of the measured surface electronic states [91] and the experimentally detected image

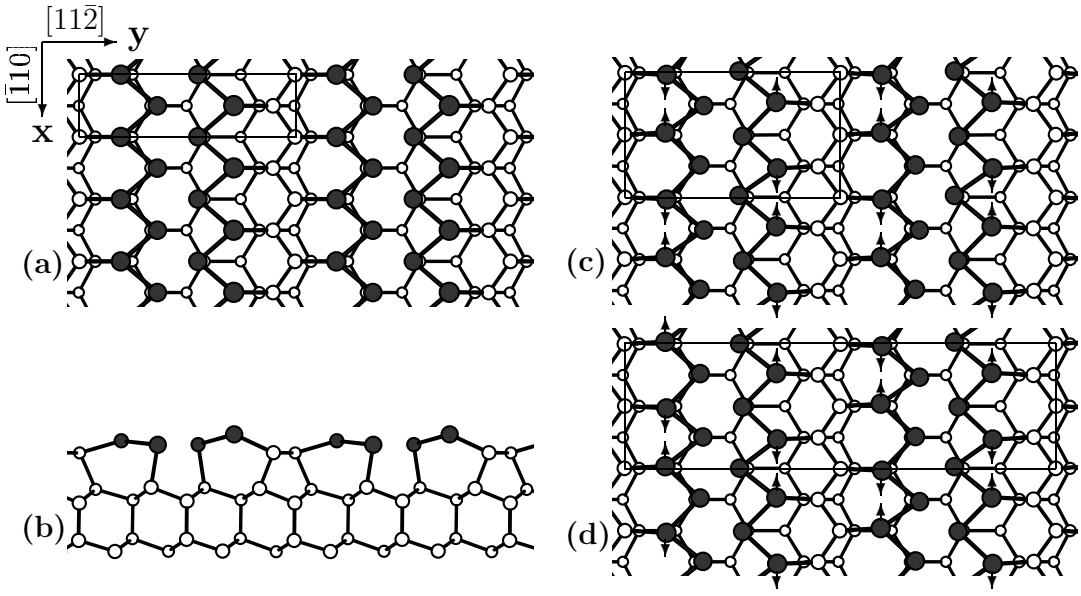

Figure 3.1: **Optimized atomic structures of the zigzag-chain model of the Si(111)-In surface: (a) top and (b) side view of the (4×1) reconstruction, (c) and (d) are the top views of (4×2) and (8×2) reconstructions, respectively. The solid/open circles are In/Si atoms. Surface unit cells are indicated. Arrows in (c) and (d) show the slight distortions of the outer indium atoms with respect to the (4×1) reconstruction.**

state anisotropy [100]. SXRD has also been used to study the low-temperature (LT) Si(111)-In surface [101]. It was proposed that the periodicity along the In chains doubles, due to trimer formation of the outer indium atoms. In addition, a glide line occurs, causing a doubling of the unit cell in the direction perpendicular to the chains. The resulting (8×2) reconstructed surface structure is somewhat in contrast to the atomic configuration changes upon cooling inferred from LT photoemission data [102]: Yeom and co-workers concluded that the inner In rows are mainly affected by the (4×1) $\longrightarrow$ (4×2)/(8×2) phase transition. However, this is not in agreement with

ab-initio calculations. Cho *et al.* [98] find that the outer indium chain atoms are displaced to form pairs, which are arranged in (4×2) or (8×2) symmetries as shown in Fig. 3.1 (c) and (d), respectively. The calculated structures for the (4×2)/(8×2) reconstruction agree in many but not all respects with the interpretation of the SXRD data in Ref. [101]. Discrepancies also exist regarding the electronic structure of the LT phase of the Si(111)-In surface. While some experimental studies find this phase to be insulating [94, 92], others see only a reduced density of states at the Fermi level [102, 103] or remain inconclusive [104]. *First-principles* calculations by the Kleinman group [98] find no gap opening at the Fermi level as a result of the (4×1) $\longrightarrow$ (4×2) phase transition. It is not clear, however, if this is indeed a correct finding or an artifact of the calculation, due to either the band gap underestimation in density-functional calculations [105], or an assumption of a wrong atomic structure for the LT phase of the Si(111)-In system.

Reflectance anisotropy spectroscopy (RAS) at the Si(111)(4×1)-In surface [106, 107, 108] shows an optical anisotropy in the energy region of 2 eV, which is significantly larger than previously reported for any semiconductor system. Very recently, it has been found that this anisotropy splits into two peaks, at 1.9 and 2.2 eV, upon formation of the LT phase of the Si(111)-In surface [109]. In this paper we present *first-principles* calculations for the nanowire-induced optical anisotropy of the Si(111)-In surface. It will be shown that the atomic structure proposed from SXRD data in Ref. [96] for the In induced (4×1) symmetry explains very well the optical anisotropy measured for the room-temperature phase of the Si(111)-In surface. Our total-energy

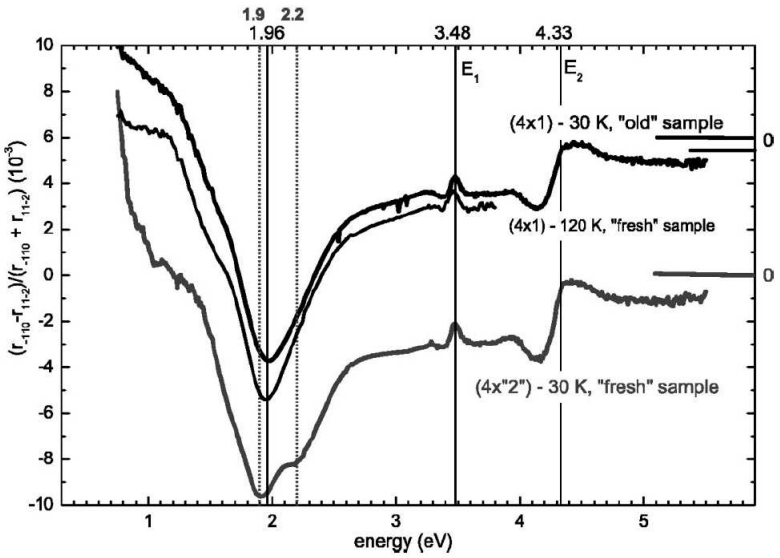

Figure 3.2: **The experimental observed RAS spectrum of the Si(111)-In 4×1 and 4×2/8×2 phases.**

and RAS calculations strongly support Kleinman's model [98] for the (4×2)/(8×2) reconstructed Si(111)-In surface.

3.2 Methodology

The calculations employ a massively parallel real-space multigrid implementation [64] of the density functional theory (DFT) within the local density approximation (LDA) [14]. The electron-ion interactions are described by non-local, norm-conserving pseudopotentials [31, 110, 111]. A partial core correction to the In pseudopotential was added in order to take into account a nonlinear effect for the exchange-correlation term [32]. The spacing of the grid used to map the wave functions, the potentials and

the charge density is 0.166 Å, corresponding to an energy cutoff of about 47 Ry in plane-wave calculations. The Si(111)-In surface is simulated by a periodic supercell that contains ten Si layers, one adsorbate layer, and a vacuum region corresponding in thickness to eight Si layers. Hydrogen atoms are employed to saturate the dangling bonds of the silicon atoms in the bottom layer. We use the calculated equilibrium lattice constant of 5.43 Å for silicon. Sets of special **k** points corresponding to 32 points in the full (1×1) surface Brillouin zone were used for the self-consistent electronic structure calculations.

The atomic surface structure gives rise to corrections ΔR in the Fresnel reflectivity R_0, which can be calculated from the diagonal slab polarizability components α_{ii} and the bulk dielectric function ϵ_b by

$$\frac{\Delta R(\omega)}{R_0(\omega)} = \frac{16\pi d\omega}{c}\mathbf{Im}\left[\frac{\alpha_{xx}(\omega) - \alpha_{yy}(\omega)}{\epsilon_b(\omega) - 1}\right],$$

where d is the slab thickness, ω is the frequency, and c the speed of light [112, 113]. Here x and y are the directions parallel and perpendicular to the indium chain, respectively, i.e., the $[\bar{1}10]$ and $[11\bar{2}]$ directions. The surface dielectric function calculations were performed with uniformly distributed **k** points corresponding to a density of 960 sampling points in the full (1×1) surface Brillouin zone. A linear cutoff function was used to eliminate spurious optical anisotropies from the bottom layer of the slab [114]. The dielectric function of metallic systems contains contributions from interband and intraband transitions [115]. The interband transitions dominate the spectra in the energy range usually investigated with RAS. Therefore, and because

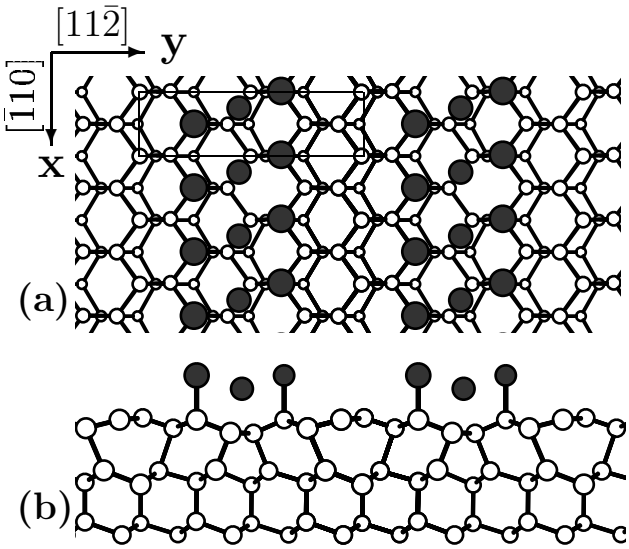

Figure 3.3: **The optimized atomic structure of the π-bonded chain stacking-fault model for Si(111)4×1-In reconstruction. Solid circles are In atoms and open circles are Si atoms. The (4×1) unit cell is indicated.**

the intraband contributions are hardly accessible to *ab-initio* calculations for systems of the size studied here, we consider only the interband part. It is calculated in the independent-particle approximation, i.e., neglecting excitonic and local-field effects. A scissors operator approach has been used to account for the band-gap underestimation of 0.5 eV for bulk Si due to the neglect of self-energy effects within DFT-LDA. Calculations for smaller systems [116, 117] have shown that many-body effects alter RAS spectra quantitatively rather than qualitatively, because RAS spectra are difference spectra, which are furthermore normalized to the bulk dielectric function. Therefore, calculations within the independent-particle approximation reliably reproduce experimental data for a wide range of semiconductors [118, 119].

3.3 Results and Discussion

The first reported RAS measurement of the Si(111)-In surface [106] was interpreted in terms of a π-bonded chain stacking-fault (π-SF) model of the (4×1) reconstruction. This model, shown in Fig. 3.3, was proposed on the basis of Auger spectroscopy, low-energy electron diffraction, and scanning tunneling microscopy [120]. The substrate Si atoms form 7656 rings, i.e., two six-fold, one seven-fold and one five-fold rings per unit cell. The In atoms on top form ridges contain three rows of In atoms. The edge In atoms saturate the dangling bonds of the underlying Si atoms and the central In row forms metallic bonds with the In atoms of the edge rows. We calculate an In-Si bond length of 2.66 Å. The minimum distance between In atoms amounts to 2.98 Å. The calculated optical anisotropy of the π-SF model is shown in Fig. 3.4. The RAS spectrum is negative for the complete energy range considered, i.e., from 1.0 to 5.5 eV. While the sign and the magnitude of the calculated optical anisotropy agree with experiments [106, 107, 108, 109], there is no agreement concerning the line shape and the energy positions of the RAS peaks. Experimentally, one very pronounced optical anisotropy dip at around 2 eV is observed. In contrast, the π-bonded chain stacking-fault model yields a series of features with the main peak at around 3.6 eV. From the calculated RAS spectrum we can thus exclude the π-SF model as the structure for In-induced (4×1) reconstructions of Si(111).

We therefore focus on the zigzag-chain model developed for interpretation of SXRD data of the Si(111)(4×1)-In surface [96]. Here, four indium atoms and two

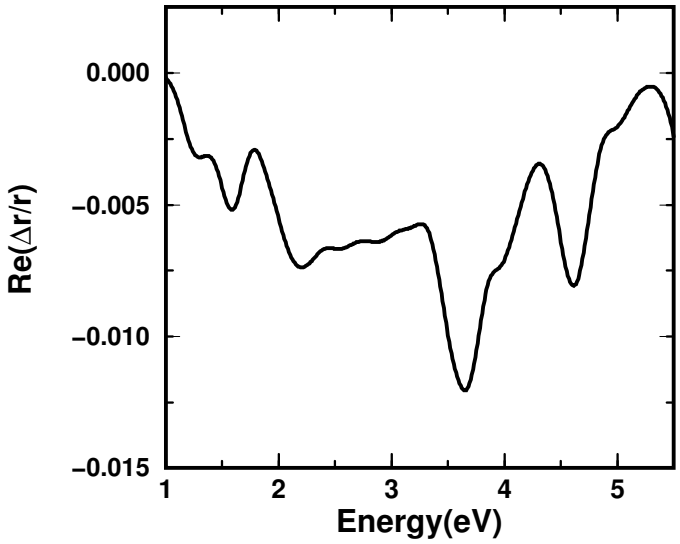

Figure 3.4: **RAS spectrum calculated for the π-SF model of the (4×1)-In reconstruction shown in Fig. 3.3.**

Si atoms per (4×1) unit cell are adsorbed on an essentially bulk-like (111) substrate surface, forming chains along the [$\bar{1}$10] direction. Starting from the experimentally determined coordinates, total-energy minimization yields the structure shown in Fig. 3.1 (a) and (b). The relaxed geometry is in good agreement with the experimental data [96] and previous calculations [97, 98, 99]. The In-In bond lengths within the chains amount to 2.98 Å and 3.00 Å. This is slightly larger than the sum of In covalent radii of 2.88 Å, but shorter than the 3.25 Å In-In distance in tetragonal bulk In, indicating the predominantly covalent character of bonding within the In chains. The distance between indium atoms in neighboring chains is 3.16 Å and the In-Si bond lengths are 2.68 Å. The calculated electronic structure of the Si(111)(4×1)-In surface confirms earlier experimental [91, 121] and theoretical results [97, 98, 99]. There are

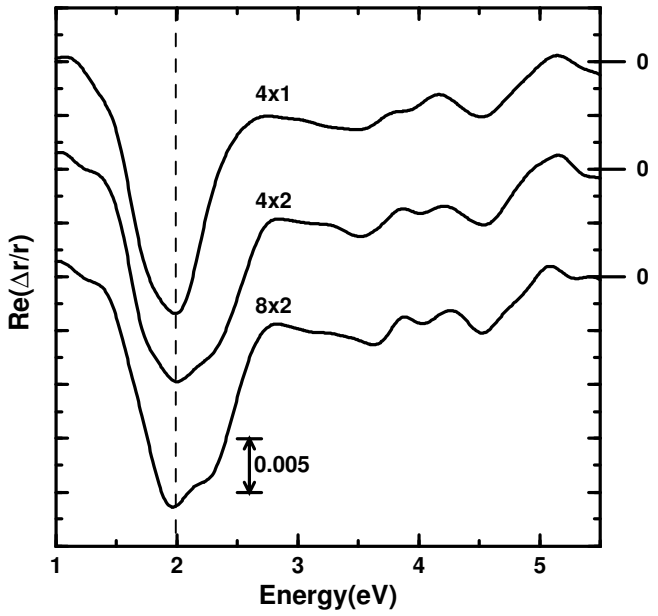

Figure 3.5: **RAS spectra calculated for the zigzag-chain models of the Si(111)-In surface reconstructions shown in Fig. 3.1. The curves are shifted vertically for clarity, with the zero position indicated on the right. The scale is shown in the inset.**

three metallic bands crossing the Fermi level along the direction parallel to the atomic chains, which show only a weak dispersion perpendicular to the In chains.

The anisotropy of the surface electronic structure is reflected in the calculated optical anisotropy, shown in Fig. 3.5. Similarly to the results obtained for the π-SF structure, we find the RAS to be negative for nearly the complete energy range considered. In the case of the zigzag-chain model, however, a very strong anisotropy peak of 2.2 % is calculated for a photon energy of 2 eV. This is in excellent agreement with the experimental findings [106, 107, 108, 109] of a very pronounced anisotropy of 1.2 – 2.0 % at 2 eV. The fact that the calculated optical anisotropy is slightly larger

than measured is expected. Our results refer to a single-domain Si(111)-In surface free of steps and other defects. Defects in the In chains and signals from minority domains reduce the peak at 2 eV, whereas surface steps give rise to additional features at the bulk critical points of Si, i.e., at 3.5 and 4.3 eV [122]. The optical anisotropy at 2 eV is unambiguously attributed to surface states, because this energy region is far below the direct optical gap of Si. However, it is not directly related to the metallicity of the nanowires, which at low frequencies is expected to result in a stronger optical coupling for light polarized in the chain direction rather than perpendicularly to the chains. Therefore, the quasi one-dimensional metallicity of the In chains should lead to positive optical anisotropies, which are indeed observed for photon energies below 1 eV [108]. Our calculations cannot safely be extended to that energy region, because of the neglect of intraband transitions.

In order to calculate the structure of the LT phase of the Si(111)-In surface, we start from symmetry-distorted geometries with (4×2) and (8×2) translational periodicities. The resulting (4×2) and (8×2) structures, shown in Fig. 3.1(c) and (d), are mainly characterized by a pairing of the outer In chain atoms. Their distance is reduced from 3.84 Å in the (4×1) structure to 3.55 Å. This reconstruction mechanism agrees with the findings of the Kleinman group [98], who predict a corresponding reduction from 3.87 to 3.59 Å. The pairing leads to an alternating relaxation of the inner chain atoms towards or away from the center of the neighboring indium chain. However, the pairing-induced In-In and In-Si bond length changes are hardly discernible, i.e., below 0.02 Å. The calculated band structures for the (4×2) and

(8×2) have no gap at the Fermi level. This may partially be related to the band gap underestimation typical of DFT-LDA calculations [105]. It appears likely that self-energy effects open a small gap for the lowest lying metallic band (cf. S_3 in Fig. 2 in Ref. [98]), because this band nearly bisects the (4×1) surface Brillouin zone. Therefore, the corresponding Fermi surface has the possibility of nesting, potentially driving the opening of a Peierls gap. This mechanism is hardly plausible, however, because of the remaining metallic surface states. Our calculations, which are in agreement with the *first-principles* results of Cho *et al.* [98], thus support the view that the doubling of the surface unit cell results in a reduction of the density of states at the Fermi level, but not in the opening of a fundamental gap. This is in agreement with some but not all of the recent experimental studies [102, 103].

In order to verify the calculated structural models for the (4×2)/(8×2) phase, we calculated their reflectance anisotropy, cf. Fig. 3.5. The spectra are similar to the one calculated for the (4×1) phase. However, for the (4×2) reconstruction a shoulder at 2.2 eV emerges, which is even more pronounced for the (8×2) surface. It corresponds exactly to the optical signature of the (4×1) $\longrightarrow$ (4×2)/(8×2) phase transition found in a recent experimental study [109]. In addition to the appearance of the 2.2 eV shoulder, the calculated minimum of the RAS shows a slight redshift by about 0.1 eV. Again, nearly quantitative agreement with the experimentally observed shift from 1.96 to 1.90 eV is obtained. Fleischer *et al.* [109] argue that the measured changes in the optical anisotropy cannot be explained as a temperature-induced sharpening of the original 2 eV peak, since the overall width of the structure is much larger for the

LT phase. The changes of the RAS spectra can thus only be explained by electronic and structural modifications of the Si(111)-In surface accompanying the (4×1) ⟶ (4×2)/(8×2) phase transition. The excellent reproduction of these changes in the calculated surface optical properties is a very strong indication for the correctness of the underlying surface structural model, i.e., the geometry originally proposed by Kumpf and co-workers [101] and modified by Cho *et al.* [98].

3.4 Summary and conclusions

The optical anisotropies for different models used to explain the formation of In nanowires at the Si(111) surface have been calculated from *first principles*. A comparison with measured data strongly supports the zigzag-chain model for the room-temperature (4×1) reconstructed phase of the Si(111)-In surface. An energetically favored pairing of In atoms gives rise to a doubling of the periodicity along the chain direction. The energy can be further lowered by arranging neighboring chains in a (8×2) superstructure. These structural changes are accompanied by changes in the calculated optical anisotropy, which are in excellent agreement with data acquired during the formation of the low-temperature phase of the Si(111)-In surface. Our results for the surface electronic structure are compatible with the view that a reduction of the density of states at the Fermi level occurs upon the (4×1) ⟶ (4×2)/(8×2) phase transition. However, there is no indication of gap opening.

Chapter 4

Resonant Coupling and Negative Differential Resistance in Metal/Ferrocenyl-Alkanethiolate/STM structures

Recent experimental studies have demonstrated that self-assembled molecules sandwiched between metallic contacts can perform logic functions based on negative differential resistance (NDR). To understand the mechanism of NDR, the electronic structure and transport properties of one such junction, ferrocenyl-alkanethiolate at-

tached to a gold surface and probed with a scanning tunneling microscope (STM) tip, have been investigated by large scale *ab initio* calculations. The I-V characteristics show strong NDR features at both positive and negative biases, in good agreement with the experimental data. The voltage-dependent transmission, potential drop profile, and molecular level alignment under bias suggest that the ferrocenyl group acts like a quantum dot and that the NDR features are due to resonant coupling between the HOMO and the density of states of gold leads. The strength of the individual NDR peaks can be tuned by changing the tunneling distance or using suitable spacer layers.

4.1 Introduction

Since the concept of using individual molecules as functional electronic devices was first proposed in 1970's [123], having individual molecules perform the basic functions of conventional electronic components became the goal of molecular electronics. Recently, this field has experienced a dramatic increase in activity, due to the emerging need for alternate routes toward smaller, faster and cheaper integrated circuits. Key electronic components, such as wires, diodes, and transistors have been successfully demonstrated, all based on single or self-assembled molecules sandwiched between electrodes. For self-assembled molecules on a single-crystal surface, the substrate serves as one of the electrodes. A variety of methods have been used to form the contact with the other electrode, such as a scanning tunneling microscope (STM) tip

[124, 125, 126, 127, 128] and atomic-scale break junctions [129, 130]. A number of novel and promising characteristics for these types of molecular devices have been reported. For example, negative differential resistance (NDR) [125, 131] and molecular memory effects [132, 133] have been demonstrated on a system of phenyl-ethylene oligomers functionalized with different side groups. Recently, a ferrocene-based molecular wire with near-perfect conductance has been reported by Getty and co-workers [134]. The ballistic conductance thus reaches the criteria for real applications of molecular electronic devices. The phenomenon of NDR has also been observed for ferrocenyl-undecanethiolate self-assembled monolayers (Fc-C_{11}S-SAM) on Au(111) surface [135, 136].

Self-assembled molecular systems with NDR are promising molecular devices due to their utility in fast switching logic and simplicity in integration. Therefore, theoretical understanding of their electron transport properties and of the underlying NDR mechanisms is critical for the design of molecular devices. However, in many cases the NDR mechanisms have not been unraveled either from experiments or from theory. The redox effect, *i.e.*, a change of the charge state with bias, has been suggested to explain the NDR for ferrocenyl-alkanethiolate SAM on Au(111) surface [135]. However, resonant tunneling may also play an important role in NDR. Another possibility is some level of bond breaking under current[137], *i.e.*, the structure of the molecule itself or of the molecule-substrate interface changes dramatically when a large bias is applied to the system.

In this chapter, we study the electronic and transport properties of the ferrocenyl-

pentanethiolate SAM on Au(111) surface. The atomic structures of Fc-C_5S and C_5S on the Au(111) surface and their saturation coverages are determined by total energy calculations. Our self-consistent quantum transport calculations for geometry-optimized structures show NDR features at both negative and positive biases, in good agreement with the experimental data. The mechanism responsible for the NDR is uncovered by analyzing the molecular level alignment, transmission spectra under bias, and charge transfer between the molecule and STM tip.

4.2 Electronic Structure Calculation

The calculations use a massively parallel real-space multigrid implementation [64] of density functional theory (DFT) [10]. Due to the relatively large sizes of the super-cells, the k-space sampling is restricted to the Γ point. The exchange and correlation terms are represented by the generalized gradient approximation (GGA) in PBE form [23]. The electron-ion interactions are described by nonlocal, ultrasoft pseudopotentials [35]. A partial core correction is added to the Au and Fe pseudopotentials in order to account for the nonlinearity in the exchange-correlation term [32]. The wave functions and localized orbitals are described on a grid with spacing of 0.20 Å. A double grid technique [81] is employed to evaluate the inner products between the non-local potentials and the wave functions, thereby substantially reducing the computational cost and memory requirement without losing accuracy.

First, the equilibrium structures of the adsorbed molecules on the Au(111) surface

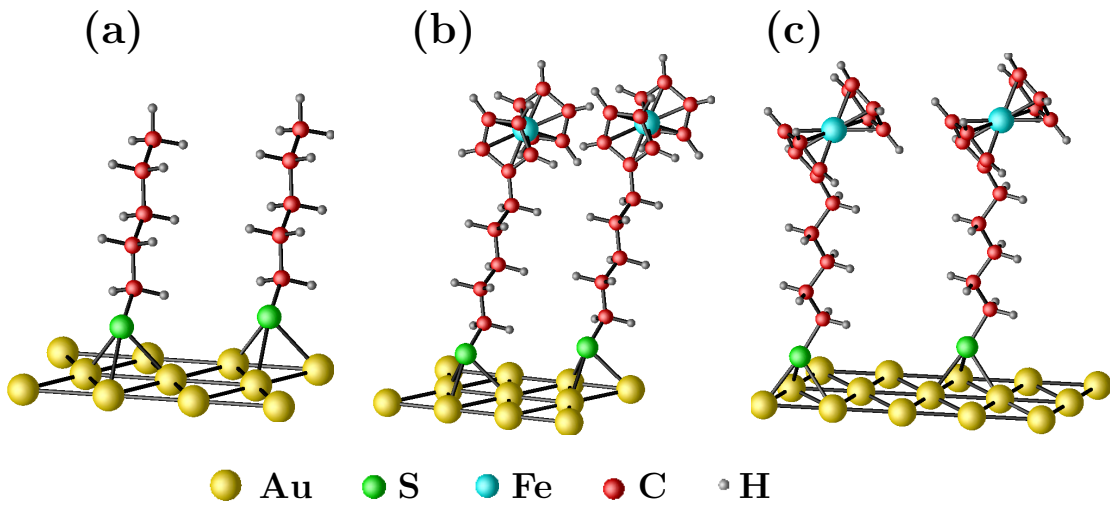

Figure 4.1: **Atomic structures of (a) C_5S-Au(111) ($\sqrt{3} \times \sqrt{3}$)R30°, (b) Fc-C_5S-Au(111) ($\sqrt{3} \times \sqrt{3}$)R30°, and (c) Fc-C_5S-Au(111) ($\sqrt{21} \times \sqrt{7}$).**

are investigated by total energy calculations. For each of the structures, the molecule is anchored with a sulfur atom to the Au(111) surface and all of the atoms are relaxed using total energy calculations and first-principles forces. For the adsorption of alkanethiolate on the Au(111) surface, two structural models have been reported. One is the ($\sqrt{3} \times \sqrt{3}$)R30° reconstruction with low molecular coverage [138]. The other is the c(4×2) reconstruction with high molecular coverage [139]. Since the molecular coverage for the c(4×2) structure is too high when ferrocenyl heads are present, we only consider the ($\sqrt{3} \times \sqrt{3}$)R30°, see Fig. 4.1(a). Our calculations show that the fcc hollow site is favorable and that the molecular axes are tilted by 27° with respect to the surface normal. These results are in agreement with recent theoretical work [140, 141] and experiments [138].

In the case of Fc-C_5S on Au(111), our calculations for the ($\sqrt{3} \times \sqrt{3}$)R30° Fc

superlattice (Fig. 4.1(b)) show that the repulsive interaction between the neighboring ferrocenyls is very strong, resulting in a positive adsorption energy. To determine the saturation coverage, we calculated adsorption energies for a number of different structures and coverages. Among the structures that were considered, two molecules per ($\sqrt{21} \times \sqrt{7}$) unit cell is the maximum coverage for which the adsorption energy is still negative (-0.3 eV per molecule in this case). For this coverage, the optimized tilt angle of the molecular axis is 15°. Clearly, the ferrocenyl head greatly decreases the molecular packing density on the Au(111) substrate.

4.3 Quantum Transport Calculation

The non-linear I-V characteristics are evaluated using the non-equilibrium Green's function (NEGF) methodology [142, 143] in a basis of optimal localized orbitals [66, 73]. The atom-centered orbitals are optimized variationally in the equilibrium geometry. Four orbitals for hydrogen, six for carbon, eight for sulfur and ten for iron and gold atoms are used, all with the radius of 4.50 Å. The Hartree potential is obtained by solving the Poisson equation with fixed boundary conditions to match the left and right leads, *i.e.*, the two semi-infinite gold leads in our case. A self-consistent calculation is carried out for each bias, which is critical in order to obtain correct I-V characteristics at large bias [144]. Due to the efficiency of our $O(N)$ methodology [66, 73], we are able to include 372 atoms and a total of nearly 4000 electrons in the calculations.

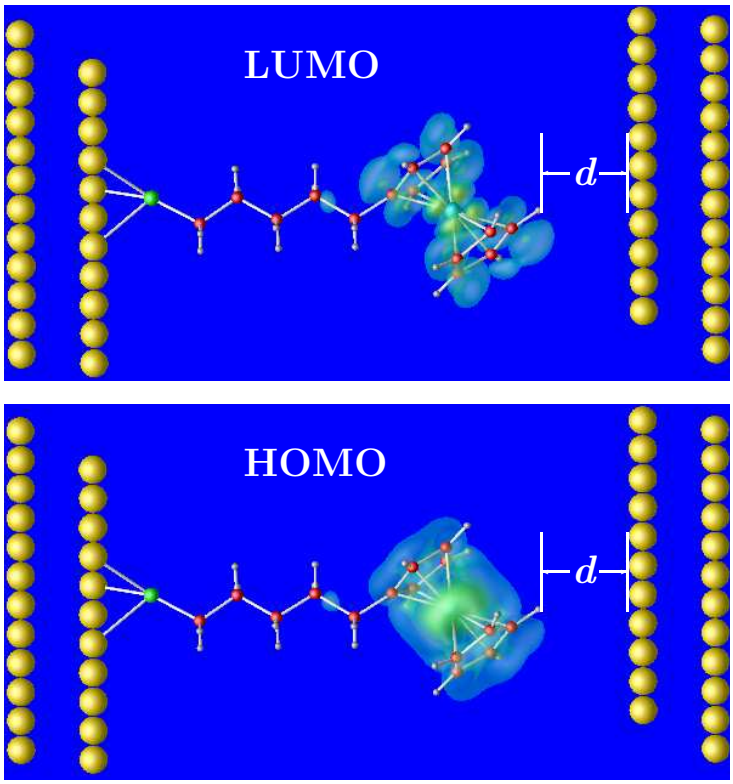

Figure 4.2: **Schematic views of the Au-SC$_5$S-Fc-Au system, where d is the distance between the molecule and the STM tip, and of isosurfaces of the HOMO and LUMO orbitals at zero bias. Both the HOMO and the LUMO are localized on the ferrocenyl head.**

Fig. 4.2 shows a schematic view of the system, together with the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO). As described above, the ferrocenyl-pentanethiolate is anchored with a sulfur atom to the Au(111) surface (left part in Fig. 4.2). In our calculations, we model the STM tip as another Au(111) surface (right part in Fig. 4.2). In experiments, the distance d between the molecule and the STM tip is determined by the setpoint current. The lower current, the larger the distance, since the magnitude of the tunneling current

depends exponentially on d.

The transmission of the molecular junction is mainly determined by molecular orbital(s) whose energy levels lie near the Fermi level of metal lead in contact with the molecule. These orbitals are usually either the HOMO or the LUMO. Their isosurfaces are plotted in Fig. 4.2. Clearly, both the HOMO and the LUMO are localized on the ferrocenyl head.

After the potential profile is self-consistently determined, the transmission spectrum under the external applied bias V can be calculated by

$$T(E,V) = \frac{2e^2}{h} Tr[\Gamma_L(E) G_C^+(E) \Gamma_R(E) G_C^-(E)],$$

where $G_C^{\pm}$ are the advanced and retarded Green's functions of the conductor, and $\Gamma_{L,R}$ are its coupling functions to the left and right leads, respectively.

The calculated transmission spectra $T(E,V)$ are shown in Fig. 4.3, with the Fermi level of the gold substrate at zero bias chosen as the energy zero. The two white lines show the chemical potentials of the left and right leads, with $\mu_L - \mu_R = V$ being the effective applied bias. We note that the transmission peak from the HOMO resonance is always near the lower chemical potential for both positive and negative biases. This is because the bias-induced potential drop across the molecular junction is very different for positive and negative biases. At negative bias, the largest potential drop occurs across the vacuum between the molecule and the right electrode (STM tip). This indicates that the electrons on the left side of the junction (the Au substrate)

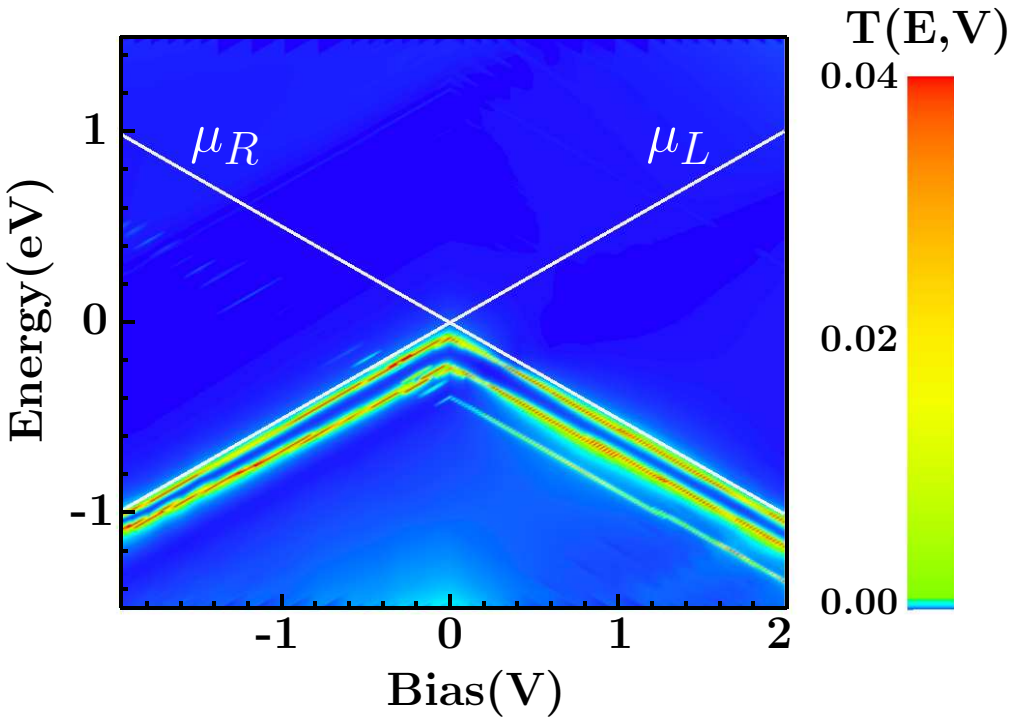

Figure 4.3: **Calculated transmission spectra $T(E,V)$ for the Au(111)-SC$_5$S-Fc-Au(111) junction with d = 3.39 Å. The two white lines are the chemical potentials of the two leads, $\mu_L = E_F + eV/2$ and ($\mu_R = E_F - eV/2$), as functions of the applied bias voltage.**

can move freely in the molecule to effectively screen the applied field. However, at positive bias, we find that the potential drops mainly in the carbon chain region of the molecule. As a result, the average potential around the ferrocenyl head is always aligned with the lower chemical potential and the HOMO, which is localized on the ferrocenyl, and is thus pinned by the lower chemical potential.

The current through the molecular junction is given by

$$I(V) = \int_{-\infty}^{\infty} T(E,V)[n_F(E-\mu_L) - n_F(E-\mu_R)]dE,$$

where n_F is the Fermi-Dirac distribution and $T(E,V)$ is the transmission coefficient

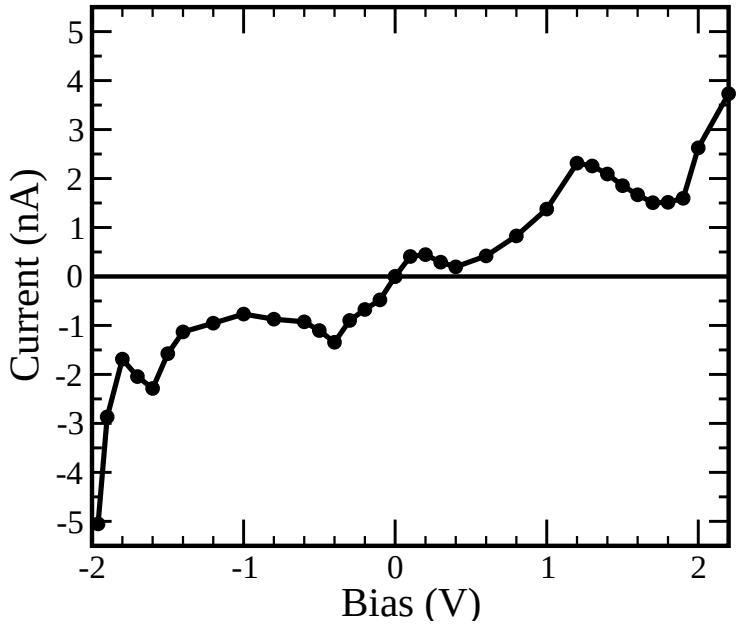

Figure 4.4: **Current-voltage characteristics of the Au(111)-SC$_5$S-Fc-Au(111) junction with d = 3.39 ÅCurrent-voltage characteristics of the Au(111)-SC$_5$S-Fc-Au(111) junction with d = 3.39 Å.**

at bias V.

The I-V curve calculated with the STM-molecule distance of $d = 3.39$ Å is shown in Fig. 4.4. The NDR features are at -1.6, -0.4, 0.2, and 1.2 V, compared with the experimentally observed NDR at -1.6 and 1.6 V. Tunneling through the ferrocenyl head is responsible for these features, since the calculated I-V curve without the ferrocenyl, *i.e.*, for pentanethiolate on Au(111), does not show NDR behavior. This is in good agreement with experiments [135, 136].

All of the NDR features can be understood in terms of resonant tunneling between the two top occupied levels and the gold density of states. In the density of states

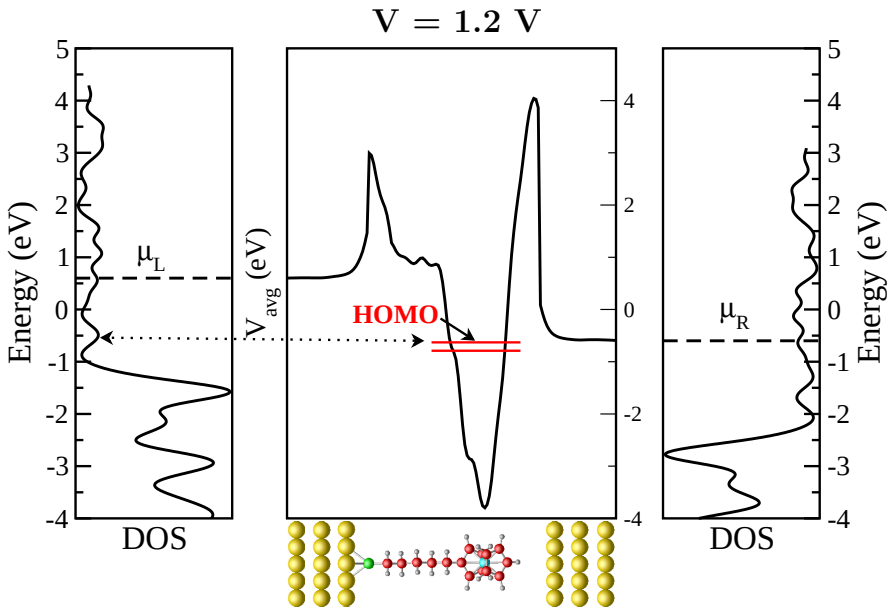

Figure 4.5: **The NDR mechanism for the Au(111)-SC$_5$S-Fc-Au(111) junction at 1.2 V. The left and right panels are densities of states (DOS) of bulk gold. The horizontal dashed lines in the DOS plot mark the left and right chemical potentials. The middle panel shows the average potential in the molecular junction. The horizontal solid lines in the middle plot mark the energies of the HOMO and the level just below the HOMO, because these orbitals dominate the transmission.**

(DOS) of bulk gold, the first peak below the Fermi energy is at around -0.2 eV and the second one is at -1.2 eV. When the HOMO level is aligned with one of these peaks, the current reaches its maximum and NDR appears. Fig. 4.5 illustrates the NDR mechanism at 1.2 V by plotting the DOS of the gold leads, the average potential, and the energy positions of the top two occupied orbitals. These orbitals are localized on the ferrocenyl by potential barriers. One barrier is around the Au-molecule interface, while the other is in the vacuum region between the ferrocenyl and the STM tip.

Therefore, the ferrocenyl acts like a quantum dot. As shown in Fig. 4.5, at bias of 1.2 V, the HOMO level is aligned with the peak in the DOS of left lead. As a result, a current maximum appears. For negative bias, the NDR can be understood in a similar way, except that the resonant coupling is between the HOMO and the right lead. Charge transfer between the molecule and the right lead also plays an important role in determining the position of the NDR. At a negative bias, charge transfer leads to an internal electric field that cancels out part of the applied external bias. As a result, the NDR appears at a higher voltage for negative bias than for the positive one, because there is no charge transfer for positive bias.

The calculated value of the current in Fig. 4.4 is about 50 times bigger than that of the STM measurement. This is mainly due to the unknown distance d between the STM tip and the molecule. In order to investigate the effect of this distance on electron transport, we carried out calculations for different values of d.

Fig. 4.6 shows the I-V curves for positive bias at distances of 2.01, 3.39 and 4.34 Å, respectively. Their overall shapes are very similar for the three different distances and the NDR positions are at about same biases. However, the absolute value of the current decreases exponentially with the increase of the distance. For example, at the bias of 0.6 V, we can fit the current by $I(d) = I_0 e^{-\beta d}$ with the decay constant $\beta = 0.90$. Another effect of the distance is in the shape of the NDR region, or its "strength." The NDR at 1.2 V is enhanced with an increase in the vacuum distance. This indicates that a low current setpoint in the STM experiment is important for the observation of NDR at a large bias. However, an increase of the distance has the

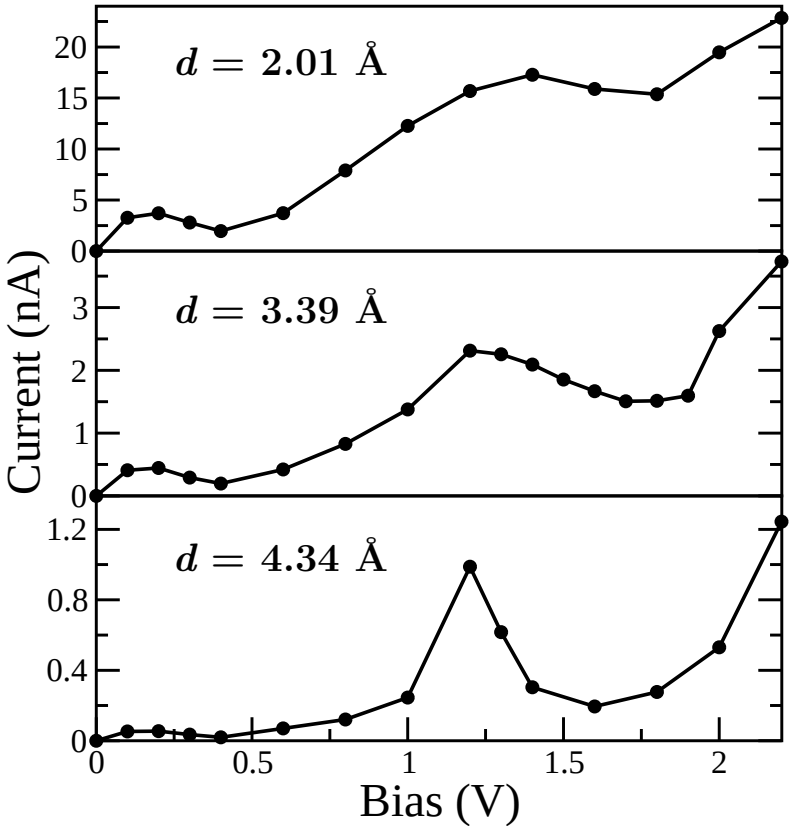

Figure 4.6: **Current-voltage characteristics of the Au(111)-SC$_5$S-Fc-Au(111) junction as a function of the STM-molecule distance.**

opposite effect on NDR at low bias; the peak at 0.2 V is attenuated with increasing d. This result can explain why the low bias NDR is not seen in the experiment, because a high current setpoint in the STM measurement would be necessary for its observation. This "NDR tuning" effect can be utilized in the design of molecular devices, for example by introducing spacer layers, which would adjust the strength of the NDR and the working current for optimal device performance.

Obviously, the length of the alkane chain in the molecule plays a major role in determining the magnitude of the current. For computational reasons, our cal-

culations were performed on ferrocenyl-pentanethiolate, while the experiment used ferrocenyl-undecanethiolate. However, we performed a few calculations for ferrocenyl-undecanethiolate, from which we learned that a current of the same magnitude as in the experiment would be obtained at the STM distance of 3.01 Å. Nevertheless, the shape and the major features of the distance-dependent I-V curves remain similar to those of ferrocenyl-pentanethiolate.

Apart from the distance and length effects, the unknown molecule/gold contact geometry and chemistry could also partially explain the remaining discrepancies between the theory and experiment. For example, if the molecule is attached to an atomic protrusion rather than an atomically flat surface, the magnitude of the current would be substantially lower [145, 146]. In addition, other variables, such as the temperature or local disorder in the metal near the contact can alter the value of current measured in experiment.

4.4 Summary and Conclusions

In summary, we investigated theoretically the geometry, electronic structure, and quantum transport properties of ferrocenyl-alkanethiolate self-assembled monolayers on Au(111) by large-scale first-principles calculations. The calculated NDR features at large biases are in good agreement with experiment. By carefully analyzing the transmission, charge transfer, and the alignment of molecular levels, we find that the HOMO is always localized at the ferrocenyl head and is pinned by the lowest

chemical potential of the electrodes. Resonant tunneling through the top two occupied orbitals dominates the I-V characteristics, with NDR effects being due to features in the density of states of the gold surface. The magnitude of the current and of the strength of the NDR are dramatically affected by the distance between the molecule and the STM tip. In practical devices, spacer layers will have a similar effect and can be used to tune the performance of the device.

Bibliography

[1] T. Kinoshita and W. B. Lindquist. Eighth-order anomalous magnetic moment of the electron. *Phys. Rev. Lett.*, 47(22):1573, November 1981.

[2] R. S. Van Dyck Jr., P. B. Schwinberg, and H. G. Dehmelt. New high-precision comparison of electron and positron g factors. *Phys. Rev. Lett.*, 59(1):26, July 1987.

[3] J. Ihm, A. Zunger, and M. L Cohen. Momentum-space formalism for the total energy of solids. *J. Phys. C*, 12:4409, 1979.

[4] P. J. H. Deteneer and W van Haeringen. The pseudopotential-density-functional method in momentum space: details and test cases. *J. Phys. C*, 18:4127, 1985.

[5] M. Born and R. Oppenheimer. Zur quantentheorie der molekeln. *Ann. Phys. (Leipzig)*, 84(20):457, 1927.

[6] V. Fock. *Z. Phys.*, 61:126, 1930.

[7] J. C. Slater. A simplification of the hartree-fock method. *Phys. Rev.*, 81:385, 1951.

[8] D. R. Hartree. *Proc. R. Soc. London*, A113:621, 1928.

[9] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864, 1964.

[10] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133, 1965.

[11] L. I. Schiff. *Quantum Mechanics*. McGraw-Hill, 1986.

[12] R. O. Jones and O. Gunnarsson. The density functional formalism, its applications and prospects. *Rev. Mod. Phys.*, 61:689, 1989.

[13] J. P. Perdew and Y. Wang. *Phys. Rev. B*, 46:12947, 1992.

[14] J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23(10):5048, 1981.

[15] D. M. Ceperley and B. J. Alder. Ground state of the electron gas by a stochastic method. *Phys. Rev. Lett.*, 45(7):566, 1980.

[16] J. Harris and R. O. Jones. The surface energy of a bounded electron gas. *J. Phys. F*, 4:1170, 1974.

[17] O. Gunnarsson and B. I. Lundqvist. *Phys. Rev. B*, 13:4174, 1976.

[18] D. C. Langreth and J. P. Perdew. *Phys. Rev. B*, 15:2884, 1977.

[19] D. C. Langreth and J. P. Perdew. Theory of nonuniform electronic systems. i. analysis of the gradient approximation and a generalization that works. *Phys. Rev. B*, 21(12):5469, 1980.

[20] J. P. Perdew and Wang Y. Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation. *Phys. Rev. B*, 33(12):8800, 1986.

[21] J. P. Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron electron gas. *Phys. Rev. B*, 33(12):8822, 1986.

[22] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behaviour. *Phys. Rev. A*, 38:3098, 1988.

[23] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865, 1996.

[24] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 78:1396, 1997.

[25] J. P. Perdew. Unified theory of exchange and correlation beyond the local density approximation. *Electronic Structure of Solids '91, edited by P. Ziesche and H. Eschrig*, pages 11–20, 1991.

[26] K. Burke, J. P. Perdew, and Y. Wang. Derivation of a generalized gradient approximation: The pw91 density functional. *Electronic Density Functional*

Theory: Recent Progress and New Directions, edited by J. F. Dobson, G. Vignale and M. P. Das, pages 81–121, 1997.

[27] J. C. Phillips. Energy-band interpolation scheme based on a pseudopotential. *Phys. Rev.*, 112:685, 1958.

[28] M. L. Cohen and V. Heine. *Solid State Physics*, Vol.24:p.37, 1970.

[29] J. C. Phillips and L. Kleinman. New method for calculating wave functions in crystals and molecules. *Phys. Rev.*, 116:287, 1959.

[30] D. R. Hamann, M. Schlter, and C. Chiang. Norm-conserving pseudopotential. *Phys. Rev. Lett.*, 43:1494, 1979.

[31] L. Kleinman and D. M. Bylander. Efficacious form for model pseudopotentials. *Phys. Rev. Lett.*, 48:1425, 1982.

[32] S. G. Louie, S. Froyen, and M. L. Cohen. Nonlinear ionic pseudopotentials in spin-density-functional calculations. *Phys. Rev. B*, 26:1738, 1982.

[33] G. B. Bachelet, D. R. Hamann, and M. Schlter. Pseudopotentials that work: From H to Pu. *Phys. Rev. B*, 26:4199, 1982.

[34] A. M. Rappe, Karin M. Rabe, E. Kaxiras, and J. D. Joannopoulos. Optimized pseudopotentials. *Phys. Rev. B*, 41:1227, 1990.

[35] D. Vanderbilt. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B*, 41:7892, 1990.

[36] X. Gonze, P. Kckell, and M. Scheffler. Ghost states for separable, norm-conserving, ab initiop pseudopotentials. *Phys. Rev. B*, 41:12264, 1990.

[37] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Saunders College, Philadelphia, 1976.

[38] L. P. Bouckaert, R. Smoluchowski, and E. Wigner. Theory of Brillouin zones and symmetry properties of wave functions in crystals. *Phys. Rev.*, 50:58, 1936.

[39] A. Baldereschi. Mean-value point in the Brillouin zone. *Phys. Rev. B*, 7(12):5212, 1973.

[40] D. J. Chadi and M. L. Cohen. Special points in the Brillouin zone. *Phys. Rev. B*, 8(12):5747, 1973.

[41] H. J. Monkhorst and J. D. Pack. Special points for Brillouin-zone integrations. *Phys. Rev. B*, 13(12):5188, 1976.

[42] D. J. Chadi. Special points for Brillouin-zone integrations. *Phys. Rev. B*, 16(4):1746, 1977.

[43] R. A. Evarestov and V. P. Smirnov. Special points of the Brillouin zone and their use in the solid state theory. *Phys. Status Solidi*, 119:9, 1983.

[44] S. Froyen. Brillouin-zone integration by fourier quadrature: Special points for superlattice and supercell calculations. *Phys. Rev. B*, 39(5):3168, 1989.

[45] I. J. Robertson and M. C. Payne. k-point sampling and the $k \cdot p$ method in pseudopotential total energy calculations. *J. Phys.: Condens. Matter*, 2:9837, 1990.

[46] I. J. Robertson and M. C. Payne. The $k \cdot p$ method in pseudopotential total energy calculations: error reduction and absolute energies. *J. Phys.: Condens. Matter*, 3:8841, 1991.

[47] R. Car and M. Parrinello. Unified approach for molecular dynamics and density-functional theory. *Phys. Rev. Lett.*, 55:2471, 1985.

[48] M. C. Payne, D. C. Allan, M. P. Teter, T. A. Arias, and J. D. Joannopoulos. Iterative minimization techniques for *ab initio* total-energy calculation: molecular dynamics and conjugate gradients. *Rev. Mod. Phys.*, 64:1045, 1992.

[49] M. P. Teter, M. C. Payne, and D. C. Allan. Solution of schringer's equation for large systems. *Phys. Rev. B*, 40:12255, 1989.

[50] T. A. Arias, M. C. Payne, and J. D. Joannopoulos. *Ab initio* molecular dynamics: Analytically continued energy functionals and insights into iterative solutions. *Phys. Rev. Lett.*, 69:1077, 1992.

[51] T. A. Arias, M. C. Payne, and J. D. Joannopoulos. *Ab initio* molecular-dynamics techniques extended to large-length-scale systems. *Phys. Rev. B*, 45:1538, 1992.

[52] P. E. Blöchl. Generalized separable potentials for electronic-structure calculations. *Phys. Rev. B*, 41:5414, 1990.

[53] F. Gygi. Electronic-structure calculations in adaptive coordinates. *Phys. Rev. B*, 48:11692, 1993.

[54] D. R. Hamann. Application of adaptive curvilinear coordinates to the electronic structure of solids. *Phys. Rev. B*, 51:7337, 1995.

[55] D. R. Hamann. Band structure in adaptive curvilinear coordinates. *Phys. Rev. B*, 51:9508, 1995.

[56] S. R. White, J. W. Wilkins, and M. P. Teter. Finite-element method for electronic structure. *Phys. Rev. B*, 39:5819, 1989.

[57] J. Bernholc, J. Y. Yi, and D. J. Sullivan. *Chem. Soc.*, 92:217, 1991.

[58] K. A. Iyer, M. P. Merrick, and T. L. Beck. *J. Chem. Phys.*, 103:227, 1995.

[59] F. Gygi and G. Galli. Real-space adaptive-coordinate electronic-structure calculations. *Phys. Rev. B*, 52:R2229, 1995.

[60] G. Zumbach, N. A. Modine, and E. Kaxiras. *Solid State Commun.*, 99:57, 1996.

[61] S. Wei and M. Y. Chou. Wavelets in self-consistent electronic structure calculations. *Phys. Rev. Lett.*, 76:2650, 1996.

[62] T. A. Arias. Multiresolution analysis of electronic structure: semi-cardinal and wavelet bases. *Rev. Mod. Phys.*, 71:267, 1999.

[63] E. L. Briggs, D. J. Sullivan, and J. Bernholc. Large-scale electronic-structure calculations with multigrid acceleration. *Phys. Rev. B*, 52:5471, 1995.

[64] E. L. Briggs, D. J. Sullivan, and J. Bernholc. *Phys. Rev. B*, 54:14362, 1996.

[65] S. Goedecker. Linear scaling electronic structure methods. *Rev. Mod. Phys.*, 71:1085, 1999.

[66] J. L. Fattebert and J. Bernholc. *Phys. Rev. B*, 62:1713, 2000.

[67] O. F. Sankey abd D. J. Niklewski. *Ab-initio* multi-center tight-binding model for molecular-dynamics simulations and other applications in covalent systems. *Phys. Rev. B*, 40:3979, 1989.

[68] P. Ordejón, E. Artacho, and J. M. Soler. Self-consistent order-n density functional calculations for very large systems. *Phys. Rev. B*, 53:10441, 1996.

[69] J. R. Chelikowsky, N. Troullier, K. Wu, and Y. Saad. Higher-order finite-difference pseudopotential method: An application to diatomic molecules. *Phys. Rev. B*, 50:11355, 1994.

[70] L. Collatz. *The Numerical Treatment of Differential Equations*, page 164. Springer-Verlag, Berlin, 1960.

[71] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B.P. Flannery. *Numerical Recipes in Fortran*, chapter 19.5. Cambridge University Press, Cambridge, 2nd edition, 1992.

[72] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B.P. Flannery. *Numerical Recipes in Fortran*, chapter 19.6. Cambridge University Press, Cambridge, 2nd edition, 1992.

[73] M. B. Nardelli, J. L. Fattebert, and J. Bernhole. *Phys. Rev. B*, 64:245423, 2001.

[74] M. B. Nardelli. Electronic transport in extended systems: Appllication to carbon nanotubes. *Phys. Rev. B*, 60:7828, 1999.

[75] R. J. Harrison. Krylov subspace accelerated inexact Newton method for linear and nonlinear equations. *J. Comput. Chem.*, 25:328, 2002.

[76] Y. Bar-Yam, S. T. Pantelides, and J. D. Joannopoulos. *Ab initio* pseudopotential solid-state calculations of highly electronegative first-row elements. *Phys. Rev. B*, 39:3396, 1989.

[77] C. Lee, D. Vanderbilt, K. Laasonen, R. Car, and M. Parrinello. *Ab initio* studies on high pressure phases of ice. *Phys. Rev. Lett.*, 69:462, 1992.

[78] K. Laasonen, R. Car, C. Lee, and D. Vanderbilt. Implementation of ultrasoft pseudopotentials in *ab initio* molecular dynamics. *Phys. Rev. B*, 43:6796, 1991.

[79] K. Laasonen, A. Pasquarello, R. Car, C. Lee, and D. Vanderbilt. Car-Parrinello molecular dynamics with Vanderbilt ultrasoft pseudopotentials. *Phys. Rev. B*, 47:10142, 1993.

[80] A. Pasquarello, K. Laasonen, R. Car, C. Lee, and D. Vanderbilt. *Ab initio* molecular dynamics for *d*-electron systems: Liquid copper at 1500 k. *Phys. Rev. Lett.*, 69:1982, 1992.

[81] T. Ono and K. Hirose. *Phys. Rev. Lett.*, 82:5016, 1999.

[82] K. P. Huber and G. Herzberg. *Constants of Diatomic Molecules*. Van Nostrand, New York, 1979.

[83] A. D. Corso, A. Pasquarello, and A. Baldereschi. Density-functional perturbation theory for lattice dynamics with ultrasoft pseudopotentials. *Phys. Rev. B*, 56:11369, 1997.

[84] F. D. Murnaghan. *Proc. Natl. Acad. Sci, U.S.A*, 3:244, 1944.

[85] J. W. Lynn, H. G. Smith, and R. M. Nicklow. Lattice dynamics of gold. *Phys. Rev. B*, 8:3493, 1973.

[86] P. Ballone and G. Galli. Accurate pseudopotential local-density-approximation computations for neutral and ionized dimers of the IB and IIB groups. *Phys. Rev. B*, 42:1112, 1990.

[87] R. Courths and S. Hüfner. Photoemission experiments on copper. *Phys. Rep.*, 112:53, 1984.

[88] A. Marini, G. Onida, and R. D. Sole. Plane-wave DFT-LDA calculation of the electronic structure and absorption spectrum of copper. *Phys. Rev. B*, 64:195125, 2001.

[89] G. Nilsson and S. Rolandson. Lattice dynamics of copper at 80 k. *Phys. Rev. B*, 7:2393, 1973.

[90] J. J. Lander and J. Morrison. Surface reactions of silicon (111) with aluminum and indium. *J. Appl. Phys.*, 36:1706, 1965.

[91] T. Abukawa, M. Sasaki, F. Hisamatsu, T. Goto, T. Kinoshita, A. Kakizaki, and S. Kono. Surface electronic structure of a single-domain Si(111)4×1-In surface: a synchrotron radiation photoemission study. *Surf. Sci.*, 325:33, 1995.

[92] T. Uchihashi and U. Ramsperger. Electron conduction through quasi-one-dimensional indium wires on silicon. *Appl. Phys. Lett.*, 80:4169, 2002.

[93] F. J. Himpsel, K. N. Altmann, R. Bennewitz, J. N. Crain, A. Kirakosian, J. L. Lin, and J. L. McChesney. One-dimensional electronic states at surfaces. *J. Phys.: Condens. Matter*, 13:11097, 2001.

[94] H. W. Yeom, S. Takeda, E. Rotenberg, I. Matsuda, K. Horikoshi, J. Schaefer, C. M. Lee, S. D. Kevan, T. Ohta, T. Nagao, and S. Hasegawa. Instability and charge density wave of metallic quantum chains on a silicon surface. *Phys. Rev. Lett.*, 82:4898, 1999.

[95] S. V. Ryjkov, T. Nagao, V. G. Lifshits, and S. Hasegawa. Phase transition and stability of Si(111)-8×'2' In surface phase at low temperatures. *Surf. Sci.*, 488:15, 2001.

[96] O. Bunk, G. Falkenberg, J. H. Zeysing, L. Lottermoser, R. L. Johnson, M. Nielsen, F. Berg-Rasmussen, J. Baker, and R. Feidenhans'l. Structure determination of the indium-induced Si(111)-(4×1) reconstruction by surface x-ray diffraction. *Phys. Rev. B*, 59:12228, 1999.

[97] R. H. Miwa and G. P. Srivastava. Atomic geometry, electronic structure and image state for the Si(111)-In(4×1) nanowire. *Surf. Sci.*, 473:123, 2001.

[98] J. H. Cho, D. H. Oh, K. S. Kim, and L. Kleinman. Weakly correlated one-dimensional indium chains on Si(111). *Phys. Rev. B*, 64:235302, 2001.

[99] J. Nakamura, S. Watanabe, and M. Aono. Anisotropic electronic structure of the Si(111)-(4×1) In surface. *Phys. Rev. B*, 6319:art. no.193307, 2001.

[100] I. G. Hill and A. B. McLean. Strongly anisotropic band dispersion of an image state located above metallic nanowires. *Phys. Rev. Lett.*, 82:2155, 1999.

[101] C. Kumpf, O. Bunk, J. H. Zeysing, Y. Su, M. Nielsen, R. L. Johnson, R. Feidenhans'l, and K. Bechgaard. Low-temperature structure of indium quantum chains on silicon. *Phys. Rev. Lett.*, 85:4916, 2000.

[102] H. W. Yeom, K. Horikoshi, H. M. Zhang, K. Ono, and R. I. G. Uhrberg. Nature of the broken-symmetry phase of the one-dimensional metallic in Si(111) surface. *Phys. Rev. B*, 65:241307, 2002.

[103] K. Sakamoto, H. Ashima, H. W. Yeom, and W. Uchida. Angle-resolved high-resolution electron-energy-loss study of In-adsorbed Si(111)-(4×1) and -(8×2) surfaces. *Phys. Rev. B*, 62:9923, 2000.

[104] O. Gallus, T. Pillo, M. Hengsberger, P. Segovia, and Y. Baer. A system with a complex phase transition: indium chains on Si(111). *Eur. Phys. J. B*, 20:313, 2001.

[105] F. Aryasetiawan and O. Gunnarsson. The GW method. *Rep. Prog. Phys.*, 61:237, 1998.

[106] F. Pedreschi, J. D. O'Mahony, P. Weightman, and J. R. Power. Evidence of electron confinement in the single-domain, (4×1) in superstructure on vicinal Si(111). *Appl. Phys. Lett.*, 73:2152, 1998.

[107] J. F. McGilp. Optical response of low-dimensional in nanostructures grown by self-assembly on Si surfaces. *Phys. Status Solidi A*, 188:1361, 2001.

[108] K. Fleischer, S. Chandola, N. Esser, W. Richter, and J. F. McGilp. Reflectance anisotropy spectroscopy of Si(111)-(4×1) In. *Phys. Status Solidi A*, 188:1411, 2001.

[109] K. Fleischer, S. Chandola, N. Esser, W. Richter, and J. F. McGilp. Phonon and polarized reflectance spectra from Si(111)-(4x1) In: Evidence for a charge-density-wave driven phase transition. *Phys. Rev. B*, 67:235318, 2003.

[110] D. R. Hamann. Generalized norm-conserving pseudopotentials. *Phys. Rev. B*, 40:2980, 1989.

[111] M. Fuchs and M. Scheffler. Ab initio pseudopotentials for electronic structure calculations of poly-atomic systems using density-functional theory. *Comput. Phys. Commun.*, 119:67, 1999.

[112] R. Del Sole. Microscopic theory of optical properties of crystal surfaces. *Solid State Commun.*, 37:537, 1981.

[113] F. Manghi, R. Del Sole, A. Selloni, and E. Molinari. Anisotropy of surface optical properties from first-principles calculations. *Phys. Rev. B*, 41:9935, 1990.

[114] A. I. Shkrebtii, N. Esser, W. Richter, W. G. Schmidt, F. Bechstedt, B. O. Fimland, A. Kley, and R. Del Sole. Reflectance anisotropy of GaAs(100): Theory and experiment. *Phys. Rev. Lett.*, 81:721, 1998.

[115] H. Ehrenreich and M. H. Cohen. Self-consistent field approach to the many-electron problem. *Phys. Rev.*, 1:786, 1959.

[116] P. H. Hahn, W. G. Schmidt, and F. Bechstedt. Bulk excitonic effects in surface optical spectra. *Phys. Rev. Lett.*, 88:016402, 2002.

[117] W. G. Schmidt, S. Glutsch, P. H. Hahn, and F. Bechstedt. Efficient $O(N^2)$ method to solve the Bethe Salpeter equation. *Phys. Rev. B*, 67:085307, 2003.

[118] W. G. Schmidt, F. Bechstedt, and J. Bernholc. Understanding reflectance anisotropy: Surface-state signatures and bulk-related features. 18:2215, 2000.

[119] W. C. Lu, W. G. Schmidt, E. L. Briggs, and J. Bernholc. Optical anisotropy of the SiC(001)-(3×2) surface: Evidence for the two-adlayer asymmetric-dimer model. *Phys. Rev. Lett.*, 85:4381, 2000.

[120] A. A. Saranin, A. V. Zotov, K. V. Ignatovich, V. G. Lifshits, T. Numata, O. Kubo, H. Tani, M. Katayama, and K. Oura. Structural model for the Si(111)-4×1 In reconstruction. *Phys. Rev. B*, 56:1017, 1997.

[121] I. G. Hill and A. B. McLean. Metallicity of in chains on Si(111). *Phys. Rev. B*, 56:15725, 1997.

[122] W. G. Schmidt and J. Bernholc. Step-induced optical anisotropy of Si(111): H surfaces. *Phys. Rev. B*, 61:7604, 2000.

[123] A. Aviram and M. A. Ratner. Molecular rectifiers. *Chem. Phys. Lett.*, 29:277, 1974.

[124] R. M. Metzger. Electrical rectification by a molecule: The advent of unimolecular electronic devices. *Acc. Chem. Res.*, 32:950, 1999.

[125] J. Chen, M. A. Reed, A. M. Rawlett, and J. M. Tour. Large on-off ratios and negative differential resistance in a molecular electronic device. *Science*, 286:1550, 1999.

[126] S. Datta, W. Tian, S. Hong, R. Reifenberger, J. I . Henderson, and C. P. Kubiak. Current-voltage characteristics of self-assembled monolayers by scanning tunneling microscopy. *Phys. Rev. Lett.*, 79:2530, 1997.

[127] Y. Xue, S. Datta, S. Hong, R. Reifenberger, J. I. Henderson, and C. P . Kubiak. Negative differential resistance in the scanning-tunneling spectroscopy of organic molecules. *Phys. Rev. B*, 59:7852, 1999.

[128] S. Hong, R. Reifenberger, W. Tian, S. Datta, J. Henderson, and C. P. Kubiak. Molecular conductance spectroscopy of conjugated, phenyl-based molecules on Au(111): the effect of end groups on molecular conduction. *Superlattices Microstruct.*, 28:289, 2000.

[129] M. A. Reed, C. Zhou, C. J. Muller, T. P. Burgin, and J. M. Tour. Conductance of a molecular junction. *Science*, 278:252, 1997.

[130] J. G. Kushmerick, D. B. Holt, J. C. Yang annd J. Naciri, M. H. Moore, and R. Shashidhar. Metal-molecule contacts and charge transport across monomolecular layers: Measurement and theory. *Phys. Rev. Lett.*, 89:086802, 2002.

[131] J. Chen, W. Wang, M. A. Reed, A. M. Rawlett, D. W. Price, and J. M. Tour. Room-temperature negative differential resistance in nanoscale molecular junctions. *Appl. Phys. Lett.*, 77:1224, 2000.

[132] J. Chen and M. A. Reed. Electronic transport of molecular systems. *Chem. Phys.*, 281:127, 2002.

[133] M. A. Reed, J. Chen, A. M. Rawlett, D. W. Price, and J. M. Tour. Molecular random access memory cell. *Appl. Phys. Lett.*, 78:3735, 2001.

[134] S. A. Getty, C. Engtrakul, L. Wang, R. Liu, S. H. Ke, H. U. Baranger, W. Yang, M. S. Fuhrer, and L. R. Sita. Near-perfect conduction through a ferrocene-based molecular wire. *Phys. Rev. B*, 71:241401, 2005.

[135] C. B. Gorman, R. L. Carroll, and R. R. Fuierer. Negative differential resistance in patterned electroactive self-assembled monolayers. *Langmuir*, 17:6923, 2001.

[136] R. A. Wassel, G. M. Credo, R. R. Fuierer, D. L. Feldheim, and C. B. Gorman. Attenuating negative differential resistance in an electroactive self-assembled monolayer-based junctio. *J. Am. Chem. Soc*, 126:295, 2004.

[137] J. L. Pitters and R. A. Wolkow. Detailed studies of molecular conductance using atomic resolution scanning tunneling microscopy. *Nano Lett.*, 6:390, 2006.

[138] F. T. Arce, M. E. Vela, R. C. Salvarezza, and A. J. Arvia. Complex structural dynamics at adsorbed alkanethiol layers at Au(111) single-crystal domains. *Langmuir*, 14:7203, 1998.

[139] J. J. Gerdy and W. A. Goodard. Atomistic structure for self-assembled monolayers of alkanethiols on Au(111) surfaces. *J. Am. Chem. Soc*, 118:3233, 1996.

[140] Y. Yourdshahyan, H. K. Zhang, and A. M. Rappe. n-alkyl thiol head-group interactions with the Au(111) surface. *Phys. Rev. B*, 63:081405, 2001.

[141] H. Gronbeck, A. Curioni, and W. Andreoni. Thiols and disulfides on the Au(111) surface: The headgroup-gold interaction. *J. Am. Chem. Soc.*, 122:3839, 2000.

[142] B. Larade, J Taylor, H. Mehrez, and H. Guo. Conductance, i-v curves, and negative differential resistance of carbon atomic wires. *Phys. Rev. B*, 64:075420, 2001.

[143] M. Brandbyge, J. L. Mozos, P. Ordejon, J. Taylor, and K. Stokbro. Density-functional method for nonequilibrium electron transport. *Phys. Rev. B*, 65:165401, 2002.

[144] W. Lu, V. Meunier, and J. Bernholc. Nonequilibrium quantum transport properties of organic molecules on silicon. *Phys. Rev. Lett.*, 95:206805, 2005.

[145] M. D. Ventra, S. T. Pantelides, and N. D. Lang. First-principles calculation of transport properties of a molecular device. *Phys. Rev. Lett.*, 84:979, 2000.

[146] Y. Q. Xue and M. A. Ratner. Microscopic study of electrical transport through individual molecules with metallic contacts: II. effect of the interface structure. *Phys. Rev. B*, 68:115407, 2003.