SCANNING TUNNELING MICROSCOPY AND SPECTROSCOPY SIMULATIONS OF THE SILICON (111)-(7X7) SURFACE

WEIMING LIU

Bachelor of Engineering, department of electronics, Jilin University (2002)

A Thesis

Submitted to the School of Graduate Studies

of the University of Lethbridge

in Partial Fulfilment of the Requirements

for the Degree of

MASTER OF SCIENCE

Department of Physics

University of Lethbridge

LETHBRIDGE, ALBERTA, CANADA

Approval and Signature page (has no page number, should be ii in prefatory page)

(to be replaced)

Abstract

Since 1982, the Si (111)-(7x7) surface has been extensively studied both theoretically and experimentally with the modern powerful tools of STM and Scanning Tunneling Spectroscopy (STS). In this work, a simple atomic orbital model for the Si (111)-(7x7) surface is developed to simulate the experimental results of STM and STS. Based on Tersoff-Hamann's theory for the tunneling current, simulations of clean Si (111)-(7x7) constant-current images are presented. The direct, real-space simulated topographic images of the surface are compared to experimental results qualitatively and quantitatively. The simulation of spectroscopic imaging and normalized conductance spectra are also included. The adsorption of atomic hydrogen atoms onto the Si (111)-(7x7) surface is also simulated.

Acknowledgements

This thesis is the result of the inspiring and thoughtful guidance and supervision of my supervisor Dr. Steve Patitsas. But for him this work could not have been completed.

I thank Dr. David Siminovitch and Dr. Michael Gerken for the valuable suggestions and encouragements during the courses of this work

I thank my colleagues Steven Horn and Pouya Maraghechi for all the support.

I thank Frank Klassen, Heinz Fischer for their technical support

I thank my father Mr. Daqing Liu and my mother Mrs Guoqin Sun for their patience and sacrifice in making me what I am today.

I thank my friends Mrs. Patricia Postman, Mr. Keanan Postman, Mrs Glenda Postman, Mr. Akaid Dirac, Mr. Rick Melvin, Mrs. Ruth Melvin for encouragement and all the support.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AH	Absolute Height relative to the CH atom
CT	Charge Transfer
CITS	Current Imaging Tunneling Spectroscopy
CH	Corner Hole
DAS	Dimer-Adatom-Stacking fault
DB	Dangling Bonds
DC	Direct Current
DOS	Density of states
FC	Corner adatom in the Faulted halves
FCC	Face Centered Cubic lattice
FM	Middle adatom in the Faulted halves
FR	Rest atom in the Faulted halves
LDOS	Local Density of States
SH	Subtracted Height
STM	Scanning Tunneling Microscopy
STS	Scanning Tunneling Spectroscopy
TH	Tip Height relative to the actual vertical position of UC
THSP	Tip Height Set Point (the specific TH corresponding to a UC when the tip is directly above the UC and the tunnel current is at the set-point value)
UC	Corner adatom in the Unfaulted halves
UM	Middle adatom in the Unfaulted half of the unit cell
UR	Rest atom in the Unfaulted halves
WKB	Wentzel-Kramers-Brillouin

Chapter 1
Introduction

1.1 Overview

Surface science has been revolutionized with the advent of two powerful tools: scanning tunneling microscopy (STM) [1] and scanning tunneling spectroscopy (STS) [2-4]. According to the Tersoff-Hamann theory of tunneling [5], an important quantity with regard to STM and STS is the local density of states (LDOS) [3]. The calculation of LDOS is essential in calculating the image for a proposed structure or set of pictures in order to interpret STM images quantitatively. We propose calculating the Si (111)-(7x7) surface [6] image by a simple atomic orbital model. In the model, the LDOS can be calculated using the combination of the Tersoff-Hamann theory of tunneling and the tight binding method.

1.2 Scanning tunneling microscope

The scanning tunneling microscope operates by rastering a very sharp metal wire tip across a surface of a material. By using this microscope, one can image the surface at an Ångstrom-level scale. This is done by bringing the tip very near to the surface while electric potential is applied between the tip and the surface. This microscope is based on several principles. One is the quantum mechanical effect of tunneling, which allows us to "see" the surface. Another principle is the piezoelectric effect. This effect allows us to precisely control the tip on the scale of Ångstroms. In practice, a feedback loop is required to both monitor the tunneling current and coordinate the current and the position of the tip.

1.2.1 Tunneling effect

Tunneling is a quantum mechanical effect. A tunneling current occurs when electrons move through a region in which the potential energy is higher than the total energy of the electrons. Classically, an electron can never be inside such a region because negative kinetic energy is not allowed according to Newtonian mechanics. However, on the scale of the order of 10 Å, Newtonian concepts fail and quantum concepts dominate the scene and events on an electron or a single atom. The quantum rules make it possible for an electron to move from one classically allowed region to the other through a classically forbidden region. This phenomenon is called electron tunneling. According to an elementary quantum mechanical calculation [7], the transmission probability through the barrier decreases exponentially with the thickness of the barrier and the square root of the potential relative to the electron energy.

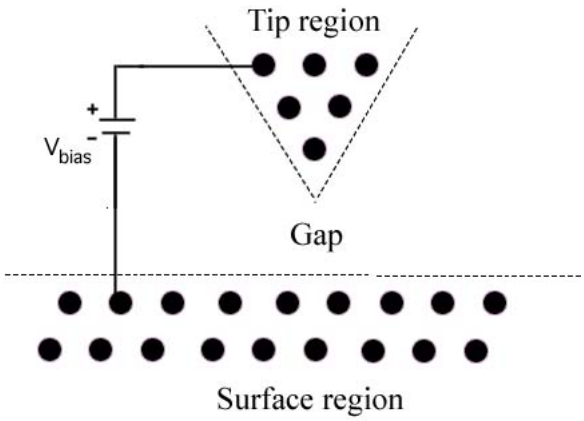


Figure 1-1: Geometrical configuration STM of the STM tunnel junction

To extend this description to STM: the sample surface and the atomically sharp tip, as shown in Figure 1-1, are two classically allowed regions. The gap (air, vacuum, liquid) between them is the potential barrier, which is a classically forbidden but quantum mechanically allowed region. The conducting tip is brought very carefully near the conducting surface. Once the

separation is within a few Å, electrons can tunnel through the gap from the tip to the sample or vice versa (depending on the sign of sample bias). The tunneling current is proportional to the transmission probability, and therefore depends exponentially on the tip-sample separation (the effective barrier thickness). This exponential dependence of tunneling current on the separation of the tip and the surface leads to exquisite spatial resolution in STM measurements, approximately 0.2nm laterally and 0.001nm vertically [8].

1.2.2 Inverse piezoelectric effect

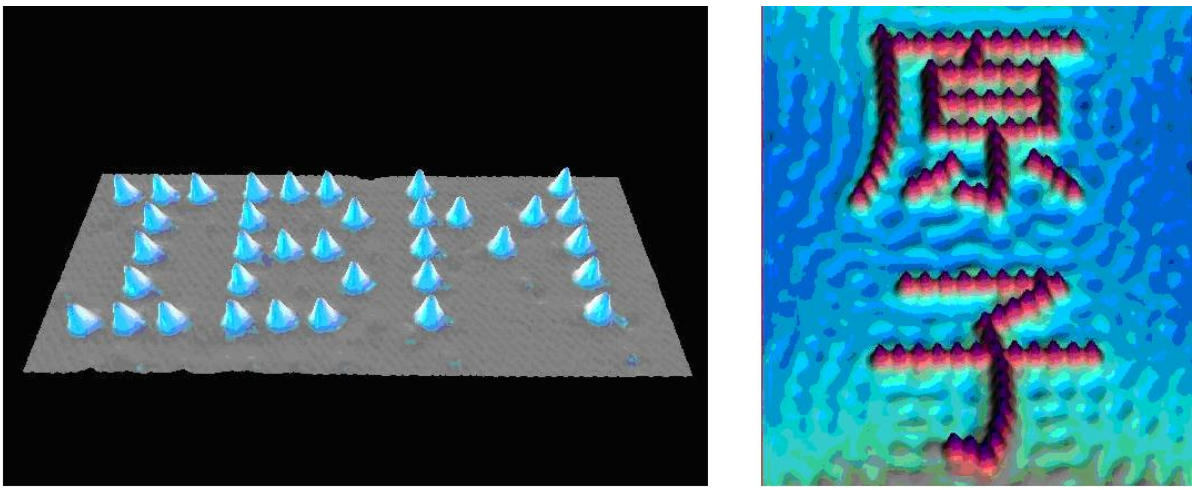
The piezoelectric effect was discovered in the 19th century [9]. It was observed that opposite charges are created on the sides of a certain crystal such as quartz when the crystal is compressed on those sides. The effect can be reversed as well. By placing a potential across a piezoelectric crystal, it will elongate or compress. It is this inverse piezoelectric effect that allows us to control the tip precisely in the three dimensions by having the tip mounted onto a piezo tube.

1.2.3 Electronics and feedback

Electronic circuits, including a feedback loop are used to measure the tunneling current, scan the tip, and translate this information into a form that we can use. A feedback loop constantly monitors the tunneling current and makes adjustments to the tip to maintain a constant tunneling current. These adjustments are recorded by the computer in a form of feedback voltage and presented as an image in the STM software. Such a setup is called a "constant current" image. In addition, for very flat surfaces, the feedback loop can be turned off and only the current is displayed. This is a "constant height" image [10].

1.3 Scanning tunneling microscopy

The scanning tunneling microscope is used to acquire images on the most extreme scale, easily resolving atomic structure in the right environments. This microscope can work at extremely low temperature around 4K as a tool for manipulating individual atoms or molecules to make remarkable structures. Figure 1-2 shows two STM images of fantastic structures from the IBM STM image gallery [11]. In order to interpret the STM image properly, we first discuss the concept of density of states.



(a) (b)

Figure 1-2: STM images. (a) Xenon atoms are manipulated to make the IBM letters on nickel (110). (b) Iron atoms are manipulated to make the Chinese characters for "atom" on copper (111). These images were taken from the IBM STM image gallery [11].

1.3.1 Density of states

Each electron within the sample has an energy level, and any given energy level can be occupied by only a certain number of electrons at any one time. Density of states is the quantity that shows the number of electrons allowed per energy level as a function of which energy level you consider. This quantity can be abbreviated as DOS(E) to indicate its dependence upon energy.

A local density of states (LDOS) is defined because the DOS can also vary as a function of position in the crystal. LDOS is a quantity that depends on energy as well as on position, LDOS (x, y, E)**.** This quantity can be interpreted physically as the density of electrons of a certain energy at that specific spatial position.

1.3.2 Topography

In practice, a topograph is measured at a fixed sample bias V with respect to the tip. All electrons with energy less than eV flow from the tip to the sample (or vice versa depending upon the sign of V). The magnitude of the tunneling current is determined by the total number of electrons with energy less than V in that specific location. Feedback is used to keep the current constant and the feedback voltage is recorded as a measure of the height of the sample at each point to produce the topographic image.

A STM constant-current topograph effectively displays the total density of states integrated between tip and sample Fermi levels as a function of location on the surface. The regions with high density of electrons are displayed as light color in greyscale images. The atoms are then bright spots in the topography as the electrons tend to gather together near them and thus produce a higher local density of electrons in that region. In order to maintain constant current, the feedback loop must pull the tip away from the surface by a small distance. As the tip scans from one location to another across the surface, the current increases or decreases depending on the LDOS and the tip height is adjusted to maintain the current constant. By making a map of where the tip height is large and where it is small, one can image the atomic positions.

In theory, the calculations of constant current topographs are challenging because for a realistic three dimensional model of the tip and the sample, calculating the tunneling current will be a fairly complex task. Fortunately, the tunneling can be treated with first-order perturbation

theory because of the weak coupling between the tip and sample (typical tip-sample separations of the order of 10 Å). Such a theory by Bardeen [12], together with the model of a tip by Tersoff-Hamann [5], leads to an expression for the tunneling current [5] (see Chapter 3 for details).

$$I \propto \sum_{v} \left|\psi_v\left(r_t\right)\right|^2 \delta\left(E_v - E_F\right) \equiv \rho\left(r_t, E_F\right) \qquad (1\text{-}1)$$

Thus the ideal STM would simply measure $\rho\left(r_t, E_F\right)$, the LDOS at E_F, i.e. the charge density from states at the Fermi level. The LDOS is evaluated for the bare surface at the position at which the tip is held. Thus within this model, STM has quite a simple interpretation as measuring a property of the bare surface, without reference to the complex tip-sample system.

For a semi-conducting surface, the DOS varies strongly with voltage and even changes discontinuously at the band edges. If the voltage between the tip and sample is small, the s-wave approximation for the tip [5] led to the very same result in equation (1-1). At larger voltages, equation (1-1) could be easily generalized to a simple expression such as

$$I\left(V\right) \propto \int_{E_F}^{E_F+V} \rho\left(E\right) T\left(E,V\right) dE \qquad (1\text{-}2)$$

where $\rho\left(E\right)$ is the LDOS at or very near the surface and a constant density of states for the tip is assumed. A barrier transmission coefficient $T\left(E,V\right)$ is included to account for the effect of the voltage on the surface wave functions. Equation (1-2) is not strictly correct because the matrix elements and the tip density of states are energy dependent, which are not considered in the equation. Discussion on the matrix elements and tip density of states can be found in the literature [5,12]. Nevertheless, equation (1-2) is a reasonable approximation for many purposes [13], as long as the voltage is much smaller than the work function.

1.4 Scanning Tunneling Spectroscopy

Measuring the DOS or LDOS can be very useful since this quantity is often used to derive information about the crystal's electronic properties. According to previous studies on STM tunneling theory, the LDOS is proportional to the differential change in tunneling current given a differential change in sample bias, or in other words, one can obtain the LDOS by measuring dI/dV.

In practice, a number of modulation techniques were developed to measure $I(V)$ and all the techniques are collectively referred to as scanning tunneling spectroscopy (STS). The technique used by Binnig and Frank [14] was to apply a high-frequency (chosen above the response of the feedback system) modulation voltage to the constant DC bias between sample and tip. Tunneling spectra, dI/dV versus V, were reported for clean and oxygen-covered Ni (100) and Si (111) surfaces [14]. Feenstra and Stroscio demonstrated the method of numerically differentiating the $I-V$ curve to obtain spectroscopic information [3]. Their work also showed that the quantity $(dI/dV)(V/I)$ is approximately proportional to the LDOS. This quantity is referred to as the normalized conductance spectrum. Wolkow and Avouris used normalized conductance spectra to study the dissociation of ammonia on the Si (111)-(7x7) surface with great precision, detecting, for example, the quenching of dangling bond states [15,16]. Finally, Hamers and Tromp obtained an energy-resolved, real-space image of occupied and unoccupied surface states of the Si (111)-(7x7) surface [2]. The technique is called current imaging tunneling spectroscopy (CITS). $I-V$ data is rapidly acquired using a sample-and-hold circuit to gate the feedback control system on and off at each fixed sample-tip position while simultaneously scanning the tip position [2].

1.5 Theoretical calculation of STM images

A proper way to interpret STM images quantitatively would be to calculate the image for a proposed structure or set of structures [17], and compare it with the actual image. It is possible to calculate STM images once the electronic structure for the tip and sample has been solved for a given surface geometry. However, for a realistic three-dimensional model of the surface, the calculation will be a fairly complex task. Fortunately, the tunneling can be treated with first-order perturbation theory because of the weak coupling between the tip and sample (typical tip-sample separations of the order of 10 Å). Such a theory by Bardeen [12], together with the model of a tip by Tersoff-Hamman [5], leads to an expression for the tunneling current [5]. This theory is presented in more detail in Chapter 3.

Tersoff's work has stimulated theoretical studies on constant-current-mode STM of a variety of conducting and semi-conducting surfaces. The essential problem is how to calculate the quantity $\rho(r_t, E_F)$. The atomic-charge superposition method [5] proposed by J. Tersoff was adopted to study the Si (111)-(7x7) surface in R. M. Tromp's work [18]. R. M. Tromp and R. J. Hamers compared experimental results to the simulated Si (111)-(7x7) images based on different models and concluded that a striking agreement is found with the DAS model [6] of Takayanagi. However, the bias dependence as well as the effect of the electric fields on wave functions has not been taken into consideration. A. Selloni and his co-workers [19] used a plane wave representation, along with a carbon local pseudo potential to conduct the STM theoretical study on surface and bulk electronic states of graphite. In their calculation, the voltage drop in the vacuum region has also been included using the WKB method. J. Pollmann [20] developed the self-consistent scattering theoretical method to investigate surface electronic properties of Si (001)-(2x1) and Ge (001)-(2x1). They calculated the contours of constant charge density at the surface plane. J. H. Wilson [21] used a tight-binding model to describe the energy dependence of

the surface wave functions for the relatively complex Si (113)-(3x2) surface. Quantitative agreement with the experiment was achieved. Joseph A. Stroscio [22] presented voltage-dependent scanning tunneling microscopy (STM) data from various semiconductor surfaces. The direct calculation was very successful especially for the GaAs (110) surface. The atomic selective imaging of the surface was confirmed by the direct calculation using equation (1-2). The direct calculations also infer the surface buckling quantitatively from the apparent atom positions in a real surface of GaAs (110), where the As atom buckles upward, and the Ga downward.

Even though all the previous works have contributed to the quantitative interpretation of STM imaging, many of them focused on calculating one-dimensional scan lines [19,22] or two-dimensional charge density or surface contours on a surface plane [19,20,23]. Only a few [18] presented the simulation of the direct, real-space observation of STM imaging of the Si (111)-(7x7) surface. The normalized conductance curve ($\left(dI/dV\right)\left(V/I\right)$ $vs.V$) plays an important role experimentally in revealing the electronic structure of the sample surface [16,24-26]. However, no work has been reported on the simulation of these curves over specific features on the Si (111)-(7x7) surface. Hydrogen atom adsorption onto the surface has also been extensively studied experimentally [25-27]. Unfortunately, simulated STM imaging of hydrogen atoms on the Si (111) surface have yet to be reported.

In our work, simulation results of both imaging and spectroscopy will be presented. Our approach to the local density of states utilizes tight-binding theory to expand the surface wave function into atomic wave functions localized on each atom nucleus. This means that we treat each atom on the surface separately. The electronic band structure of orbitals described by the wave function is modeled using J. D. Joannopoulos group's theoretical results [28].

1.6 Thesis outline

This thesis addresses the applicability of computer simulations of STM and STS imaging to the STM study of clean or atom-covered Si (111)-(7x7) surfaces. Chapter 2 introduces the STS technique and includes mathematical modeling of essential elements, including the feedback loop. In Chapter 3 we discuss a quantum mechanical model for the STM tunneling current that we believe is widely applicable to semiconductor systems. Chapter 4 demonstrates a simple atomic orbital model for the Si (111)-(7x7) surface and includes modeling of surface band states and bulk states. Chapter 5 presents the simulation results of STM and STS and compares them to the experimental results. Finally, conclusions and future work are included in Chapter 6.

Chapter 2
Experimental technique

2.1 Overview

As discussed in the introduction, a powerful application of STM is to measure, with extremely high lateral resolution, the local electronic density of states of a surface. To date, several techniques have been developed to obtain this information. These techniques can be divided into two categories according to whether the feedback loop is active or not during the spectroscopy. An example is the technique of CITS developed by Hamers *et al*. in which the feedback loop is deactivated for a short time while the tunneling current is measured at a prescribed set of voltage values. From this data, the dI/dV function can be calculated to give the electronic density of states at the precise location of the tip. Alternatively, a lock-in amplifier can be used to directly measure dI/dV [14,29]. Generally, one has a choice of conducting these spectroscopic measurements during the raster scan or after the raster scan. Taking the measurements after the raster scan has the drawback that one may be uncertain of the precise lateral position of the tip during the spectroscopy. This uncertainty can arise from thermal drift and piezo creep. Another drawback is that sophisticated control software is required to precisely position the STM tip. Taking the measurements during the raster scan alleviates these problems. However, taking complete spectral measurements at each pixel of an image can be very time consuming.

Taking spectral measurements during the raster scan and while the feedback loop is on is a relatively simple and fast way to produce a local density of states map (image) of the surface being studied. This lock-in amplifier technique was first used by Binnig *et al.* to study Ni (110)

[10,14]. This is the technique currently used in our lab to obtain local density of states information about molecules adsorbed to silicon surfaces. The main drawback to this technique is that the feedback loop has some response to the ac ripple applied to the sample. This generally means that the frequency of the ripple must be outside of the bandwidth of the feedback loop. For example, if the response of the feedback loop to the ripple is perfect, the tunneling current will not change i.e. dI/dV will be zero. The goal of this chapter is to analyze this response with the aim of maximizing the desired spectroscopic signal.

2.2 General description of feedback loop in time domain

As the main apparatus of a STM, the feedback loop is used to adjust the tip height relative to the sample to maintain a constant current between the tip and the sample. The one used in our laboratory is a commercial feedback system constructed by RHK Instruments (Model SPM 100). The feedback loop consists of a tunneling junction, a piezo element, and analog components as shown in Figure 2-1. Analog circuit components include a summer, a linear gain block, an integrator and a high voltage amplifier.

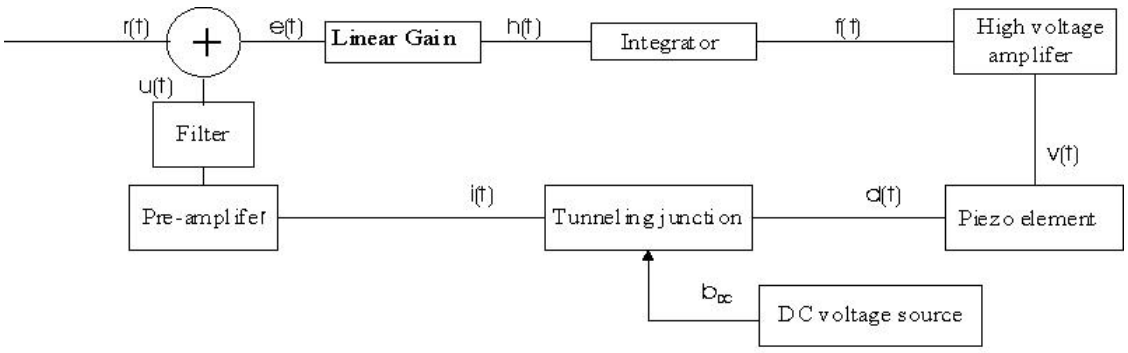


Figure 2-1: Diagram of RHK controller feedback loop

2.2.1 The tunneling junction

A tunneling junction refers to a conductive junction formed by a sharp tungsten tip and a sample with a small separation between them. The energy level in the vacuum between the tip and sample (vacuum energy level) is zero, which is greater than both of the Fermi levels of the tip and the sample. This means that a potential barrier exists between the tip and sample and their separation is the width of barrier. One rough approach to this problem would assume a rectangular potential barrier of width L (tip-sample distance) over the junction.

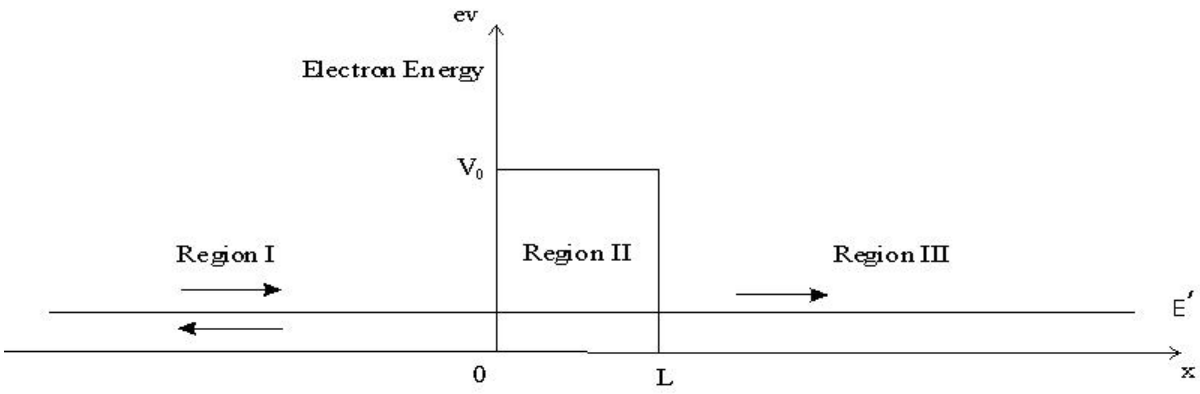


Figure 2-2: Steady state scattering by a rectangular barrier of width L

According to the basic quantum mechanics, the wave functions in the three distinct regions are expressed by:

$$\psi_I = A_0 e^{ikx} + Ae^{-ikx} \qquad (2\text{-}1)$$

$$\psi_{II} = Be^{-\alpha x} + Ce^{\alpha x} \qquad (2\text{-}2)$$

$$\psi_{III} = De^{ikx} \qquad (2\text{-}3)$$

k is the wave-vector of a particle with energy of E' i.e. $k = \sqrt{2mE'}/\hbar$. V_0 is the height of a rectangular barrier and α is the inverse length of decay of the wave function in region II: $\alpha = \sqrt{2m(V_0 - E')}/\hbar$. We define penetrability as $\chi = |D/A_0|^2$. Utilizing the boundary conditions, and after some algebra [7], we obtain the following simplified expression for the penetrability: $\chi \propto e^{-2\alpha L}$. The tunneling current i is proportional to the penetrability:

$$i \propto e^{-2\alpha L} \qquad (2\text{-}4)$$

If the effect of the voltage placed across the junction is taken into consideration, the potential barrier will change. There are three possible ways to apply the potential to the junction, which correspond to three different energy diagrams of the junction as shown in Figure 2-3. We represent the voltage in terms of sample bias.

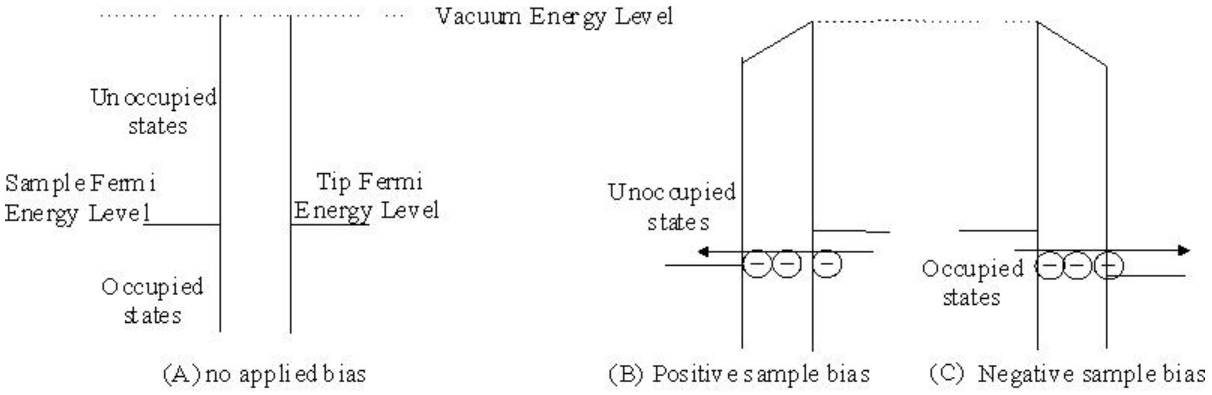


Figure 2-3: Energy diagram of tunneling junction.

If no voltage is applied onto the sample, the net tunneling current is zero between the tip and the sample. The reason is that the electrons would tunnel from both sides of the junction and

the tunneling in the two opposite directions balances each other. Figure 2-3 (B) shows the case where the sample bias is positive. In this case, the electrons tunnel from the occupied states of the tip into the unoccupied states of the sample, and so the tunneling current flows from the sample to the tip. The last case is where electrons in the occupied states of the sample tunnel through the barrier into unoccupied states of the tip in Figure 2-3 (C) when a negative voltage is applied on the sample. This means that the tunneling current flows from the tip to the sample. Therefore, two conditions are necessary for the occurrence of a tunneling current between the junction: (1) a potential is applied between the tip and the sample and (2) the tip and the sample are close to each other, which can also be seen from the tunneling current's exponential dependence on the tip-sample distance in equation (2-4). The second condition means that the current is very sensitive to the distance between the tip and sample. This makes it necessary to adjust the tip position precisely with reference to the sample surface (the sample is assumed to be fixed). The latter condition has been realized by the utilization of a piezo element on which the tip is mounted.

2.2.2 The piezo element

Piezo elements exhibit a piezoelectric effect: the length of a piezo element is adjustable if a voltage is applied. The variation in the length of the piezo tube is proportional to the voltage. Here we assume that the voltage signal which we apply is dependent on time. Therefore, the variation in the length of the piezo tube is a function of time and is given by:

$$d(t) = (l/T)\, d_{31} \mathrm{v}(t) \qquad (2\text{-}5)$$

where l is the relaxed length of the tube, T is the thickness of the wall, and d_{31} is the piezoelectric coefficient. Equation (2-5) ignores dynamical effects such as piezo tube resonance. Piezo tube resonance will be considered in section 2.3.1.

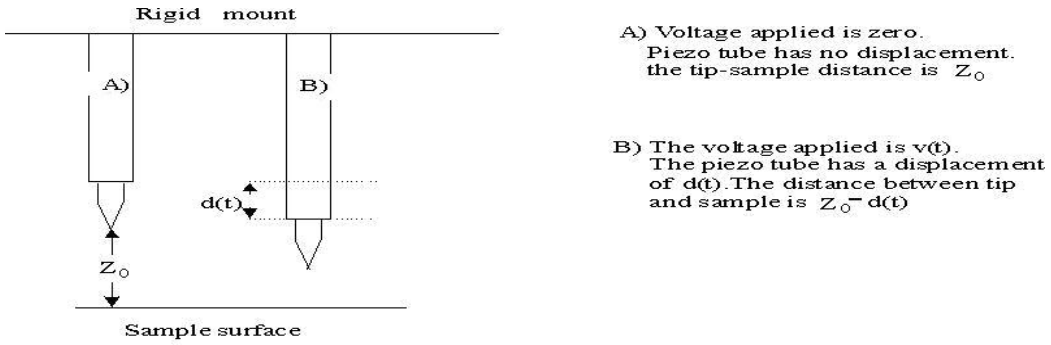


Figure 2-4: The separation between the tip and sample at (a) 0 and (b) V (t) Volts.

The tip-sample distance is z_0 when the piezo element stays in relaxation at $t \le 0$ on the left of Figure 2-4. Here we assume that z_0 is constant in time. i.e. we ignore thermal drift processes and we consider the situation where the tip is not scanning laterally across the surface (i.e. we are not imaging). On the right of the same figure, at time $t > 0$, v(t) is applied on the piezo tube and it results in the extension $d(t)$ of the piezo tube. The new tip-sample separation z is:

$$z = z_0 - d(t) \qquad (2\text{-}6)$$

Suppose that at $t \le 0$, the tunneling current is b_{DC}/R_J, where b_{DC} is the applied bias between sample and tip, and R_J is the effective resistance of the tunnel junction. Since the tip-sample distance is z_0, according to equation (2-4), we have

$$b_{DC}/R_J \propto e^{-2\alpha z_0} \qquad (2\text{-}7)$$

At time $t > 0$, the tip-sample distance changes into $z = z_0 - d(t)$ and the tunneling current correspondingly changes according to

$$i(t) \propto \exp\left[-2\alpha\left(z_0 - d(t)\right)\right] \qquad (2\text{-}8)$$

Combining equations (2-5), (2-7) and (2-8) we find:

$$i(t) = b_{DC}/R_J \, e^{2\alpha d(t)} = b_{DC}/R_J \, e^{2\alpha(l/T)d_{31}\mathrm{v}(t)} \qquad (2\text{-}9)$$

The following assumption is made: $2\alpha(l/T)d_{31}\mathrm{v}(t) << 1$ i.e. we assume that we will make only small perturbations from the situation described by z_0 , b_{DC}/R_J . With this assumption we can linearize the exponential in equation (2-9).

$$e^{2\alpha(l/T)d_{31}\mathrm{v}(t)} \approx 1 + 2\alpha\left(l/T\right)d_{31}\mathrm{v}\left(t\right) \qquad (2\text{-}10)$$

With the assumption that there is no filter between the summer and the pre-amplifier in Figure 2-1, the voltage output $u(t)$ from the pre-amplifier with resistance of R_{pa} is simply given by:

$$u(t) = i(t)R_{pa} = \left(b_{DC}/R_J\right)R_{pa}\left(1 + 2\alpha\left(l/T\right)d_{31}\mathrm{v}\left(t\right)\right) \qquad (2\text{-}11)$$

2.2.3 Analog circuit components

Our introduction of analog circuit components begins with a summer. Three voltage signals are related to this component. The two input voltage signals are the voltage output $u(t)$ from pre-amplifier and a pre-set voltage signal $r(t)$, while the output voltage signal is called the error signal $e(t)$. The feedback loop acts to try to maintain $e(t)$ at zero. The error signal is obtained from the difference of the two input signals

$$e(t) = r(t) - u(t) \qquad (2\text{-}12)$$

This error voltage signal is then processed by the linear gain amplifier (the gain is p). Its analog circuit inside the block is illustrated as follows.

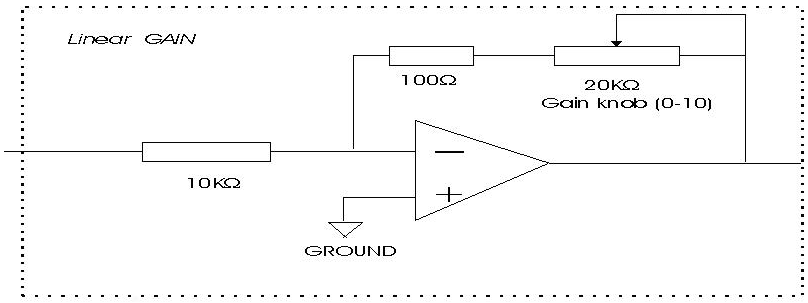


Figure 2-5: The analog circuit of linear gain

$$h(t) = pe(t) \tag{2-13}$$

The next component in the circuit is an integrator. The parameters for the integrator shown in Figure 2-6 are taken from the Model SPM 100 manual.

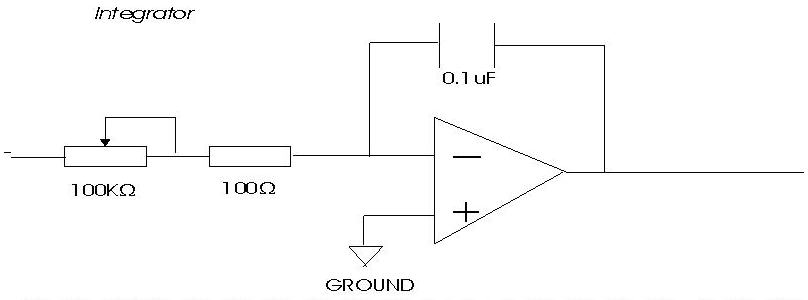


Figure 2-6: The analog circuit of integrator

The output signal from the integrator is then given by

$$f(t) = K_i \int_{-\infty}^{t} h\left(t'\right)dt' \tag{2-14}$$

where K_i is the integral gain. The signal of $f(t)$ is then amplified by a high voltage component by a factor of g_v to obtain the voltage signal $v(t)$ that is called the high voltage gain.

$$v(t) = g_v f(t) \tag{2-15}$$

This high voltage is then applied to the piezo tube to control the position of the tip. Combining equations (2-11)-(2-15) , we have

$$e(t) = r(t) - \frac{b_{DC}}{R_J} R_{pa} \left(1 + 2\alpha \left(l/T \right) d_{31} g_v K_i p \int_{-\infty}^{t} e(t')dt' \right) \qquad (2\text{-}16)$$

Note that we have "looped" around and we are left with an integral equation for $e(t)$ with an input function $r(t)$. As a simple example let us suppose $r\left(t\right)$ is a step function, defined by

$$r\left(t\right) = \frac{b_{DC} R_{pa}}{R_J} \text{ if } t \le 0 \text{ and } r\left(t\right) = r_1 \quad \text{if } t > 0 .$$

We are interested in finding the transient response in the error signal $e\left(t\right)$. We define:

$$1/\tau = \left(b_{DC} / R_J \right) R_{pa} 2\alpha (l/T) d_{31} g_v p K_i \qquad (2\text{-}17)$$

The physical meaning of the symbol τ is the time constant of the feedback. This will become apparent soon. Solving the equation (2-16) by Laplace transformation (Appendix B), the result turns out to be

$$e\left(t\right) = \left(r_1 - R_{pa}\, b_{DC} / R_J \right) e^{\frac{-t}{\tau}} \qquad (2\text{-}18)$$

We can see here that the physical quantity of τ can be identified as the time constant of the feedback loop. The disruption at $t = 0$ results in an exponentially decaying error signal. The system eventually re-adjusts to the new set-point value. In practice, the parameters p , K_i are user controlled settings on the RHK controller.

2.3 Frequency domain analysis of feedback loop with lock-in amplifier

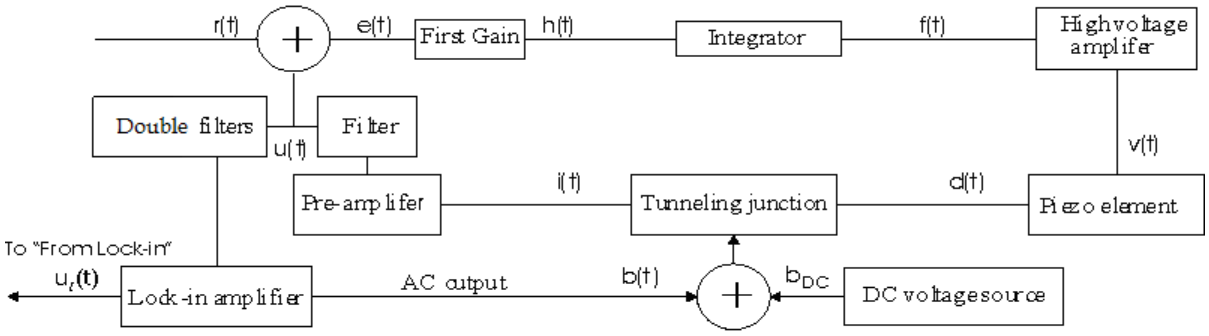


Figure 2-7: Feedback loop with Lock in amplifier

In STS (Scanning Tunneling Spectroscopy), a lock-in amplifier is hooked up to the circuit as shown in Figure 2-7. It supplies an ac voltage signal of certain frequency ω_0 to the tunneling junction. This ac signal will affect the tunneling current $i(t)$ and therefore the output signal from a double filter, which is also the input signal of the lock-in amplifier. From this input signal that may mix with a noise signal, the lock-in amplifier sorts out the signal of frequency ω' into its output channel. The sorted signal is then recorded into computer and the data is related to the STM spectroscopic imaging, also called conductance imaging. If we assume $G = dI/dV$ to be the conductance of the tunneling junction and ac signal voltage $b(t)$, the tunneling current is: $i(t) = b_{DC}/R_J + b(t)G$, at $t \le 0$ and at $t > 0$:

$$i(t) = \left(b_{DC}/R_J + b(t)G\right)e^{2\alpha d(t)} \approx \left(b_{DC}/R_J + b(t)G\right)\left(1 + 2\alpha d\left(t\right)\right) \qquad (2\text{-}19)$$

According to linear response theory, since $b(t)$ and $2\alpha d(t)$ are both small, we can ignore the non-linear term involving their product.

$$i(t) = b_{DC}/R_J + b_{DC}2\alpha d(t)/R_J + b(t)G \quad (2\text{-}20)$$

It is convenient to analyze the feedback loop in the frequency domain. Therefore the Fourier transform is used to transform the functions from the time domain into the frequency domain. The Fourier transform is defined as:

$$I(\omega) = \int_{-\infty}^{\infty} i(t)e^{-j\omega t}dt \quad (2\text{-}21)$$

Letters in upper case describe frequency domain functions while those in lower case represent time domain functions correspondingly. Thus, we can convert equations (2-12) to equation (2-15) from the time domain into the frequency domain as

$$E(\omega) = R(\omega) - U(\omega) \quad (2\text{-}22)$$

$$H(\omega) = pE(\omega) \quad (2\text{-}23)$$

$$F(\omega) = K_i H(\omega)/j\omega \quad (2\text{-}24)$$

$$V(\omega) = g_v F(\omega) \quad (2\text{-}25)$$

Using equation (2-21), equation (2-20) is transformed into the frequency domain,

$$I(\omega) = b_{DC}/R_J\,\delta(\omega) + b_{DC}2\alpha D(\omega)/R_J + B(\omega)G \quad (2\text{-}26)$$

Since there is no filter between the summer and the pre-amplifier,

$$U(\omega) = I(\omega)R_{pa} \quad (2\text{-}27)$$

Introducing equation (2-26)and equation (2-27) into equation (2-22), we obtain

$$E(\omega) = R(\omega) - R_{pa}\left(b_{DC}/R_J\,\delta(\omega) + b_{DC}2\alpha D(\omega)/R_J + B(\omega)G\right) \quad (2\text{-}28)$$

According to equation (2-5) and equations (2-23)-(2-25),

$$D(\omega) = d_{31} V(\omega) l / T = d_{31} g_v p K_i E(\omega) l / j\omega T \qquad (2\text{-}29)$$

Introducing equation (2-17) and equation (2-29) into equation (2-28), we obtain the following equation:

$$E(\omega)\left(1 + 1/j\omega\tau\right) = R(\omega) - R_{pa}\delta(\omega) b_{DC} / R_J - R_{pa} GB(\omega) \qquad (2\text{-}30)$$

The reference signal is typically constant in the STM constant current mode and this constant is chosen to satisfy the condition that the error signal is zero at $t \le 0$, therefore we consider

$$r(t) = r_0 = b_{DC} R_{pa} / R_J \qquad (2\text{-}31)$$

By transforming equation (2-31) into the frequency domain,

$$R(\omega) = b_{DC} R_{pa} / R_J \, \delta(\omega) \qquad (2\text{-}32)$$

With the introduction of equation (2-32) into (2-30), the first two terms cancel so that

$$E(\omega) = -j\omega\tau R_{pa} GB(\omega) / \left(1 + j\omega\tau\right) \qquad (2\text{-}33)$$

This expression given by equation (2-33) shows how, in frequency domain, the relationship between the error signal and the driving ac signal. Based on this equation, we have discussed the feedback response to a variety of signals such as a sine wave signal (see Appendix A for more detail).

2.3.1 The effect of the piezo tube resonance

As introduced earlier, the piezo tube plays an important role in the whole feedback loop. Therefore, for a strict analysis, it is necessary to take into consideration the resonance frequency of the piezo tube. A proper approach to this problem would describe the piezo tube as a single damped harmonic oscillator as illustrated in the Figure 2-8.

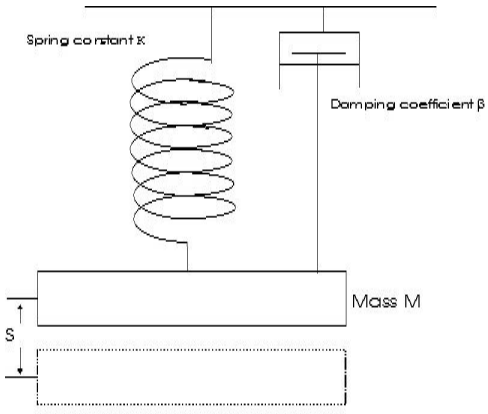


Figure 2-8: Damped single harmonic oscillator

The position $s(t)$, of the mass is described by Newton's second Law as:

$$m\ddot{s}(t) = -Ks(t) - \beta\dot{s}(t) + \phi(t) \qquad (2\text{-}34)$$

where $\phi(t)$ represents an external force, K is the spring constant and β is the damping coefficient.

Taking the Fourier transform of equation (2-34) leaves us with the following equation:

$$S(\omega) = \frac{\Phi(\omega)}{-m\omega^2 + K + \beta j\omega} \qquad (2\text{-}35)$$

The following two descriptions for the extension, $d = (l/T)d_{31}\mathrm{v}$ and $s = \frac{\phi}{K}$ must be the same in order to model the continuous piezo element as a harmonic oscillator, and therefore $\Phi(\omega) = K(l/T)d_{31}\mathrm{V}(\omega)$. Equation (2-35) can be rewritten:

$$S(\omega) = \frac{K(l/T)d_{31}\mathrm{V}(\omega)}{K + j\beta\omega - m\omega^2} \qquad (2\text{-}36)$$

The resonance frequency of the piezo tube is assumed to be ω_p, which can be related to the spring constant: $K = m\omega_p^2$. The theoretical calculation of the resonance frequency in longitudinal and bending modes can be found in Appendix D.

$$S(\omega) = \frac{m\omega_p^2 (l/T) d_{31} \mathrm{V}(\omega)}{m(\omega_p^2 - \omega^2) + j\beta\omega} \qquad (2\text{-}37)$$

If we define the quality factor as $Q = \frac{m\omega_p}{\beta}$ then we can rewrite $S(\omega)$ in equation (2-37) as:

$$S(\omega) = \frac{l d_{31} \mathrm{V}(\omega)/T}{\left(1 - \omega^2/\omega_p^2\right) + j\,\omega/\left(\omega_p Q\right)} \qquad (2\text{-}38)$$

The length variations of the piezo tube as described by equation (2-29) and equation (2-38) differ by a factor of $1/\left[\left(1 - \omega^2/\omega_p^2\right) + j\,\omega/\left(\omega_p Q\right)\right]$. This factor accounts for the effect of the piezo tube resonance. To simplify our discussion, we combined this factor into d_{31}. The new time constant τ' is thus related to τ in the frequency domain, according to equation (2-17).

$$1/\tau' = (1/\tau)/\left[\left(1 - \omega^2/\omega_p^2\right) + j\,\omega/\left(\omega_p Q\right)\right] \qquad (2\text{-}39)$$

Therefore, we can easily rewrite equation (2-33) as follows:

$$E(\omega) = -j\omega\tau' R_{pa} GB(\omega)/\left(1 + j\omega\tau'\right) \qquad (2\text{-}40)$$

Bringing equation (2-39) into (2-40) and rearranging the terms leads us to:

$$E(\omega) = \frac{\omega\left(\omega_p^2 - \omega^2 + j\,\omega\omega_p/Q\right) R_{pa} GB(\omega)}{\omega^3 - j\,\omega^2\omega_p/Q - \omega\omega_p^2 + j\,\omega_p^2/\tau} \qquad (2\text{-}41)$$

When $\omega \to 0$, the error signal goes to zero according to equation (2-41). This means that feedback loop response is very good and looks like a high pass filter. More discussion about equation (2-41).can be found in Appendices A and E.

2.4 Frequency optimization

The performance of STM spectroscopic imaging is related to the output ac voltage signal from a lock-in amplifier and thus is frequency dependent. The question that arises is what frequency (or frequencies) corresponds to the best imaging performance.

In this section, we approach the problem theoretically. Based on a theoretical model of the circuit (refer to appendix F), we can obtain the output ac voltage signal as a function of frequency. If the effect of noise is not considered, we favour the frequency that corresponds to the highest magnitude of the output signal since this magnitude determines how clear the image will be.

The modeling of the circuit that is involved in the spectroscopic imaging is based on experimental measurements. The analog circuit is shown in Figure 2-9. More information about the modeling can be found in Appendix F.

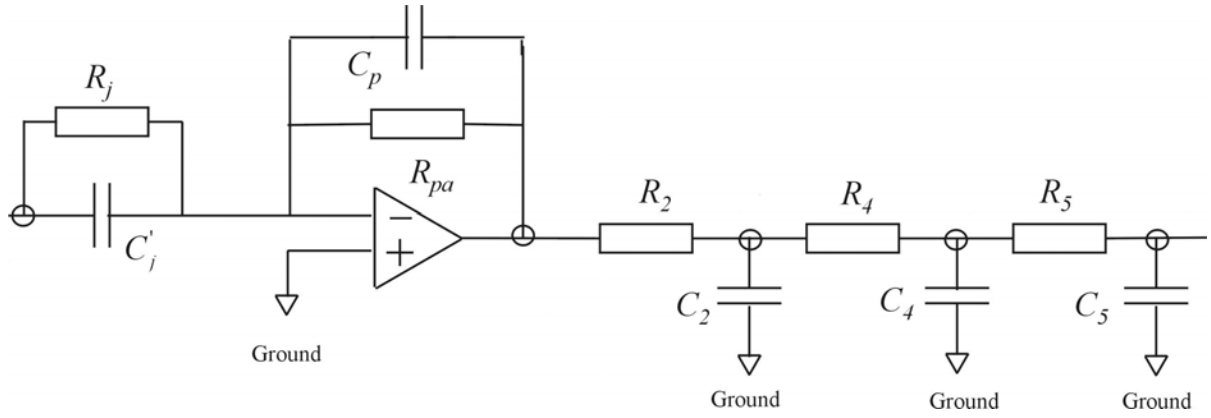


Figure 2-9: The model of the spectroscopic imaging circuit

In this model, we consider the capacitive coupling of C_j between the sample and the tip and the stray capacitance C_p over the resistance of pre-amplifier. It is not a difficult task, in the frequency domain, to include in the previous equations the two capacitances and the three filter stages. Due to the existence of the filter stage (*R2, C2*) between the pre-amplifier and the summer as well as the stray capacitance C_p, equation (2-27) becomes:

$$U(\omega) = I(\omega) R_{pa}^{'} F_2 \qquad (2\text{-}42)$$

where $R_{pa}^{'} = R_{pa} \big/ \left(1 + j\omega R_{pa} C_p\right)$ and $F_2 = 1 \big/ \left(1 + j\omega R_2 C_2\right)$. Correspondingly, equation (2-40) becomes:

$$E(\omega) = \frac{-j\omega\tau^{"} F_2 R_{pa}^{'} G^{'} B(\omega)}{1 + j\omega\tau^{"}} \qquad (2\text{-}43)$$

where $1/\tau^{"} = \left(1/\tau^{'}\right) F_2 \big/ \left(1 + j\omega R_{pa} C_p\right)$ and $G^{'} = G + j\omega C_j$.

According to equation (2-22)

$$U(\omega) = R(\omega) + \frac{j\omega\tau^{"} F_2 R_{pa}^{'} G^{'} B(\omega)}{1 + j\omega\tau^{"}} \qquad (2\text{-}44)$$

In Figure 2-7, not only does this voltage produce an error signal together with a reference voltage to activate the feedback loop as we discussed before, but after passing through a double filter stage located inside the RHK controller (*R4, C4*) and (*R5, C5*), the signal becomes the source signal from which the signal of interest with a specific frequency is sorted out by the lock-in amplifier. After a double filter stage, the signal becomes:

$$U^{'}(\omega) = F_4 F_5 U(\omega) \qquad (2\text{-}45)$$

where $F_4 = 1 \big/ \left(1 + j\omega R_4 C_4\right)$, $F_5 = 1 \big/ \left(1 + j\omega R_5 C_5\right)$ are from filter terms. If the voltage signal output from the lock-in to the sample bias is a sine wave, $b_0 e^{j\omega_0 t}$ then $B(\omega) = b_0 \delta(\omega - \omega_0)$, and $R(\omega) = r_0 \delta(\omega)$

In the time domain,

$$u'(t) = f_4 f_5 r_0 + \frac{f_2 f_4 f_5 R_{pa}\left(G + j\omega_0 C_j\right) b_0 e^{j\omega_0 t}}{1 + j\omega_0 R_{pa} C_p + f_2 \Big/ \left\{ \left(j\omega_0 \tau \right) \left[1 - \left(\omega_0 / \omega_p \right)^2 + j\,\omega_0 \Big/ \left(\omega_p Q \right) \right] \right\}} \quad (2\text{-}46)$$

where $f_4 = 1/\left(1 + j\omega_0 R_4 C_4\right)$, $f_5 = 1/\left(1 + j\omega_0 R_5 C_5\right)$ and $f_2 = 1/\left(1 + j\omega_0 R_2 C_2\right)$.

The first term denotes a dc signal, which will be filtered out by the lock-in amplifier and thus only the second term will be left in the lock-in output channel. The frequency domain function of the second term is :

$$U_{output}(\omega) = \left(G + j\omega C_j\right)\overline{Z} \quad (2\text{-}47)$$

where $\overline{Z} = \dfrac{F_2 F_4 F_5 R_{pa} B(\omega)}{1 + j\omega R_{pa} C_p + F_2 \Big/ \left\{ \left(j\omega\tau \right) \left[1 - \left(\omega / \omega_p \right)^2 + j\,\omega \Big/ \left(\omega_p Q \right) \right] \right\}}$

In the time domain

$$u_{output}(t) = \left(G + j\omega_0 C_j\right)\overline{z} \quad (2\text{-}48)$$

where $\overline{z} = \dfrac{f_2 f_4 f_5 R_{pa} b_0 e^{j\omega_0 t} / 2\pi}{1 + j\omega_0 R_{pa} C_p + f_2 \Big/ \left\{ \left(j\omega_0 \tau \right) \left[1 - \left(\omega_0 / \omega_p \right)^2 + j\,\omega_0 \Big/ \left(\omega_p Q \right) \right] \right\}}$.

In practice, one first retracts the tip almost 1 micron meter away so that the tunnel barrier is huge and the conductance is essentially zero. We obtain $U_{output}(t) = j\omega_0 C_j \overline{z}$. After that, we adjust the offset phase until an output phase of 90 degrees is achieved in the output channel of the lock-in amplifier. i.e. we effectively multiply (rotate) U_{output} by $e^{j\varphi}$ until $\overline{z}$ is real and positive. Under this condition, $\overline{z}$ has been adjusted to be a positive real number: $\overline{z} = |\overline{z}|$. Keeping everything unchanged, we make the tip approach the sample until a tunneling current exists. We

record the signals from the X-channel of the lock-in amplifier into the computer to generate the spectroscopic images. The X-channel output (real part output) is:

$$\frac{U_X}{GR_{pa}b_0} = \left| \frac{\omega_0 \tau f_4 f_5 \left[1 - \left(\omega_0 / \omega_p \right)^2 + j\,\omega_0 / \left(\omega_p Q \right) \right] / 2\pi}{\left(1 + j\omega_0 R_{pa} C_p \right)\left(1 + j\omega_0 R_2 C_2 \right)\left(j\omega_0 \tau \right)\left[1 - \left(\omega_0 / \omega_p \right)^2 + j\,\omega_0 / \left(\omega_p Q \right) \right] + 1} \right| \qquad (2\text{-}49)$$

Table 2-1: Parameter of the filter stages in the circuit (see Appendix F). The right column shows the cut off frequency (kHz) of each filter stage**.**

R_2	31.6 k	C_2	0.001 μF	0.50
R_4	10.0 k	C_4	220 pF	72.34
R_5	10.0 k	C_5	220 pF	72.34
R_{pa}	0.1 G	C_p	0.6 pF	26.53

Table 2-2: Important feedback loop parameters. The expected values are in parenthesis**.**

l	$0.5^{"}$	T	$0.025^{"}$
$d_{31}\,(m/V)$	$215*10^{-12}$	Q	(70)
K_i (s^{-1})	-199.6	p	-0.41
$\omega_p\,(kHz)$	(225.74)	$I\,(nA)$	1.0
α (m^{-1})	($1.1*10^{10}$)	g_v	13

The values for the parameters l, T and d_{31} of the piezo tube (BM527 type) are given in its handbook and displayed in Table 2.2. The quality factor Q is estimated to be 70. The values of integral gain and linear gain are obtained under a typical setup of buttons (gain: 2.0 and time constant: 5.0) in STM controller. The piezo tube resonance is chosen according to our theoretical calculation (see Appendix D for more detail). Here, we only use the lowest resonant frequency in the longitudinal mode. A typical tunneling current set-point value is 1.0 nA. The high voltage

gain, g_v has a value of 13. The value for α is calculated by assuming the work function of the tungsten tip used is $4.5eV$.

Using the parameters listed in table 2-1 and 2-2, we can now plot $U_X/\left(GR_{pa}b_0\right)$ as a function of $f_0 = \omega_0/\left(2\pi\right)$.

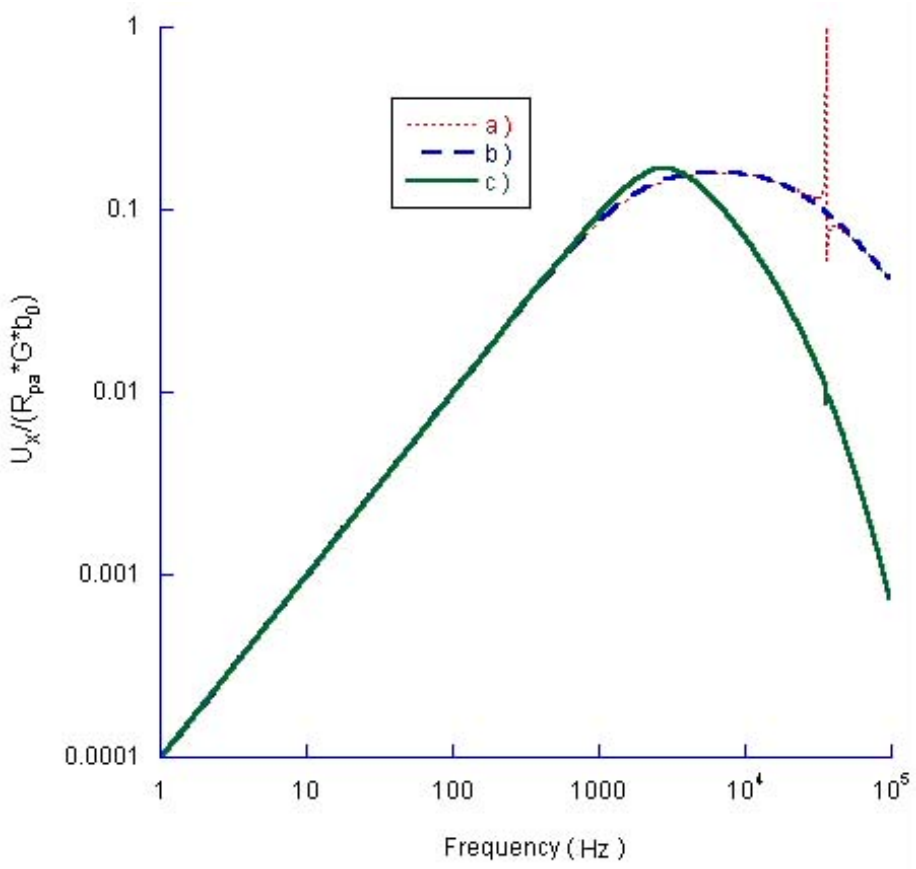


Figure 2-10: X-channel output as a function of the frequency. We obtain curves under the condition that we include a) no filter stages but piezo tube resonance b) none of them c) both of them into our calculation.

By comparing curve a) to curve b), the effect that the piezo tube resonance has on U_X is huge, as can be seen from the sharp peak at the resonant frequency of 35.93 kHz. After

including the filter stages, the peak observed in curve a) almost disappears in curve c). Also, after a frequency of around 4.0 kHz , curve c) drops off much more quickly than curve a) and curve b). The main reason is that the filter stage R_2C_2 has a comparatively low cut off frequency of 5.0 kHz and greatly filters out those signals with frequencies outside the cut-off frequency. Curve c) also shows us that the maximized magnitude of $U_X \big/ \left(GR_{pa}b_0 \right)$ corresponds to a frequency of around 3.0 to 4.0 kHz . This result is close to 4.5 kHz used in a previous study by Becker *et al.*[29].

2.5 Summary

Each component of the STM feedback loop has been examined and modeled. The feedback loop has been modeled in both the time domain and the frequency domain. The expression for the time constant of the feedback loop was obtained. This expression is useful for optimizing the gain settings. More importantly, the frequency response of the feedback loop to an AC signal has been calculated. The complicating effect of piezo tube resonance was also treated in the theoretical analysis. These calculations can be used to optimize the ripple frequency used in the experimental procedure for measuring local density of states. In the following chapter a theory is developed to relate the experimentally measured tunneling current to the local density of states.

Chapter 3
Theory for STM tunneling current

3.1 Overview

In this chapter, a theoretical approach for calculating the STM tunneling current is described [30]. In particular, it is shown how the experimentally measured tunneling current depends on the electronic local density of states. The starting point is the very general theory for tunneling developed by Bardeen, followed by a development of the Tersoff-Hamann model, with a refinement due to Selloni. This theory can be used to calculate the STM tunneling current, in general. Finally we follow Wilson *et al.* by using a tight-binding approximation for the electron wave functions of the sample, to add an important extension to the Tersoff-Hamann model.

3.2 Bardeen's approach

The tunneling process can be thought of as a many-electron problem. The probability $P_{i,f}$ of the system changing its state from a certain initial multi-electron state ψ_i into a final multi-electron state ψ_f can be used to describe this process. Therefore, it is necessary to calculate the many-particle, time-dependent wave functions of the entire system. This difficult task cannot be solved generally unless additional approximations are introduced. Here, the single-electron approximation will be used. This approximation implies that effects by multi-electron tunneling are not considered, since they are of minor importance to this work. In the single-electron picture the tunneling could be described as transitions from single-electron states on one side of the vacuum barrier into states on the other side. This makes it necessary to describe the single electron state in terms of eigen-functions of two separate systems. The following description of the tunneling process is based on Bardeen's approach to tunnelling [12].

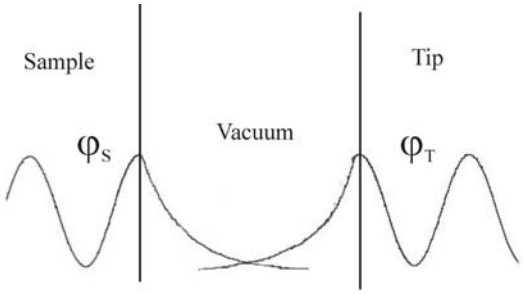


Figure 3-1: Bardeen's approach to the tunneling process.

This approach first applies time-dependent perturbation theory to the problem and then makes additional approximations. In Figure 3-1, the entire system is divided into two sub-systems: one for the sample and the other for the tip. Both sample and tip wave functions decay into the vacuum and their overlap accounts for the tunneling phenomena.

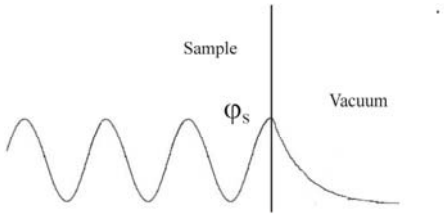


Figure 3-2: Wave function in the separated sample system

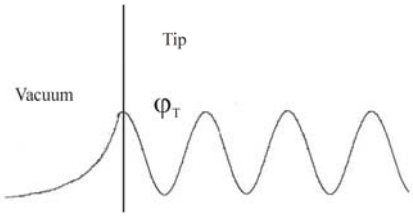


Figure 3-3: Wave function in the separated tip system

There exist two independent Schrödinger equations for the two systems as shown in Figure 3-2 and Figure 3-3 that are completely separated from each other

$$\left(T+U_s\right)\varphi_s=\varepsilon_s\varphi_s \qquad (3\text{-}1)$$

$$\left(T+U_T\right)\varphi_T=\varepsilon_T\varphi_T \qquad (3\text{-}2)$$

where U_S and U_T are the potentials of the tip and sample system, respectively, and T represents the operator of the kinetic energy of a single electron. The single-electron wave function of the total system is determined by:

$$H=T+U_S+U_T \qquad (3\text{-}3)$$

H is the Hamiltonian of the entire system. Time-dependent perturbation theory can be applied now to describe the tunneling of a single electron from the sample to the tip. The tunneling of an electron from the tip to the sample can be treated in exactly the same way. Because the initial state of the tunneling process in the sample is not influenced by the tip, there exists an eigenstate φ_s^μ of the sample system with $\left|\varphi\left(t\to\infty\right)\right\rangle=\left|\varphi_S^\mu\right\rangle$. The superscript μ represents the quantum numbers of a sample state. For example, μ could represent a band index and a wave vector. The time dependence of the state is determined by the Hamiltonian.

$$i\hbar\frac{\partial}{\partial t}\left|\varphi\left(t\right)\right\rangle=H\left|\varphi\left(t\right)\right\rangle \qquad (3\text{-}4)$$

According to basic quantum mechanics, the tunneling probability is simply given by the overlap of this time-dependent wave function with a tip wave function $\left|\varphi_T^\nu\right\rangle$.

$$\left\langle\varphi_T^\nu\right|\left(i\hbar\frac{\partial}{\partial t}\left|\varphi\left(t\right)\right\rangle\right)=\left\langle\varphi_T^\nu\right|H\left|\varphi\left(t\right)\right\rangle \qquad (3\text{-}5)$$

The Schrödinger equation (3-2) for the tip state is then used to obtain:

$$
\begin{aligned}
i\hbar\frac{\partial}{\partial t}\left\langle \varphi_T^{\nu}\middle|\varphi\right\rangle &= \left\langle \varphi_T^{\nu}\middle|\left(i\hbar\frac{\partial}{\partial t}\middle|\varphi(t)\right\rangle\right) - \left(\left\langle \varphi_T^{\nu}\middle|i\hbar\frac{\partial}{\partial t}\right)\middle|\varphi(t)\right\rangle \\
&= \left\langle \varphi_T^{\nu}\middle|H\middle|\varphi(t)\right\rangle - \left\langle \varphi_T^{\nu}\middle|H_T\middle|\varphi(t)\right\rangle \\
&= \left\langle \varphi_T^{\nu}\middle|U_S\middle|\varphi(t)\right\rangle
\end{aligned}
\qquad (3\text{-}6)
$$

The superscript ν represents the quantum numbers of a tip state. On the right side of the equation, we replace $|\varphi\rangle$ with $|\varphi(t\to\infty)\rangle = |\varphi_S^{\mu}\rangle$ and get a first-order perturbation theory

$$
i\hbar\frac{\partial}{\partial t}\left\langle \varphi_T^{\nu}\middle|\varphi\right\rangle = \left\langle \varphi_T^{\nu}\middle|U_S\middle|\varphi_S^{\mu}\right\rangle \qquad (3\text{-}7)
$$

Since φ_T^{ν} and φ_S^{μ} are eigenstates of two different Hamiltonians- $H_T = T + U_T$ and $H_S = T + U_S$, they do not form a complete orthogonal basis of the eigenspace of the total Hamiltonian $H = T + U_S + U_T$ of the entire system. Therefore, the total time dependence of $|\varphi\rangle$ cannot be determined, due to the insufficiency of the matrix elements at the right side, which is a basic weakness of Bardeen's approach. However, it has been shown in many applications [31,32] that this approach produces reliable results for systems that are well separated, i.e. systems where the overlap of the two wave functions is small. One may question whether perturbation theory can be applied at all because the sample potential is not small in the sample system. However, it is the matrix element of $\langle \varphi_T | U_S | \varphi_S \rangle$ that actually determines the strength of the perturbation of the initial state. Because the final wave function is localized in the tip region where the sample potential is very weak, this perturbation might still be considered as a small perturbation. Therefore, time-dependent perturbation theory will lead to reasonable results. The time dependence of the two states is separated:

$$
|\varphi_S^{\mu}\rangle = e^{i\varepsilon_{\mu}t}|\psi_S^{\mu}\rangle \text{ and } |\varphi_T^{\nu}\rangle = e^{i\varepsilon_{\nu}t}|\psi_T^{\nu}\rangle
$$

Then, the integration of equation (3-7) is performed and the time limit is applied to get the expression for the tunneling probability per time interval:

$$P_{\mu\nu}^{ST} = \lim_{t\to\infty} \frac{4}{t\hbar^2} \int_0^t \left| \left\langle \psi_T^{\nu} \left| U_S \right| \psi_S^{\mu} \right\rangle \right|^2 dt = \lim_{t\to\infty} \frac{4\sin^2\left(\left(\varepsilon_\mu - \varepsilon_\nu\right)t\right)}{\hbar\left(\varepsilon_\mu - \varepsilon_\nu\right)^2 t} \left| M_{\mu\nu}^{ST} \right|^2 \qquad (3\text{-}8)$$

where the matrix element $M_{\mu\nu}^{ST}$ is given by $\left\langle \psi_T^{\nu} \left| U_S \right| \psi_S^{\mu} \right\rangle$, the stationary-state matrix element of the potential. If a continuous range of energy levels ε_μ (or ε_ν) is assumed, the limit of equation (3-8) can be evaluated directly. One obtains

$$P_{\mu\nu}^{ST} = \frac{2\pi}{\hbar} \delta\left(\varepsilon_\nu - \varepsilon_\mu\right) \left| M_{\mu\nu}^{ST} \right|^2 \qquad (3\text{-}9)$$

This result looks like 'Fermi's Golden Rule' of standard time-dependent perturbation theory. It describes elastic tunneling from a sample state with energy ε_μ into a tip state with energy ε_ν only. The two values of energy have to be equal: $\varepsilon_\nu = \varepsilon_\mu$, which is taken care of by the delta function. In order to evaluate this matrix element, Bardeen introduces an additional approximation [12].

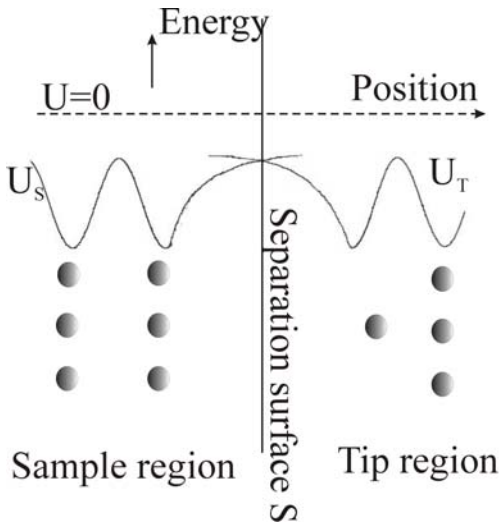


Figure 3-4: Potential setup used in the Bardeen approach to tunneling.

In his assumption, the tip potential U_T is zero in the sample region of space. Similarly the sample potential U_S should be zero in the tip region. A separation surface S is introduced to separate the regions where the two potentials U_T and U_S differ from zero. This can be described by the expression $U_S U_T = 0$ for any point in space. Of course, the smaller that the two potentials get as one goes further and further beyond the surface S, the better this approximation will become. Thus, it is expected that the approximation is valid for tunneling processes through a wide barrier, i.e. at large tip-sample distances, whereas the results of this approach will become invalid for very small tip-sample distances.

Since the sample potential is not zero only in the sample system, we can rewrite the matrix element as an integral over the sample region only instead of all space:

$$\begin{aligned} M_{\mu\nu}^{ST} &= \int_{sample} \psi_T^{\nu\,*} U_S \, \psi_S^{\mu} dV \\ &= \int_{sample} \psi_T^{\nu\,*} \left(\varepsilon_\mu + \frac{\hbar^2}{2m} \vec{\nabla}^2 \right) \psi_S^{\mu} dV \end{aligned} \qquad (3\text{-}10)$$

The Schrödinger equation (3-1) for the sample wave function allows the transformation in equation (3-10). The expression can be written in a more symmetric form.

$$\begin{aligned} M_{\mu\nu}^{ST} &= \int_{sample} \psi_T^{\nu *} \left(\varepsilon_\nu + \frac{\hbar^2}{2m} \vec{\nabla}^2 \right) \psi_S^\mu dV \\ &= \int_{sample} \left\{ \psi_T^{\nu *} \overleftarrow{\left(T + U_T \right)} \psi_S^\mu + \psi_T^{\nu *} \frac{\hbar^2}{2m} \vec{\nabla}^2 \psi_S^\mu \right\} dV \\ &= -\frac{\hbar^2}{2m} \int_{sample} \left\{ \psi_S^\mu \vec{\nabla}^2 \psi_T^{\nu *} - \psi_T^{\nu *} \vec{\nabla}^2 \psi_S^\mu \right\} dV \end{aligned} \qquad (3\text{-}11)$$

In the first step of these transformations, the eigenvalue ε_μ was replaced by ε_ν to satisfy the condition of energy conservation required in elastic tunneling. The Schrödinger equation for the tip wave function was used in the second transformation. Since the tip potential U_T is assumed to be zero in the sample region, it was omitted in the last step.

The integral in equation (3-11) can be transformed into a surface integral over S by using Green's theorem and the boundary condition that the tip wave function is zero at infinite distance from the tip.

$$M_{\mu\nu}^{ST} = -\frac{\hbar^2}{2m} \int_{surface} \left\{ \psi_S^\mu \vec{\nabla} \psi_T^{\nu *} - \psi_T^{\nu *} \vec{\nabla} \psi_S^\mu \right\} dS \qquad (3\text{-}12)$$

3.3 The tunneling current

We have so far obtained the probability rate of the transition of an electron from a sample state into a tip state. Based on equation (3-9), the probability rate can be written as

$$P_{\mu\nu}^{ST} = \frac{2\pi}{\hbar} \delta \left(\varepsilon_\mu^S - \varepsilon_\nu^T + eV \right) \left| M_{\mu\nu}^{ST} \right|^2 \qquad (3\text{-}13)$$

where the term eV is introduced to account for the potential V placed across the tip and the sample. The calculation of the tunneling current requires the addition of all possible tip and

sample states. Meanwhile, the tunneling of electrons in two opposite directions (from tip to sample or from sample to tip) should both be considered.

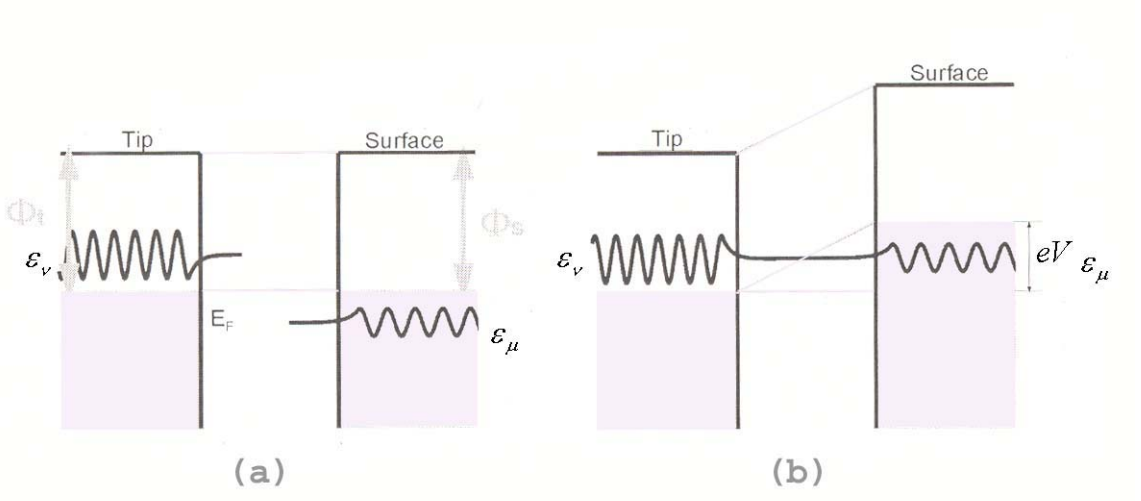


Figure 3-5: Schematic tip-sample energetics [33] (a) no bias applied (b) a positive voltage applied to the tip.

The total current is then given by

$$\begin{aligned} I &= I^{S\to T} - I^{T\to S} \\ &= 2e\sum_{\mu\nu} f_S(\varepsilon_\mu)\left(1-f_T(\varepsilon_\nu)\right)P_{\mu\nu}^{ST} - 2e\sum_{\mu\nu}\left(1-f_S(\varepsilon_\mu)\right)f_T(\varepsilon_\nu)P_{\nu\mu}^{TS} \\ &= 2e\sum_{\mu\nu}\left(f_S(\varepsilon_\mu)-f_T(\varepsilon_\nu)\right)P_{\mu\nu}^{ST} \\ &= 2e\sum_{\mu\nu}\left(f_S(\varepsilon_\nu - eV)-f_T(\varepsilon_\nu)\right)P_{\mu\nu}^{ST} \end{aligned} \qquad (3\text{-}14)$$

All sample states are labelled by u and all tip states are labelled by v. The function, $f(\varepsilon)$, represents the Fermi-distribution function that ensures that only tunneling from occupied to unoccupied states can occur. The factor 2 accounts for the two spin states of an electron. In equation (3-14), the symmetry of the tunneling probability $P_{\mu\nu}^{ST} = P_{\nu\mu}^{TS}$ is used in the last step of transformation. This symmetry can be deduced from equation (3-12). To evaluate the current, it

is necessary to calculate the matrix-elements $M_{\mu\nu}^{ST}$. In the next section, it will be shown how the matrix elements $M_{\mu\nu}^{ST}$ can be evaluated analytically after additional approximations regarding the tip orbitals are introduced.

3.4 Tersoff-Hamann's model of tip

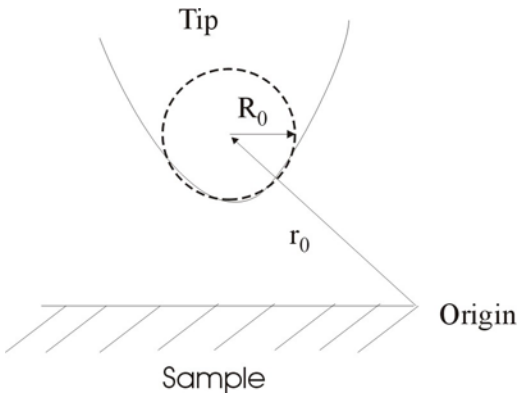


Figure 3-6: Tersoff-Hamann's model of tip structure.

Tersoff and Hamann [5] proposed that some assumptions be made on the tip system to simplify the calculation of Bardeen's matrix element $M_{\mu\nu}^{ST}$. This approach is most widely used in the interpretation of STM images. In their proposal, the electronic structure of the tip is described by a simple model system as shown in Figure 3-6. The geometrical structure at the end of the tip is assumed to be a sphere centered at position r_0. The radius of the sphere is R_0. The wave function of the outmost tip atom is assumed to be an atomic s wave function. The orbitals localized at the outermost tip atom will play the most important role in the tunneling process. We reach this conclusion on the basis of the following two facts: the tunneling current is related to the overlap of the wave function of the tip and the sample and the wave function decays

exponentially into the vacuum. In Tersoff and Hamann's model, the potential is assumed to be a spherical potential well as shown in Figure 3-7. This means that the potential has a constant value at any point inside the sphere and zero outside. This figure only describes the potential in one dimension of three-dimensional space

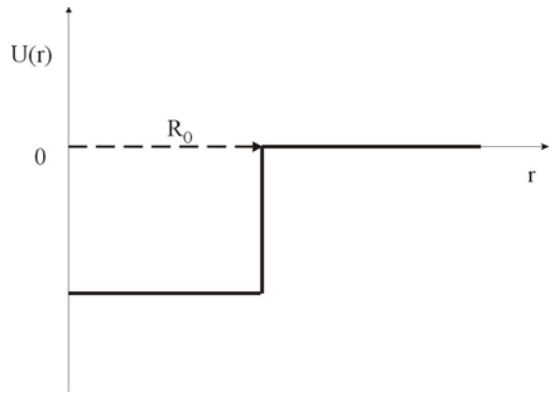


Figure 3-7: Tersoff-Hamann's model of potential.

Based on this assumption, ψ_T^v has the following form in the gap between the tip and the sample, i.e. for $r > R_0$, the tip radius , the general form of a wave function with s-symmetry centered at position $\vec{r}_0$ and with a vacuum-decay constant of κ_v. More details can be found in Appendix C.

$$\psi_T^v\left(\vec{r}-\vec{r}_0\right) = C\exp\left(-\kappa_v\left|\vec{r}-\vec{r}_0\right|\right)\Big/\kappa_v\left|\vec{r}-\vec{r}_0\right| \qquad (3\text{-}15)$$

$$G_v\left(\vec{r}-\vec{r}_0\right) = \frac{\kappa_v}{4\pi C}\psi_T^v\left(\vec{r}-\vec{r}_0\right) \qquad (3\text{-}16)$$

$$\left(\nabla^2-\kappa_v^2\right)G_v\left(\vec{r}-\vec{r}_0\right) = -\delta\left(\vec{r}-\vec{r}_0\right) \qquad (3\text{-}17)$$

The wave function of sample is also a solution of the vacuum Schrödinger equation in the tip region

$$\left(\nabla^2-\kappa_\mu^2\right)\psi_\mu^S = 0 \qquad (3\text{-}18)$$

If the energies of the tip and sample state are the same, the decay constant for the sample wave function will be the same as in equation (3-15).

$$\kappa_\nu = \kappa_\mu = \kappa \qquad (3\text{-}19)$$

Green's theorem is used to transform the integral in the expression for $M_{\mu\nu}^{ST}$ back into an integral over the tip region of space. We can obtain an expression:

$$\begin{aligned} M_{\mu\nu}^{ST} &= -\frac{\hbar^2}{2m}\int_{tip}\left\{\psi_T^{\nu *}\vec{\nabla}^2\psi_S^{\mu} - \psi_S^{\mu}\vec{\nabla}^2\psi_T^{\nu *}\right\}dV \\ &= -\frac{4\pi\hbar^2 C}{2m\kappa}\int_{tip}\left\{G_\nu\left(\vec{r}-\vec{r}_0\right)\kappa^2\psi_S^{\mu} - \psi_S^{\mu}\left(G_\nu\left(\vec{r}-\vec{r}_0\right)\kappa^2 - \delta\left(\vec{r}-\vec{r}_0\right)\right)\right\}dV \\ &= -\frac{2\pi\hbar^2 C}{m\kappa}\psi_S^{\mu}\left(\vec{r}_0\right) \end{aligned} \qquad (3\text{-}20)$$

This result involves the sample wave function evaluated at the position of the end of the tip. The constant C can be obtained by normalizing the wave function given by

$$C = c_t\kappa_\nu R_0 \cdot \exp\left(\kappa_\nu R_0\right)\Big/\sqrt{V_t}$$

Details for this normalization procedure are outlined in Appendix C. V_t is the volume of the tip. After substituting the expression in equation (3-20) for the matrix element in the equation (3-14) and replacing the summation over the eigen values by integration over the energy, one obtains

$$\begin{aligned} I &= 2e\cdot\frac{2\pi}{\hbar}\cdot\left(\frac{2\pi\hbar^2}{m\kappa}\right)^2\cdot C^2\int_{-\infty}^{\infty}\left(f\left(\varepsilon_\nu - eV\right) - f\left(\varepsilon_\nu\right)\right)n_T\left(\varepsilon_\nu\right)n_S\left(\varepsilon_\nu,\vec{r}_0\right)d\varepsilon_\nu \\ &= 2e\cdot\frac{2\pi}{\hbar}\cdot\left(\frac{2\pi\hbar^2}{m\kappa}\right)^2\cdot C^2\int_{-\infty}^{\infty}\left(f\left(\varepsilon\right) - f\left(\varepsilon + eV\right)\right)n_T\left(\varepsilon + eV\right)n_S\left(\varepsilon,\vec{r}_0\right)d\varepsilon \end{aligned} \qquad (3\text{-}21)$$

where n_T denotes the density of states (DOS) of the tip and

$$n_S\left(\varepsilon,\vec{r}_0\right) = \sum_\mu \left|\psi_\mu^S\left(\vec{r}_0\right)\right|^2 \delta\left(\varepsilon - \varepsilon_\mu\right) \qquad (3\text{-}22)$$

is the local density of the sample at the position $\vec{r}_0$ of the tip apex atom. In addition, to simplify the expression even further one normally assumes the density of states for the tip n_T to be a

constant and the Fermi distribution function to be a step function under the condition of low temperature. The current is then given as follows:

$$I = 16ec_t^2 \cdot \frac{\hbar^3\pi^3}{m^2} \cdot R_0{}^2 \exp\left(\kappa R_0\right) n_T / V_t \int_{\varepsilon_F}^{\varepsilon_F + eV} n_S\left(\varepsilon - eV, \vec{r}_0\right) d\varepsilon \qquad (3\text{-}23)$$

where $\kappa = \hbar^{-1}\left(2m\phi\right)^{1/2}$, is the minimum inverse decay length for the wave functions in vacuum, and ϕ is the work function of the tip. For tunneling at small voltage, where eV is small enough that $n_S\left(\varepsilon, \vec{r}_0\right)$ changes negligibly from ε_F to $\varepsilon_F + eV$, we obtain

$$I = 16e^2 V c_t^2 \cdot \frac{\hbar^3\pi^3}{m^2} \cdot R_0{}^2 \exp\left(\kappa R_0\right) n_T / V_t \, n_S\left(\varepsilon_F, \vec{r}_0\right) \qquad (3\text{-}24)$$

This equation developed by Tersoff and Hamann [5], provides a useful tool for the interpretation of STM images. As the tip is scanned laterally across the surface only $n_S\left(\varepsilon_F, \vec{r}_0\right)$ varies so the STM image can be interpreted as the local density of states of the sample. For STM studies on metal surfaces this equation is usually sufficient as one can tunnel at low biases. For semiconductors the bias is usually higher, typically 1.0 to 2.0 volts so we must use equation (3-23) and integrate over the sample DOS. This expression is commonly attributed to Selloni [19].

3.5 Tight-binding extension of Tersoff-Hamann's model

Here we have developed an extension of the Tersoff-Hamann model. The calculation of sample surface wave functions can be considerably simplified if we introduce an additional approximation. The tight binding approximation will be used. In equation (3-22), the state index μ is now substituted by band index i and wave vector $\vec{k}$.

$$
\begin{aligned}
n_S\left(\varepsilon,\vec{r}_0\right) &= \sum_{\mu}\left|\psi_{\mu}^{S}\left(\vec{r}_0\right)\right|^2 \delta\left(\varepsilon-\varepsilon_{\mu}\right) \\
&= \sum_{i,\vec{k}}\left|\psi_{i,\vec{k}}\right|^2 \delta\left(\varepsilon-\varepsilon_i\left(\vec{k}\right)\right)
\end{aligned}
\qquad (3\text{-}25)
$$

We first develop our extension of the Tersoff and Hamann theory for the case of a mono-atomic Bravais unit cell. The sample wave function can be expanded into a series of atomic wave functions, ϕ_i, centered at our Bravais surface lattice sites labelled by $\vec{R}$.

$$
\psi_{i,\vec{k}}\left(\vec{r}_0\right) = c\sum_{\vec{R}} e^{i\vec{k}\cdot\vec{R}}\phi_i\left(\vec{r}_0-\vec{R}\right)
$$

In order to obtain the normalization constant, c, we first square the wave function:

$$
\left|\psi_{i,\vec{k}}\left(\vec{r}_0\right)\right|^2 = |c|^2\sum_{\vec{R}}\left|\phi_i\left(\vec{r}_0-\vec{R}\right)\right|^2 + |c|^2\sum_{\vec{R}}\sum_{\vec{R}\neq\vec{R}} e^{i\vec{k}\cdot\left(\vec{R}-\vec{R}'\right)}\phi_i\left(\vec{r}_0-\vec{R}\right)\phi_i^*\left(\vec{r}_0-\vec{R}'\right) \qquad (3\text{-}26)
$$

If the tight-binding approximation is good then the terms in the second sum on the right side of the equation above are small because they contain the product of two atomic wave functions centered at different sites.

$$
\left|\psi_{i,\vec{k}}\left(\vec{r}_0\right)\right|^2 \approx |c|^2\sum_{\vec{R}}\left|\phi_i\left(\vec{r}_0-\vec{R}\right)\right|^2 \qquad (3\text{-}27)
$$

Now we can see that the $\vec{k}$ dependence is gone and the constant c can be determined by performing a normalization on both sides of equation (3-27). If we assume that N atomic sites are involved in our problem, the constant c is given by

$$
|c|^2 = \frac{1}{N} \qquad (3\text{-}28)
$$

Introducing equations (3-27) and (3-28) into (3-25) leads to

$$n_S(\varepsilon,\vec{r}_0)=\sum_{i,\vec{k}}\left|\psi_{i,\vec{k}}\right|^2\delta\left(\varepsilon-\varepsilon_i\left(\vec{k}\right)\right)$$
$$=\sum_i\sum_{\vec{R}}\left|\phi_i\left(\vec{r}_0-\vec{R}\right)\right|^2\frac{1}{N}\sum_{\vec{k}}\delta\left(\varepsilon-\varepsilon_i\left(\vec{k}\right)\right) \qquad (3\text{-}29)$$

If we define $g_i(\varepsilon)=\frac{1}{N}\sum_{\vec{k}}\delta_{\vec{k}}\left(\varepsilon-\varepsilon_i\left(\vec{k}\right)\right)$ as the normalized DOS for band i

i.e. $\int_{-\infty}^{\infty}g_i(\varepsilon)\,d\varepsilon=1$, then

$$n_S(\varepsilon,\vec{r}_0)=\sum_i g_i(\varepsilon)\sum_{\vec{R}}\left|\phi_i\left(\vec{r}_0-\vec{R}\right)\right|^2 \qquad (3\text{-}30)$$

Equation (3-23) can then be rewritten as

$$I=P\int_{\varepsilon_F}^{\varepsilon_F+eV}\sum_i g_i(\varepsilon)\sum_{\vec{R}}\left|\phi_i\left(\vec{r}_0-\vec{R}\right)\right|^2 d\varepsilon \qquad (3\text{-}31)$$

where $P=16ec_t^2\cdot\frac{\hbar^3\pi^3}{m^2}\cdot R_0{}^2\exp\left(\kappa_v R_0\right)n_T/V_T$.

Equations (3-30) and (3-31) are readily generalized to the case where the primitive surface unit cell contains more than one atom:

$$n_S(\varepsilon,\vec{r}_0)=\sum_{i,\alpha}g_{i,\alpha}(\varepsilon)\sum_{\vec{R}}\left|\phi_{i,\alpha}\left(\vec{r}_0-\vec{R}-\vec{d}_\alpha\right)\right|^2 \qquad (3\text{-}32)$$

$$I=P\int_{\varepsilon_F}^{\varepsilon_F+eV}n_S(\varepsilon,\vec{r}_0)\,d\varepsilon \qquad (3\text{-}33)$$

3.6 Free electron model for tip DOS

The pre-factor P is estimated by using free gas theory to describe the behaviour of electrons in the probe tip. In the free electron gas theory, the density of states is given by

$$\frac{n_T(\varepsilon_F)}{V_T} = \frac{3}{2}\frac{n}{\varepsilon_F} \qquad (3\text{-}34)$$

a_0 is the Bohr radius and r_s is the radius of an atomic s orbital. A value of $10.00\ eV$ for the Fermi energy and a value of $n \approx 10^{28}\ (m^{-3}/eV)$ for the density of valence electrons are typical for elements such as tungsten, gold, and platinum, commonly used for tips in STM studies. The work function of tungsten is $\phi = 4.5\ eV$, which gives $\kappa = \hbar^{-1}(2m\phi)^{1/2} = 1.087\times10^{10}\ m^{-1}$. If we choose the tip radius R_0 to be 1.0 Angstrom, the pre-factor has a value of $3.806\times10^{-12}\left(m^3\right)(C/s)$.

3.7 Summary

The STM tunneling current can be theoretically calculated using the theory developed by Bardeen, Tersoff and Hamann. An expression for this tunneling current, most useful for our purposes, is shown in equation (3-23). This expression earlier shows how the tunneling current depends on the local density of states function, n_s. In order to calculate the tunneling current for the semiconductor surface such as silicon, we have used the tight-binding approximation and developed expressions (3-32) and (3-33) for the tunneling current. In particular, these equations allow us to calculate the tunneling current for the Si (111)-(7x7) surface.

Chapter 4

STM tunneling current model for the Si (111)-(7x7) surface

4.1 Overview

The extended Tersoff-Hamann model, developed in the previous chapter, leads to a simple expression for the tunneling current. As can be seen from the expression (3-32), calculating the current requires both the atomic orbital wave function at each atomic site on the surface as well as the electronic density of states for the atomic orbital bands. Therefore, it is necessary to understand the crystal structure and electronic structure of the sample surface, Si (111)-(7x7) in our case. The Si (111)-(7x7) surface is obtained by a process called reconstruction. When the silicon (111) surface is heated to a sufficiently high temperature under UHV (ultra high vacuum) conditions, the surface atoms rearrange into a more energetically stable configuration. One of the configurations is the 7x7 reconstruction, which has been determined to be the most energetically favourable Si (111) geometry [23]. Other possible configurations include 2x1, 5x5, and 9x9 reconstructions. The 7x7 reconstruction involves several types of atomic positions in the top three atomic layers to form a unit cell much larger than the unit cell of the unreconstructed surface. To describe this configuration, many models have been introduced after the reconstructed surface was first observed experimentally. At present, the generally accepted model for the Si (111)-(7x7) surface is the Dimer-Adatom-Stacking (DAS) fault model by Takayanagi [6]. This acceptance is largely due to a set of theoretical calculations performed by Tromp *et al.* [18]. In these calculations STM images were simulated for various candidate models for the atomic structure of the surface. Comparison to the measured STM images showed striking agreement with the DAS model and all of the other models were ruled out. Though the simple approach of Tromp *et al.* was highly successful, some important effects were left out.

These include density of states variations from atom to atom, the electrical field due to the STM tip, as well as relaxations of the atomic positions. Apart from the electrical field effect, these effects have been considered in the calculations by the Joannopoulos group [28]. This group uses highly sophisticated density functional calculations for the surface wave functions. It is our goal to include all of these effects into a simple atomic orbital model for the simulations of STM images.

4.2 Introduction to the Si (111)-(7x7) surface

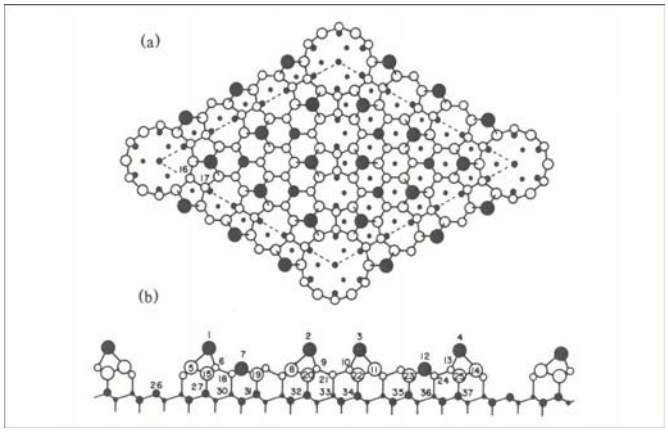


Figure 4-1: The DAS model of one unit cell of Si(111)-(7x7) surface a) top view b) side view [23].

The DAS model contains sub-surface dimers, twelve adatoms, and a stacking fault in half of the unit cell. The driving force behind this reconstruction is a minimization of the surface free energy, primarily by a reduction in the number of surface dangling bonds. Figure 4-1 a) shows one unit cell of the DAS model. The complexity of this model reflects the considerable atomic rearrangement that the system undergoes in order to minimize the number of broken covalent bonds on the surface without introducing excessive strain. Each unit cell of the DAS model contains 19 dangling bonds. The same surface area holds 49 dangling bonds before the rearrangement. This dramatic decrease in the number of surface dangling bonds accounts for the

more stable surface configuration with a large reduction of surface free energy. The nineteen dangling bonds in the surface layer stem from nineteen surface atoms including twelve adatoms (large dark dots), six rest atoms (small dark dots), and one corner hole (CH) atom (at the bottom of the hole). Each adatom and rest atom share three covalent bonds called "back bonds" with their three neighbouring atoms. The CH atom can be identified as the surface atom on the fourth atomic layer in the side view of Figure 4-1 b). For some reasons that will be apparent later, the adatoms are divided into two groups. Those next to CHs are referred to as "corner adatoms", while the rest we refer to as "middle adatoms". Another two terms frequently used in the DAS model are "faulted" and "unfaulted". In Figure 4-1 a) the unit cell is split into two sub-cells. The one on the left we refer to as the stacking faulted half while the other one is the unfaulted half. The faulted and unfaulted halves are "zipped" together by dimers at the interface between the halves.

4.3 Description of atomic positions in unit

4.3.1 Ideal positions

In order to apply equations (3-32) and (3-33) to the Si (111)-(7x7) surface there are several considerations that need to be made. Here we address the summation α over intra-unit-cell Si atoms with positions. As just discussed, there are many such atoms in this unit cell. The DAS model has been successful in describing the Si (111)-7x 7 surfaces qualitatively. However, the model is not sufficient to give a quantitative description of the surface. For example, the positions of surface atoms cannot be determined precisely due to the complexity of the physics on the surface such as minimized surface free energy and surface stress. Therefore, in order to localize each surface atom, we will introduce an additional approximation. We assume that all the surface atoms on the top three atomic layers are in the same perfectly tetrahedral arrangement

as those in the bulk. This means that for now we will not consider effects such as bond stretching or bending.

This extended model has a regular geometry and the location of any atom with reference to an origin can be easily determined. For illustrative purposes, we only show how the positions of adatoms and rest atoms are determined. The figure below shows the atomic positions in the top two atomic layers.

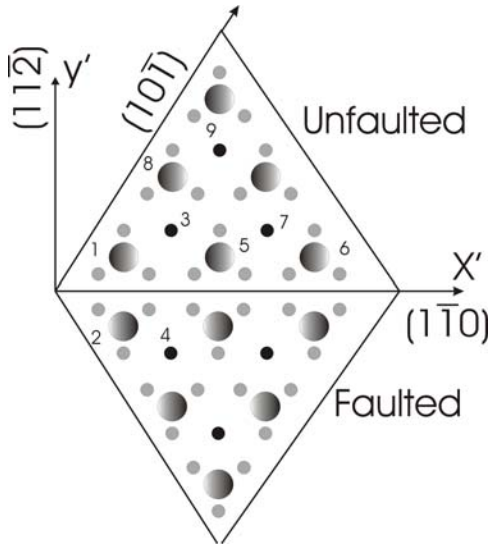


Figure 4-2: The top view of the top two atomic layers in the extended DAS model

The small solid circles are localized in the second layer. The small dots in gray are the neighbouring atoms of adatoms identified as the big solid shaded dot in the in the top layer. The small dark dots denote the rest atoms. The distance between the two layers is $\sqrt{6}a/12$ and a is the length of the standard cubic lattice side. The $x^{'}$ axis divides the unit cell into two sub-cells: the one above the axis is the unfaulted half cell and the other one is the faulted half cell. A single unit cell is focused on for now. The surface is the (111) surface, which means that the unit vector

of $(1\ 1\ 1) = 1/\sqrt{3}(\hat{x}+\hat{y}+\hat{z})$ is perpendicular to the page and points outward. The $\hat{x}$, $\hat{y}$, $\hat{z}$ vectors lie along the standard directions associated with the conventional FCC bulk unit cell. There are three vectors $(1\ \bar{1}\ 0)$, $(1\ 0\ \bar{1})$ and $(1\ 1\ \bar{2})$ on the surface as shown in Figure 4-2. In the unfaulted half cell, adatoms labelled by 1, 5 and 6 on the top atomic layer are evenly aligned along $(1\ \bar{1}\ 0)$, and adatoms labelled by 1, 8 and 9 are evenly aligned along $(1\ 0\ \bar{1})$. In each "line", the distance between two neighbour atoms is $\sqrt{2}a$. This means that if we can find out the position of adatom 1, the rest of adatoms can be located simply by stepping over a certain distance along a known vector. The same principle will be applied to find out the positions of the rest atoms labelled by 3, and 7. We treat the atoms in the faulted half cell exactly the same way. The question now becomes how to find out the positions of one adatom and one rest atom.

The origin is set up at the position of one of the CH atoms on the fourth layer. The relative location of the first adatom to the origin is shown in Figure 4-3 and Figure 4-4.

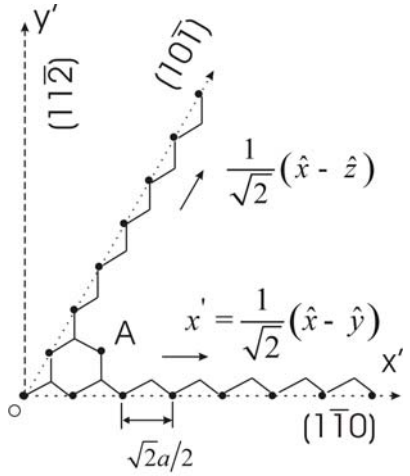


Figure 4-3: The top view of the arrangement of some atoms in the fourth layer.

In Figure 4-3, we indicate the directions of the x', y' axes on the 4th layer and the z' axis points out of the plane. The atoms are evenly distributed aligned along the x' axis and along the $(1\ 0\ \bar{1})$ direction. Along these directions, the distance between two neighbouring atoms is $\sqrt{2}a/2$. The site A is the projection point of the adatom 1(on the top layer), which means the adatom 1 is directly above the site A. The addition of vectors $(1\ \bar{1}\ 0)$ and $(1\ 0\ \bar{1})$ and basic knowledge of geometry, gives us the expression for vector $\overrightarrow{OA}$

$$\overrightarrow{OA}/|OA| = \frac{1}{\sqrt{6}}(2\hat{x} - \hat{y} - \hat{z}) \qquad (4\text{-}1)$$

Since the distance $|OA|$ between the two points O, A is $\sqrt{6}a/2$, therefore

$$\overrightarrow{OA} = \frac{a}{2}(2\hat{x} - \hat{y} - \hat{z}) \qquad (4\text{-}2)$$

.

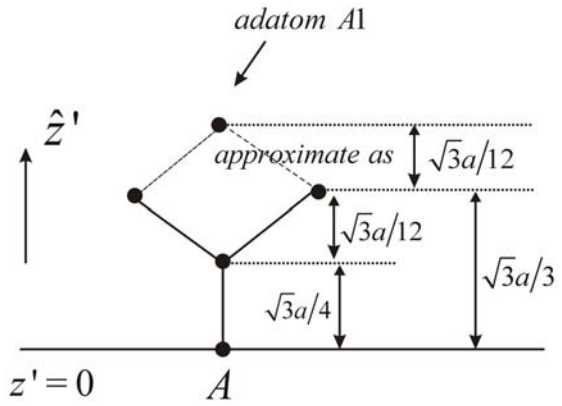


Figure 4-4: Side view of part of the top four layers of the DAS model.

In Figure 4-4, adatom A1 is on the top layer and A is on the fourth layer. The distances between two neighbouring layers are simply obtained by geometrical calculation. In our extended DAS model, the distance between the first and second layer is assumed to be the same as that between the second and third layer. Figure 4-4 apparently gives us the vector of $\overrightarrow{AA1}$, which is along the $\hat{z}'$ or $(1\,1\,1)$ direction.

$$\overrightarrow{AA1} = \left(\frac{\sqrt{3}a}{12} + \frac{\sqrt{3}a}{3}\right)\frac{1}{\sqrt{3}}(\hat{x}+\hat{y}+\hat{z}) = \frac{5a}{12}(\hat{x}+\hat{y}+\hat{z}) \qquad (4\text{-}3)$$

The vector pointing toward adatom $A1$ from the origin can be obtained as

$$\overrightarrow{OA1} = \overrightarrow{OA} + \overrightarrow{AA1} = \left(\frac{17a}{12}\hat{x} - \frac{a}{12}\hat{y} - \frac{a}{12}\hat{z}\right) \qquad (4\text{-}4)$$

This vector is the $\vec{d}_\alpha$ vector in equation (3-32) for this particular adatom dangling bond. To get the other adatoms' positions in this unfaulted half of the unit cell, we move in steps of

$$\sqrt{2}a\frac{1}{\sqrt{2}}(\hat{x}-\hat{y}) = a(\hat{x}-\hat{y}) \text{ and } a(\hat{x}-\hat{z})$$

i.e. $ia(\hat{x}-\hat{y}) + ja(\hat{x}-\hat{z})$ where $i = 0,1,2$ and $j = 0,1,2$ $i+j \le 2$.

We have so far, found out all the locations, $\vec{d}_\alpha$, of the six adatoms in the unfaulted half cell. To find the position of the rest atom R_1, the only thing we need to do is to find the vector connecting the adatom A_1 to R_1.

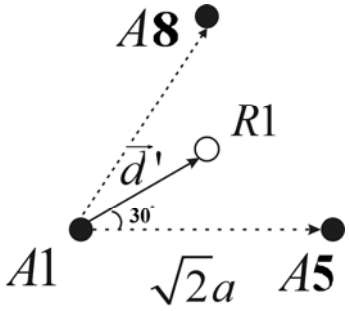


Figure 4-5: Top view of the rest atom R_1 relative to three nearby adatoms. The lateral distance d' between A1 and R1 is $\sqrt{2/3}\,a$.

The rest atom R_1 is on the second atomic layer while the three adatoms are in the top atomic layer. The distance between two adatoms is $\sqrt{2}a$. $\vec{d}'$ (added arrow) is the top-layer vector that connects the point A_1 to R_1's projection point in the top atomic layer.

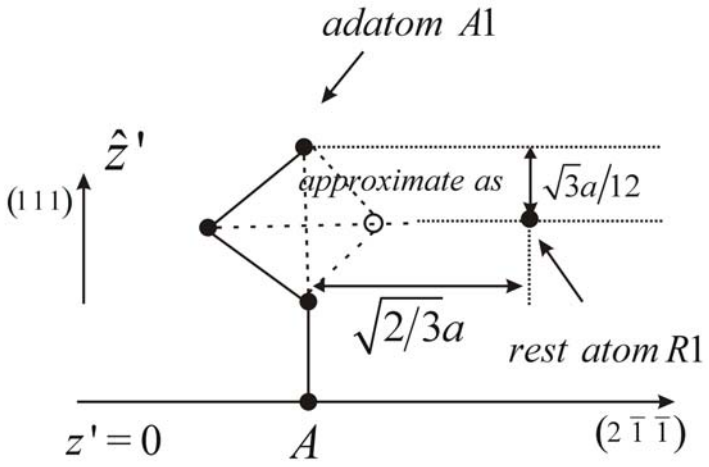


Figure 4-6: Side view of the relative positions of R_1 and A_1. The distance between A_1 and R_1 in the direction of $\left(2\ \bar{1}\ \bar{1}\right)$ is $\sqrt{2/3}a$, which can be obtained in the Figure 4-5.

From Figure 4-6, one can easily find the vector,

$$\overrightarrow{A1R1} = \frac{7a}{12}\hat{x} - \frac{5a}{12}\hat{y} - \frac{5a}{12}\hat{z} \qquad (4\text{-}5)$$

$$\overrightarrow{OR1} = \overrightarrow{OA1} + \overrightarrow{A1R1} = 2a\hat{x} - \frac{a}{2}\hat{y} - \frac{a}{2}\hat{z} \qquad (4\text{-}6)$$

We add $a(\hat{x}-\hat{y})$ to $\overrightarrow{OR1}$ and $a(\hat{x}-\hat{z})$ to $\overrightarrow{OR1}$ to get the positions of the other two rest atoms. We have so far found all the positions of adatoms and rest atoms in the unfaulted half of the unit cell. In the faulted half cell, as I have mentioned before, the task can be accomplished using the same method.

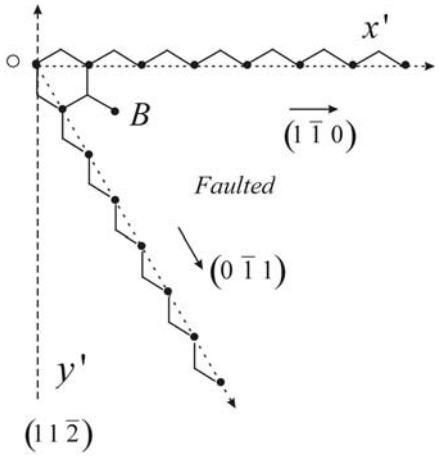


Figure 4-7: The top view of the arrangement of some atoms in the fourth atomic layer.

In Figure 4-7, the setup of x', y' axes happens on the 4th layer and z' axis points out of the plane. The atoms are evenly distributed along $(1\ \bar{1}\ 0)$ and $(0\ \bar{1}\ 1)$. In each line, the distance between two neighbour atoms is $\sqrt{2}a/2$. The site B is the projection point on the plane of the adatom $B1$ (on top layer), which means the adatom $B1$ is directly above the site B.

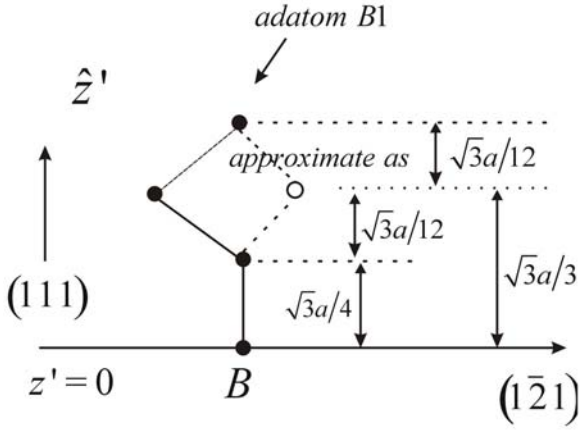


Figure 4-8: Side view of part of the top four layers.

Adatom $B1$ is on the top atomic layer and B is on the fourth atomic layer. The distances between two neighbouring layers are simply obtained by geometrical calculation. In our extended DAS model, the distance between the first and second layer is assumed to be the same as that between the second and third layer.

$$\overrightarrow{OB1} = \overrightarrow{OB} + \overrightarrow{BB1} = \left(\frac{11a}{12}\hat{x} - \frac{7a}{12}\hat{y} + \frac{11a}{12}\hat{z} \right) \qquad (4\text{-}7)$$

To get the other adatoms in this half of the unit cell, we move in steps of

$\sqrt{2}a\frac{1}{\sqrt{2}}(\hat{x}-\hat{y}) = a(\hat{x}-\hat{y})$ and $a(-\hat{y}+\hat{z})$

i.e. $ia(\hat{x}-\hat{y}) + ja(-\hat{y}+\hat{z})$ where $i = 0,1,2$ and $j = 0,1,2 \quad i+j \le 2$.

All of the six adatoms in the faulted half of the unit cell are now located.

4.3.2 Corrected positions

In a real Si (111)-(7x7) surface, surface silicon atoms are not in their ideal bulk positions due to the surface relaxation process. The fully relaxed atomic geometry can be determined

theoretically by using *ab initio* theory. For example, the Joannopoulos group has undertaken such an investigation and presented their results of the relaxed atomic positions in the top 4 layers [34]. In this section, only the relaxation of adatoms and rest atoms is taken into consideration. The relaxation of other atoms is neglected since the atoms are localized in the lower layers and play a less important role in STM imaging.

Table 4.1 shows that the ideal locations of adatoms and rest atoms calculated in 4.3.1 are transformed into the reduced coordinates [35] (X, Y, Z) in the Joannopoulos group's work [34] and compared correspondingly to the relaxed positions.

Table 4-1: Idealized locations of adatoms and rest atoms compared to relaxed positions in parentheses

		Atom	X (Å)	Y (Å)	Z (Å)
Adatoms	Faulted half	1	1.5(1.5)	1.5(1.5)	1.00(1.662)
		2	4.5(4.492)	4.5(4.491)	1.00(1.594)
	Unfaulted half	3	6.0(6.002)	6.0(6.002)	1.00(1.555)
		4	9.0(9.005)	9.0(9.005)	1.00(1.603)
Rest atoms	Faulted	7	2.5(2.481)	2.5(2.481)	0.00(0.348)
	Unfaulted	12	8.0(8.019)	8.0(8.019)	0.00(0.345)
CH		26	0.0(0.001)	0.0(0.001)	-4.0(-3.986)

This table tells us that these atoms relax mainly out of the silicon bulk along the vertical Z direction and the lateral relaxation is relatively small.

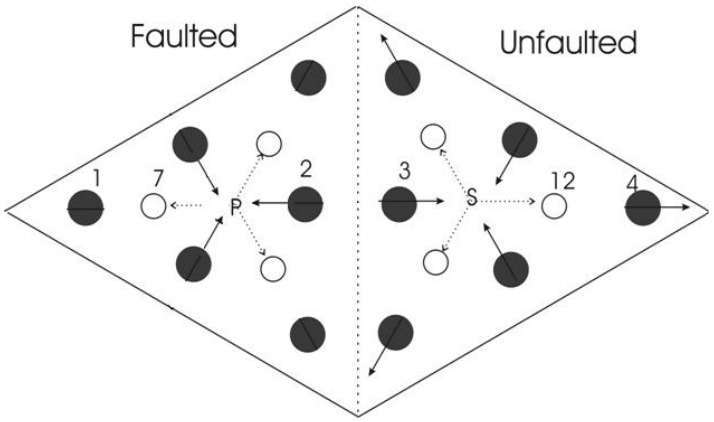


Figure 4-9: The relaxation of adatoms (large filled circles) and rest atoms (small open circles). The arrows show the lateral movement of the atoms on the surface

In Figure 4-9, P and S are the geometrical centers of the faulted half and the unfaulted half respectively. Three corner adatoms in the unfaulted half relax to be closer to their CHs while those in the faulted halves do not relax laterally. The middle adatoms in the unfaulted halves and the faulted halves tend to approach to point S and P respectively. The rest atoms always tend to spread out as shown by the dashed arrows.

4.4 Atomic orbital model

Next we consider the precise nature of the atomic orbitals, $\phi_{i,\alpha}$ in equation (3-32). We describe the Si (111)-(7x7) surface by an atomic orbital model. In this picture, the surface is thought of as a system that consists of a large number of silicon atoms. In quantum mechanics, the atomic orbitals are the quantum states that electrons surrounding an atom may reside in. These orbitals can be described as a wave function over space by the quantum numbers n, l, and m of the orbital, or by the names of the orbitals such as $1s, 2s, 2p$ etc., as used in electron configurations. The simplest atomic orbitals are those that occur in an atom with only one

electron, such as the hydrogen-atom. In this case the atomic orbitals are the eigenstates of the exactly solvable hydrogen atom Hamiltonian. An atom of any other element ionized down to a single electron in the outmost shell is very similar to hydrogen, and the orbitals take the same form. For atoms like silicon with two or more valence electrons, the governing equations can only be solved by using methods of iterative approximation. Orbitals of multi-electron atoms are qualitatively similar to those of hydrogen, and in the simplest model, they are taken to have the same form. The outermost shell of a silicon atom holds four valence electrons in four atomic orbitals, one $3s$ orbital and three $3p$ orbitals ($3p_x$, $3p_y$, and $3p_z$) respectively. For simplicity, we describe the four basic orbitals of a silicon atom by hydrogen-like wave function. For a more rigorous and precise analysis, a numerical approximation must be used. In three-dimensional space with the atomic nucleus of an isolated atom as the origin, the wave functions for the four atomic orbitals are given by

$$\varphi_{3s} = \frac{1}{81\sqrt{3\pi}}\left(\frac{Z_{eff}}{a_0}\right)^{3/2}\left(27-18\frac{Z_{eff}r}{a_0}+2\frac{Z_{eff}^2r^2}{a_0^2}\right)\exp\left(-\frac{Z_{eff}r}{3a_0}\right) \quad (4\text{-}8)$$

$$\varphi_{3p_z} = \frac{\sqrt{2}}{81\sqrt{\pi}}\left(\frac{Z_{eff}}{a_0}\right)^{3/2}\left(6-\frac{Z_{eff}r}{a_0}\right)\frac{Z_{eff}r}{a_0}\exp\left(-\frac{Z_{eff}r}{3a_0}\right)\frac{z}{r} \quad (4\text{-}9)$$

$$\varphi_{3p_x} = \frac{\sqrt{2}}{81\sqrt{\pi}}\left(\frac{Z_{eff}}{a_0}\right)^{3/2}\left(6-\frac{Z_{eff}r}{a_0}\right)\frac{Z_{eff}r}{a_0}\exp\left(-\frac{Z_{eff}r}{3a_0}\right)\frac{x}{r} \quad (4\text{-}10)$$

$$\varphi_{3p_y} = \frac{\sqrt{2}}{81\sqrt{\pi}}\left(\frac{Z_{eff}}{a_0}\right)^{3/2}\left(6-\frac{Z_{eff}r}{a_0}\right)\frac{Z_{eff}r}{a_0}\exp\left(-\frac{Z_{eff}r}{3a_0}\right)\frac{y}{r} \quad (4\text{-}11)$$

where Z_{eff} is the effective nuclear charge of the silicon atom, and a_0 is the Bohr radius. The overlap of the atomic orbitals of the participating atoms has been used in valence bond theory to

explain the formation of a covalent bond between two atoms [36]. However, the atomic orbitals for bonding may not be basic atomic orbitals directly from the solution of the Schrödinger equation. Often, the bonding atomic orbitals have a hybrid character of several possible types of orbitals. Mixing one s and three p atomic orbitals produces a set of four equivalent sp^3 hybrid atomic orbitals. The four sp^3 hybrid orbitals points toward the vertices of a tetrahedron. As a result of the tetrahedral arrangement of silicon atoms in our study, we know that the silicon atom has four atomic hybrid orbitals formed by the hybridization of four silicon atomic orbitals- $3s$ orbital, $3p_x$ orbital, $3p_y$ orbital and $3p_z$ orbital. The wave functions describing these hybrid orbitals are simply linear combinations of the wave functions for the four basic atomic orbitals. Assuming that a hybrid orbital is oriented along the unit vector $\langle a\ b\ c \rangle$, and then the general expression for its wave function is:

$$\varphi_{a\,b\,c} = \tfrac{1}{2}\varphi_{3s} + \tfrac{\sqrt{3}}{2}\left(a\,\varphi_{3p_x} + b\,\varphi_{3p_y} + c\,\varphi_{3p_z}\right) \qquad (4\text{-}12)$$

According to the DAS model for the Si(111)-(7x7) surface, the dangling bond orbitals on adatom and rest atom sites are indeed sp^3 hybrid orbitals along the (111) direction. Therefore, the wave function describing the dangling bond orbital is given by

$$\varphi_{111} = \tfrac{1}{2}\left(\varphi_{3s} + \varphi_{3p_x} + \varphi_{3p_y} + \varphi_{3p_z}\right) \qquad (4\text{-}13)$$

According to molecular orbital theory, the covalent bond between two silicon atoms can be considered as a molecular orbital, which can be obtained by combining the hybrid atomic orbitals on the two silicon atoms. One of the molecular orbitals, called the bonding orbital, σ, is constructed by adding wave functions for the two hybrid atomic orbitals. Another orbital named

as the anti-bonding orbital, σ^*, is formed by subtracting one of these wave functions from the other. In the extended DAS model, the hybrid orbitals on each atom (those orbitals overlapping to form the bond) are oriented directly toward the other atom. So the orientation of the two hybrid orbitals, opposite to each other, can easily be found from the relative coordinates of the two atoms. Following equation (4-12), we can quickly write down the wave functions for the two hybrid orbitals and therefore the wave functions for the bonding and anti-bonding orbitals between the two nuclei. The wave functions for the bonding and anti-bonding orbitals of a specific adatom back bond (the position of adatom is labelled by (x_a, y_a, z_a) and one of its neighbour atoms is located at (x_n, y_n, z_n)) can be determined. The orientation of the hybrid orbital on the adatom is determined to be along the unit vector (x', y', z'), where

$$x' = \frac{(x_n - x_a)}{\sqrt{(x_a - x_n)^2 + (y_a - y_n)^2 + (z_a - z_n)^2}} \qquad (4\text{-}14)$$

$$y' = \frac{(y_n - y_a)}{\sqrt{(x_a - x_n)^2 + (y_a - y_n)^2 + (z_a - z_n)^2}} \qquad (4\text{-}15)$$

$$z' = \frac{(z_n - z_a)}{\sqrt{(x_a - x_n)^2 + (y_a - y_n)^2 + (z_a - z_n)^2}} \qquad (4\text{-}16)$$

According to equation (4-12), the wave function for the hybrid orbital centered on the adatom is given by

$$\varphi_{x'y'z'} = \frac{1}{2}\varphi_{3s} + \frac{\sqrt{3}}{2}\left(x'\varphi_{3p_x} + y'\varphi_{3p_y} + z'\varphi_{3p_z}\right)$$

Because the two hybrid orbitals on the two atoms point toward each other, the wave function for the other hybrid orbital centered at one of the three neighbour atoms is given by

$$\varphi_{-x'-y'-z'} = 1/2\,\varphi_{3s} + \sqrt{3}/2\left(-x'\varphi_{3p_x} - y'\varphi_{3p_y} - z'\varphi_{3p_z}\right)$$

Therefore, the wave functions φ_σ (φ_{σ^*}) for the bonding (anti-bonding) orbital of the adatom backbonds are:

$$\varphi_\sigma(x,y,z) = \left[\varphi_{x'y'z'}(x-x_a, y-y_a, z-z_a) + \varphi_{-x'-y'-z'}(x-x_n, y-y_n, z-z_n)\right]/\sqrt{2}$$

$$\varphi_{\sigma^*}(x,y,z) = \left[\varphi_{x'y'z'}(x-x_a, y-y_a, z-z_a) - \varphi_{-x'-y'-z'}(x-x_n, y-y_n, z-z_n)\right]/\sqrt{2}$$

In summary, from the point of view of atomic orbitals, the surface is a large system that consists of hybrid atomic orbitals, bonding orbitals and anti-bonding orbitals. The latter two types of orbitals can be obtained by addition or subtraction of hybrid atomic orbitals.

4.5 Effect of electric field

R. M. Tromp [18] simulated the Si (111)-(7x7) image and his results favoured the DAS model by Takayanagi [6], which has now been widely accepted. However, in his calculation, the effect of electric fields is not considered. The electric field in the realistic tip-sample system is fairly complicated. For simplicity, we assume the tip-surface system to be two closely-spaced metal plates. This assumption implies a simple one-dimensional triangular potential barrier over the tip-surface system, which is shown in Figure 4-10.

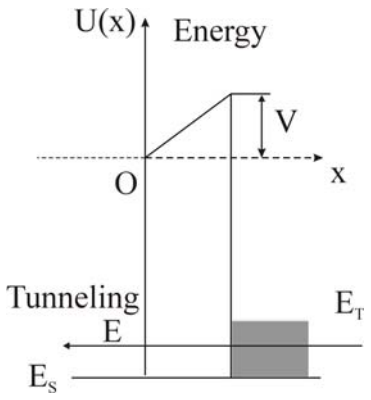


Figure 4-10: One dimension triangle potential barrier.

The sample on the left is positively biased. The potential difference is V (volt). The potential in the vacuum region between tip and sample is a linear function of the distance from the sample. The gradient is V / d_0, where d_0 is the tip-sample distance, or the barrier width. Electrons with energy E tunnel from an occupied state in the tip region to an unoccupied state in the sample region. According to the tunneling effect, electrons will tunnel across the potential barrier. The decay of the wave function of electrons from the sample is characterized by a tunneling coefficient. The WKB method leads us to an expression for the tunneling coefficient:

$$t = \exp(\frac{1}{\hbar}\int_0^{d_0}\sqrt{2m\left(U\left(x\right)-E\right)}dx) \qquad (4\text{-}17)$$

where $U\left(x\right)=\frac{V}{d_0}x$.

In our discussion, d_0 is actually the coordinate r_0 between the tip center and an atom. The reason is that the tight binding approximation we use treats the sample surface as the summation over each individual atom. In this picture, each atom is treated separately and we can say that the potential difference is placed across the tip and an individual atom on the sample

surface. Therefore we make the tunneling problem one-dimensional by taking the width of the potential barrier d_0 as the distance between the tip center and the specific atom. The integral in expression (4-17) can be evaluated explicitly as

$$t = \exp\left(\frac{d_0}{3m\hbar V}\left(t_1 - t_2\right)\right)$$

where $t_1 = \left[2m\left(V-E\right)\right]^{3/2}$ and $t_2 = \left[2m\left(-E\right)\right]^{3/2}$. The term $\exp\left(-\frac{\bar{Z}_{eff} r}{3a_0}\right)$ in equations (4-8) to (4-11) describing the exponential decay of the wave function into all space has to be replaced by the tunneling coefficient to account for the effect of the electric field. This substitution gives the correct functional form for the wave function decay far away from the surface. The new set of wave functions are then given by

$$\varphi_{3s} = \frac{t}{81\sqrt{3\pi}}\left(\frac{\bar{Z}_{eff}}{a_0}\right)^{3/2}\left(27 - 18\frac{\bar{Z}_{eff} r}{a_0} + 2\frac{\bar{Z}_{eff}^2 r^2}{a_0^2}\right) \qquad (4\text{-}18)$$

$$\varphi_{3p_z} = \frac{\sqrt{2}t}{81\sqrt{\pi}}\left(\frac{\bar{Z}_{eff}}{a_0}\right)^{3/2}\left(6 - \frac{\bar{Z}_{eff} r}{a_0}\right)\frac{\bar{Z}_{eff} r}{a_0}\frac{z}{r} \qquad (4\text{-}19)$$

$$\varphi_{3p_x} = \frac{\sqrt{2}t}{81\sqrt{\pi}}\left(\frac{\bar{Z}_{eff}}{a_0}\right)^{3/2}\left(6 - \frac{\bar{Z}_{eff} r}{a_0}\right)\frac{\bar{Z}_{eff} r}{a_0}\frac{x}{r} \qquad (4\text{-}20)$$

$$\varphi_{3p_y} = \frac{\sqrt{2}t}{81\sqrt{\pi}}\left(\frac{\bar{Z}_{eff}}{a_0}\right)^{3/2}\left(6 - \frac{\bar{Z}_{eff} r}{a_0}\right)\frac{\bar{Z}_{eff} r}{a_0}\frac{y}{r} \qquad (4\text{-}21)$$

The question arises how to determine the value of the effective charge $\bar{Z}_{eff}$. Considering a silicon atom isolated in vacuum, the tail of its wave function goes like $\exp\left(-\frac{\bar{Z}_{eff} r}{na_0}\right)$ by

solving the Schrödinger equation, where n is the principal quantum number. Using the WKB method, the tail of the wave function depends on $\exp\left(-\frac{\sqrt{2m(-E)}r}{\hbar}\right)$ (there is no potential difference between the tip and the sample, i.e. a rectangular potential barrier is assumed). Since the two methods should lead to the same decay length, we should have:

$$\frac{Z_{eff}}{na_0} = \frac{\sqrt{2m(-E)}}{\hbar}.$$

Since those states close to the Fermi level contribute most to the tunneling process, a proper choice of the energy would be the silicon Fermi energy (i.e. $E = -4.5eV$). The four silicon valence electrons stay in the third shell, and n is equal to 3. This will lead us to the effective nuclear charge: $Z_{eff} \approx 1.8$. We note that this value is considerably smaller than the value of 4.15 which is obtained using Slater's rules [36] for atomic silicon. The reason for this disparity is that a surface silicon atom experiences much more shielding and is also immersed in a surface dipole, i.e. much the same reason for why work functions are generally lower than ionization potentials.

4.6 Charge transfer

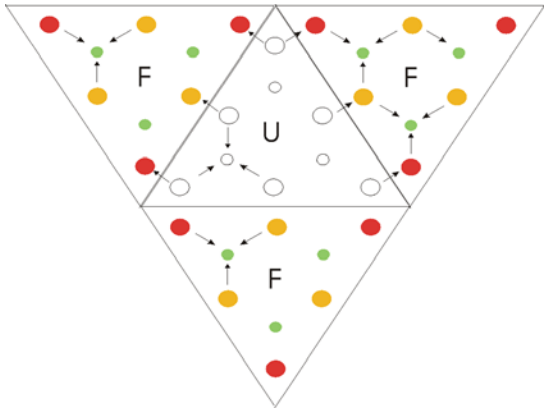


Figure 4-11: Charge transfer on the Si (111)-(7x7) surface.

In Figure 4-11, three faulted half cells and one unfaulted half cell are shown and labelled by F and U respectively. The large dots denote adatoms and small dots represent rest atoms. The arrows indicate the direction in which charges or electrons are transferred. The charge transfer process [37] has been used to explain the experimental observations that the middle adatoms are not as bright as the corner adatoms in STM imaging. (The intensity or brightness is proportional to the measured surface charge density at the tip position). Another experimental STM observation is that the adatoms in the faulted halves are brighter than those in the unfaulted halves. This is most likely explained by the charge transfer between the adatoms in different halves.

Figure 4-11 illustrates how the charges are possibly transferred on the Si (111)-(7x7) surface. The electrons transfer from adatoms to rest atoms presumably because the bonds holding the adatoms are more strained than the rest atom back-bonds. The results of Wolkow and Avouris [16] supported this transfer process by showing that the rest-atom db state is fully occupied since its corresponding strong occupied band is localized at about 0.8 eV below the Fermi level. It is also suggested in his work that the charge transfer is more efficient for middle

adatoms because they have two neighbouring rest atoms while the corner adatoms have only one rest atom. Charge transfer probably happens between two closest neighbour adatoms in different halves of cell. The direction of the transfer is from the adatoms in unfaulted half of cells to the adatoms in the faulted. The driving force behind the charge transfer is the minimization of surface free energy since all the electrons tend to reside in lower energy quantum states. The dangling bonds (adatom or rest atom), before charge transfer, are all half-filled states, which means that the states hold a single electron. The energy level of the rest atom db state is about -0.8 eV , which is considered to be sufficiently low with reference to the sample Fermi level. That is to say, the rest atom dangling bond states should be taken as occupied states and completely filled. The additional electron required to fill the state of the rest atom dangling bond can come partially from adatoms surrounding the rest atom.

4.7 Modeling of the band DOS

Now we consider the density of states for the bands at each atomic position $g_{i,\alpha}(\varepsilon)$ in equation (3-32). In order to calculate the tunneling current, the bands and their density-of-states have to be modeled. We start with a brief introduction of bands and DOS.

The electrons of an isolated atom occupy atomic orbitals, which have a discrete set of energy levels. When several atoms are brought together to form a molecule, their atomic orbitals split into a number of molecular orbitals proportional to the number of atoms. If a very large number of atoms is brought together into a solid, the number of orbitals becomes very large, and the difference in energy between them becomes very small. However, some intervals of energy contain no orbitals, no matter how many atoms are aggregated. Any solid has a large number of bands. In theory, it can be said to have infinitely many bands (just as an atom has infinitely many energy levels). However, all but a few lie at energies so high that they correspond to electron

states escaping from the solid. These bands are usually disregarded. Bands have different widths, based upon the properties of the atomic orbitals from which they arise. While the density of energy states in a band is very large, it is not uniform. It approaches zero at the band boundaries, and is generally greatest near the middle of a band.

In this section, our study focuses on the (111) surface of silicon. The termination of the silicon with a surface leads to a change of the electron band structure from the bulk material to the vacuum. Because of the sharp transition, surface states are formed and are localised only at the atomic layers closest to the surface. In the DAS model, surface states include states for the adatom dangling bonds, rest atom dangling bonds, and adatom backbonds since adatoms and rest atoms are localized on the topmost atomic layer and exposed to vacuum. The overlap of a large number of adatom or rest atom dangling bond orbitals and adatom backbond orbitals leads to corresponding surface bands. It is necessary to model the band structure and the density of states for all the surface states. Fortunately the most important surface bands are narrow and can be thought of simply as somewhat broadened atomic levels. Finally, the bulk bands that exist in the silicon bulk will be discussed.

4.7.1 Local density of states for surface bands

(a) Band for adatom dangling bonds

The Joannopoulos group described the behaviour of the local density of states (LDOS) of the features as a function of energy near the Fermi-level for adatom and rest atom sites [28]. The results are illustrated in Figure 4-12.

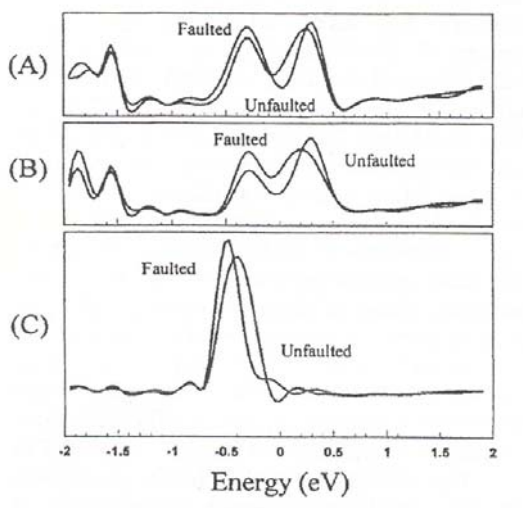


Figure 4-12: Differences in density of states between faulted and unfaulted states at different dangling bonds [28]: (A) corner adatoms db; (B) center adatoms db; (C) rest atoms db. In their further work, the peak in plot (C) is positioned at -0.8 eV [38].

The analysis of the curves A and B would focus on the sections near the Fermi-level (0 eV) (e.g. -0.5 to +0.5) in which adatom dangling bond states are located. Curve fitting can be a proper way to extract the DOS function of energy from the curves. Determination of the scale factor for the vertical axis can be obtained if we consider the physical meaning of the areas under the curve. After successfully reproducing the curves, we perform curve fitting operations on them using either two overlapping Gaussian functions (adatoms) or a single Gaussian function (rest atom). The results of our modeling of the density of states are shown below.

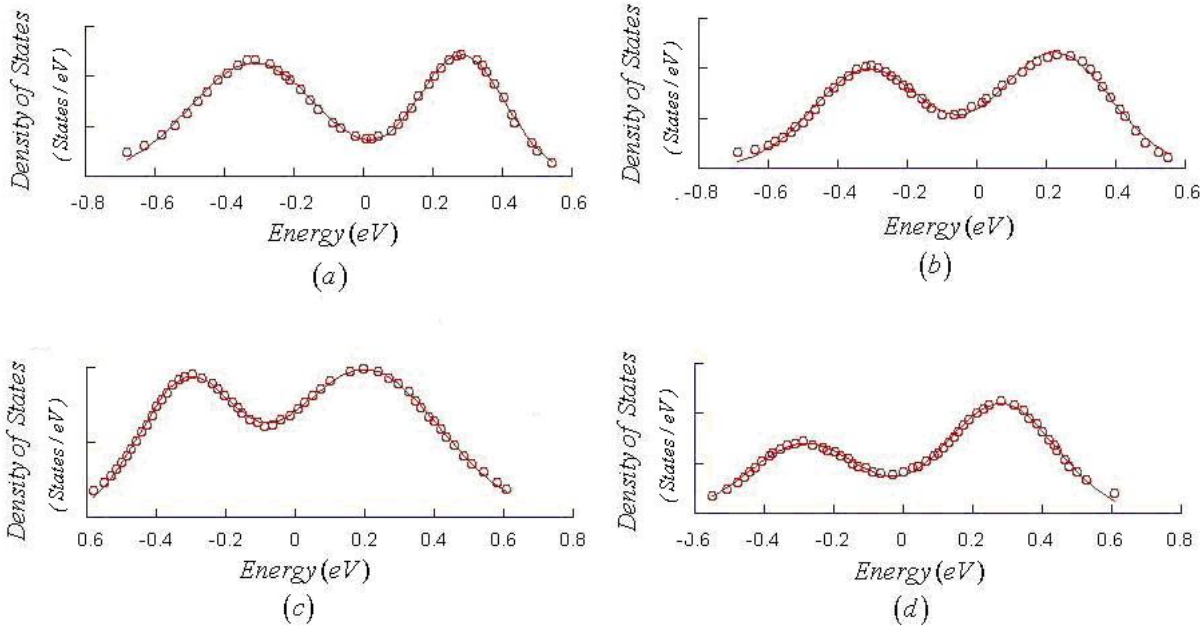


Figure 4-13: Modeling of density of states for dangling bond orbitals on (a) corner adatom in faulted halves (b) corner adatom in the unfaulted halves (c) middle adatom in faulted halves (d) middle adatom in the unfaulted halves.

The symbols in the graphics are data points extracted from the Joannopoulos group's result and the smooth curve is generated by curve fitting. In all pictures, the curve fitting looks very successful. The successful curve fitting in Figure 4-13 presented above proves that two overlapping Gaussian functions can be used to describe the density of states for adatom dangling bond orbitals, and a single Gaussian function for a rest atom dangling bond state. The normalized Gaussian function takes the form of:

$$h(\varepsilon,\bar{\varepsilon},\sigma)=\frac{1}{\sqrt{2\pi\sigma}}\exp\left(-\frac{(\varepsilon-\bar{\varepsilon})^2}{2\sigma^2}\right)$$

where $\bar{\varepsilon}$ is defined as the energy, and σ as the width.

Here we introduce some indices that will be quoted later: "F" means the faulted halves of a unit cell of the DAS model, while "U" denotes the unfaulted halves. "o" represents the

"occupied states" (below Fermi level referred as to zero in energy level) and "u" indicates the "unoccupied states" (above Fermi level). "c" denotes corner adatom and "m" represents the middle adatom, and w as the weight. The density of states for an adatom dangling bond in the faulted half $g_c^F(\varepsilon)$ is the addition of two Gaussian functions:

$$g_c^F(\varepsilon) = w_{c,o}^F h\left(\varepsilon, \bar{\varepsilon}_{c,o}, \sigma_{c,o}^F\right) + w_{c,u}^F h(\varepsilon, \bar{\varepsilon}_{c,u}, \sigma_{c,u}^F)$$

The symbols of $w_{c,o}^F$ and $w_{c,u}^F$ are the weights of the two normalized Gaussian functions. The symbols $\bar{\varepsilon}_{c,o}$ and $\bar{\varepsilon}_{c,u}$ are the energy states corresponding to the two peaks of $g_c^F(\varepsilon)$ and their localization relative to the sample Fermi level is : $\bar{\varepsilon}_{c,o} < \varepsilon_F$ and $\bar{\varepsilon}_{c,u} > \varepsilon_F$. The number of electrons n in the occupied states and the number of holes p in the unoccupied states of the adatom dangling bonds in the faulted half are given by

$$n_c^F = \int_{-\infty}^{\varepsilon_F} g_c^F(\varepsilon) d\varepsilon \qquad (4\text{-}22)$$

$$p_c^F = \int_{\varepsilon_F}^{\infty} g_c^F(\varepsilon) d\varepsilon \qquad (4\text{-}23)$$

where ε_F is the Fermi level of the sample.

Similarly, the other three sets of parameters correspond to corner adatoms in the unfaulted halves, middle adatoms in the faulted halves and middle adatoms in the unfaulted halves respectively. Following the indexes, we can denote them by

$$n_c^U,\ p_c^U,\ g_c^U,\ w_{c,o}^U,\ w_{c,u}^U$$

$$n_m^F,\ p_m^F,\ g_m^F,\ w_{m,o}^F,\ w_{m,u}^F$$

$$n_m^U,\ p_m^U,\ g_m^U,\ w_{m,o}^U,\ w_{m,u}^U$$

The energies $\overline{\varepsilon}$, widths σ and weights w' of Gaussian functions are obtained from the curve fitting. All the results are listed in the following table.

Table 4-2: Modeling results of the band and DOS for adatom and rest atom dangling bond. The data in parenthesis is from a further work by the same group [38].

Atom	*c*				*m*				Rest atom	
F / *U*	*F*		*U*		*F*		*U*		*F*	*U*
States	*o*	*u*	*o*	*u*	*o*	*u*	*o*	*u*	*o*	*o*
$\overline{\varepsilon}$ (eV)	-0.31	0.28	-0.31	0.22	-0.31	0.20	-0.28	0.28	(-0.8)	(-0.8)
σ (eV)	0.19	0.13	0.16	0.16	0.14	0.22	0.16	0.16	0.14	0.18
w'	1.058	0.764	0.8	0.934	0.606	1.068	0.564	0.872	1.726	2.12
w	1.1	0.79	0.83	0.97	0.63	1.11	0.587	0.907	1.795	2.205
n (o) or p (u)	1.06	0.83	0.89	0.92	0.82	0.92	0.60	0.92	1.795	2.205

The weights w are obtained by rescaling the weights w' to satisfy the condition that the total number of electrons in the rest atom dangling bonds is 4.0. Since we have acquired the functions for the density of states, the number of electrons in the occupied states or the number of the holes in the unoccupied states can be obtained by adding equation (4-22) and (4-23).

(b) Bands of rest atom dangling bonds

A single Gaussian function describes the LDOS for a rest atom very well. According to the results of curve fitting, its peak is positioned around -0.5~--0.4 (e V), which does not agree with -0.8 (e V) that was obtained through experiments [2,16]. In a later paper [38] by Joannopoulos' group, the peak is properly positioned at -0.8 eV. In this work, we use the experimental result, which is believed to be more convincing. The peak width is extracted from Reference 38 since no direct experimental data is available.

(c) Band of adatom back-bonds

For simplicity, a Gaussian function is again used to describe the density of states for the bonding or anti-bonding bands between adatoms and their three neighbour atoms. It is believed that the adatom backbonds states are located at 2.0 eV below the Fermi level and at 1.5 eV above the Fermi level [16,37]. The width is not available and is roughly estimated to be 0.5 eV.

4.7.2 Density of states for the bulk bands

Calculating the density of states for sigma bonds (bonding or anti bonding) that 'weave' the silicon surface is a complicated problem. Density functional theory and tight binding are possible approaches to this problem. However, they would require a significant amount of computation. An easy theoretical approach would assume that surface sigma bonds behave like those sigma bonds in the bulk. The problem is now reduced to finding the density of silicon bulk states. Silicon bulk's DOS has been obtained and collected in a related hand book [39]. The function describing this DOS is not simple and can not be easily obtained by curve fitting procedures. In this work, we can use a numerical look-up table to solve this problem.

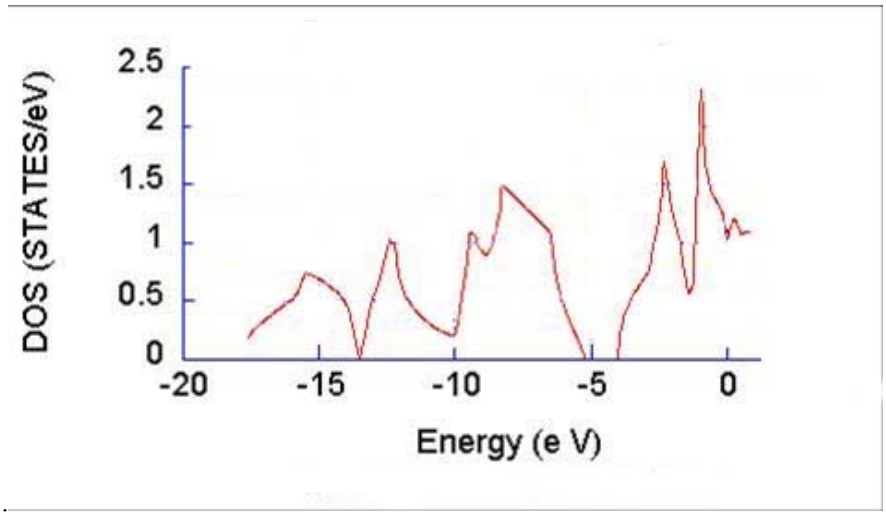


Figure 4-14: Modeling of the density of silicon bulk states [39].

In Figure 4-14, the zero energy is the vacuum level. The Fermi level for silicon is at about

-4.5 eV, as can be seen from the separation of the valence band and the conductance band. The physical meaning of the area under the curve is the number of states available. In our case, the occupied number of states is eight because of the eight valence electrons arising from two silicon atoms in the primitive unit cell for the bulk diamond structure for silicon.

4.8 Summary

According to the DAS model of the Si (111)-(7x7) surface, the ideal position of each atom of the unit cell can be determined. The adatoms and rest atoms have been taken as examples for the purpose of illustration. The ideal positions of the adatoms and rest atoms were corrected according to the Joannopoulos group, which has successfully calculated the average relaxation of each atomic layer relative to the ideal level.

A simple atomic orbital model has been used to calculate the surface electron wave functions. In this model, adatom and rest atom dangling bond orbitals are hybrid orbitals of silicon basic atomic orbitals assumed to be hydrogen-like wave functions. The sigma bonds between two silicon atoms are described using molecular orbital theory and valence bond theory. The effect of electrical field on the wave functions has been considered by assuming a triangular potential barrier between the STM tip and the sample surface.

Due to the different features observed experimentally, which were explained by charge transfer, the adatoms were divided into four categories according to their geometrical environment in the unit: corner adatoms, and middle adatoms in the faulted or unfaulted halves. The adatom dangling bands were successfully modeled by two overlapping Gaussian functions

according to the theoretical calculations by the Joannopoulos group. The bands are very narrow and localized close to the Fermi level, which explains the saturation of adatoms in a STM topographic image. This tight binding approximation made in chapter 3 works very well for the semiconductor surfaces since the bands corresponding to the uppermost atoms dominate the STM image. This is because these bands are narrow and close to the Fermi level. Though the tight binding approximation may be poor for atoms deeper into the surface, these atoms and their corresponding wider bands make a much smaller contribution to the STM image.

All the details concerning the atomic and electronic structure presented in this chapter have been incorporated into a FORTRAN program. Unlike sophisticated first-principles calculations, our empirical approach does not require enormous computer power, yet still allows for accurate simulation of a very complicated surface. This is the first simulation we know of that simulates STM images of the silicon (111)-(7x7) surface and includes the effect of the electrical field from the STM tip. Furthermore, this empirical approach is very flexible for adding extra atoms such as adsorbates. The next chapter will present our calculated images and compares them to experimental results.

Chapter 5
Results of the modeling

5.1 Overview

Computer code, written in FORTRAN, was used to calculate STM images of the Si (111)-(7x7) surface. The calculated topographic images are compared to experimental images both quantitatively and qualitatively. The uneven intensities among adatoms in the images obtained at negative sample bias, which have been experimentally observed, are well simulated. Quantitatively, the tip corrugations along the longest diagonal of the unit cell are calculated and compared to the experimental measurements.

Spectroscopic images are also simulated and compared to the experimental images. The simulated images are compared to the experimental CITS images. Even though the images are obtained in different techniques, the essential spectroscopic features are supposed to be the same. Believed to be a tool of giving the electronic structure of bands, the normalized conductance curves over five unique atomic positions are generated. Due to the fact that most of the experimental normalized conductance curves are not measured and recorded quantitatively, we have compared most of the simulations only qualitatively to experiment. The only quantitative comparison between our results and the result from Boland's results are quite successful.

Finally, the adsorption of one hydrogen atom onto the Si (111)-(7x7) surface is simulated. Specifically, one hydrogen atom is preferably adsorbed directly above the adatom site or rest atom site. The images and curves corresponding to the hydrogen adsorbed surface are presented and compared to the experimental measurements.

5.2 The tunneling current set-point

In order to describe the tip-sample separation, we introduce a term called tip height set-point. Tip height set-point is defined as the distance between the tip and the actual location of the unfaulted corner adatom in the vertical direction, when the tip is directly above that adatom with the set-point current flowing. The actual location of the unfaulted corner adatom is 4.39 Å higher than the CH atom 26 as shown in figure 4-1. Because of the complexity of realistic tip-sample systems, it is impossible to calculate the tunneling current exactly. Even in the highly simplified tunneling current expression (3-33) given by Tersoff and Hamann, the task of calculating the tunnel current is still difficult due to lack of knowledge of the density of states of tip and tip geometry. Therefore, related theoretical works have rarely been reported [40].

We use the free electron gas theory to model the density of states of the tip to determine the pre-factor of equation (3-33) of section 3.6. We have found from comparisons to experimental data that this value for the pre-factor is quite a bit too large. It is known that at a typical tunneling current near 1 nA and bias voltages around 1 volt, the STM tip is about 5 to 7 Å above the adatom [10]. With these parameters we find currents of about 10 μA in our calculations. There are likely several reasons for this discrepancy. For example, the tip density of states discussed in section 3.6, may be overestimated. In particular the s-wave model for the tip may be suspect [41]. A bigger problem may arise from ignoring the effects of electronic screening on the hydrogen-like wave functions we have used in our model in equations (4-8) to (4-11). Screening would have the effect of decreasing the range of the coulomb potential and should cause the wave functions (4-8) to (4-11) to tail off more rapidly. Based on these considerations we adjust the tunneling current to ensure that the tip is about 5.0-7.0 Å away from the adatoms, which is the typical distance between the tip and the sample in real systems [10]. This deviation, although huge, is not that unexpected since many assumptions have been involved

in our calculation. Therefore, the precise magnitude of the tunneling current won't be a big concern to us as long as the tip-sample separation is simulated to be in the typical range.

5.3 Topography and comparison to experiment

5.3.1 Brief description of the image calculation

The calculation of the images in the constant current mode is briefly described in this paragraph. The image resolution and sample surface area have a significant impact on the running time of the image calculation program. We wanted to calculate images with high resolution without sacrificing running time of the program. In this work, we create 128x128 images of a Si (111)-(7x7) surface whose area is 40Å x 40Å. This resolution requires 16384 (128x128) pixels (measurement points). At each pixel, the set-point tip height is searched for in a range of (5Å~15Å) until it corresponds to the desired tunneling current. The tip height is then stored to be one element in a 128x128 matrix. The element's position in the matrix is determined by the position of the pixel in the image. After all the measurements are finished, the matrix can be output onto a text file, which is later opened by ImageJ to produce the simulated image.

5.3.2 Qualitative comparison

In our simulation, the sample bias is chosen to be -2.0 V and 2.0 V. We obtain two simulated images and compare them to their experimental images. The comparison is shown

in Figure 5-1.

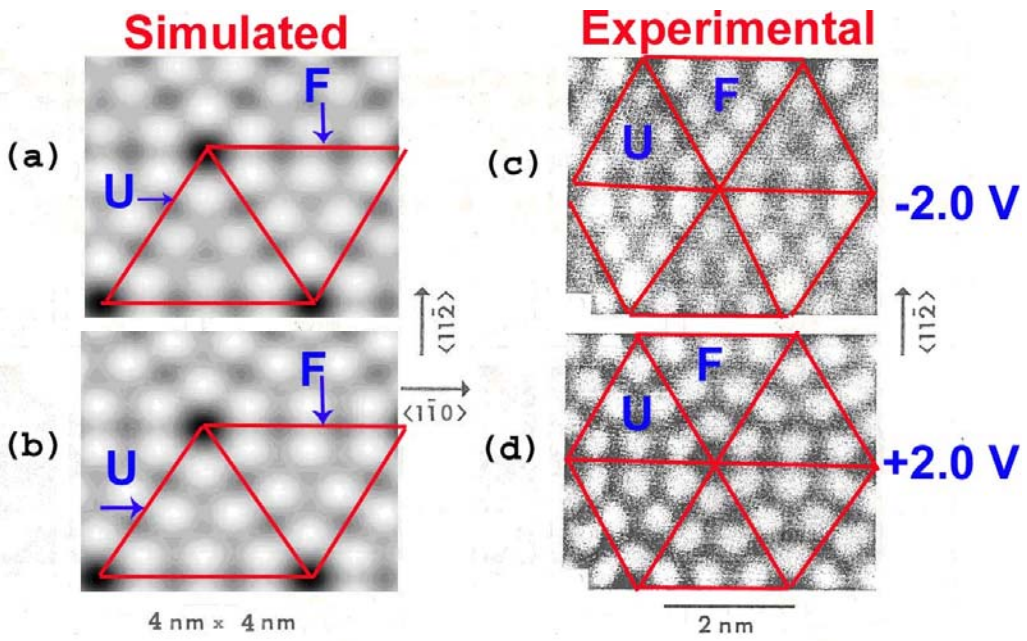


Figure 5-1: STM topographic images of the Si (111)-(7x 7) surface: simulated images acquired at sample bias of (a) -2.0 volt (b) +2.0 volt and experimental images [18] acquired at sample bias (c) -2.0 volt (d) +2.0 volt.

In Figure 5-1, according to the DAS model, the bright dots that saturate the image correspond to adatoms. The dark region surrounded by six bright dots is in accordance with a CH. The explanation for the adatom's dominance is that the states of the band that stems from the adatom dangling bonds are very close to the silicon Fermi level. Figure 5-1 (a) shows a simulated image obtained at a negative sample bias. This means that the electrons tunnel from the occupied states of the sample to the unoccupied states of the tip. Therefore, the features resolved in Figure 5-1 (a) are mainly related to the occupied states of the adatom dangling bonds band. We compare it to the experimental image in Figure 5-1 (c) and conclude that the following features are successfully simulated: (1) adatoms in the faulted sub-cell appear brighter overall than those in the unfaulted sub-cell. (2) The six adatoms in one sub-cell are imaged with unequal

intensity (at negative sample bias).The three corner adatoms (i.e., the ones closer to the CHs) are more intense than the middle adatoms (The reason for separating adatoms into two groups: corner adatoms and middle adatoms, is now apparent).

In Figure 5-1 (b), electrons tunnel from the occupied states of the tip into the unoccupied states of the sample since the sample is positively biased. Therefore, the features resolved are related to the unoccupied states of the adatom dangling bond band. Our simulation result again well match its experimental counterpart in Figure 5-1 (d) and both results show that the adatoms in an image obtained at positive bias have almost the equal intensity. Thus, we are confident to conclude that our simulation has been very successful in reproducing the observed features of the experimental STM images qualitatively.

5.3.3 Quantitative comparison

The next step is to quantitatively compare simulations to experimental STM topographs. We first define the tip height set-point (THSP) as the vertical height of the tip above the UC adatom when the tunnel junction is at the prescribed tunnel current set-point value. It is understood that at all pixels of these STM topographs, the tunnel current is constant while the AH (or TH) varies. The THSP is generally used as a fitting parameter here. An excellent graphical method for comparing topographs is to plot the TH (or AH) data over a scan line spanning a long diagonal of the 7x7 unit cell from one CH to another i.e. from points A to B in Figure 5-2 (top-view) and Figure 5-3 (side-view).

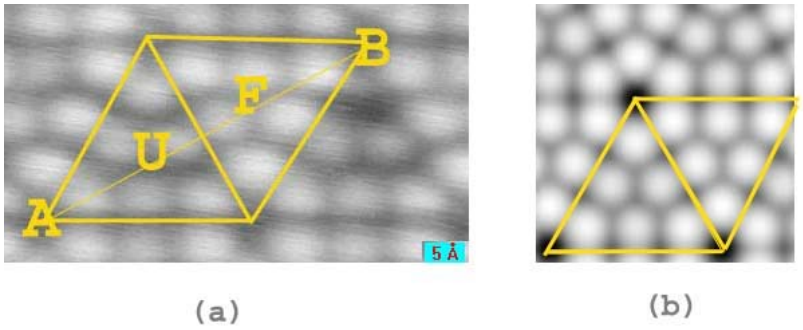


Figure 5-2: STM images of the clean Si (111)-(7x7) surface, obtained at -0.7 V : (a) experimental image obtained by S. Horn (b) simulated image

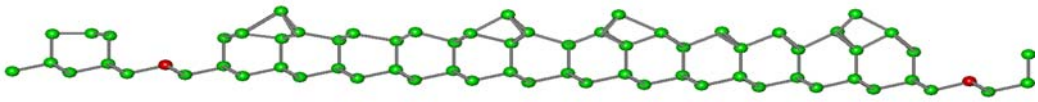


Figure 5-3: The side view of crystal structure along the longest diagonal of one unit cell.

This scan line includes one of each of the four distinct types of adatoms, in the following order from A to B: UC, UM, FM, FC. Plot (b) in Figure 5-4 shows the scan line for the unrelaxed atomic positions. We see four pronounced peaks corresponding to the four adatoms, as well as deep minima at the CH positions. We note that the CHs are physically located about~4.39 Å lower than the adatoms are while the image contrast from adatom to CH is more like 3.0 Å. The tip never does get deep enough into the corner hole to pick up signal from the CH atom. Instead what we detect are the "shoulders" of the six adatom db orbitals surrounding the CH. The scan line also shows two small peaks labelled SU and SF. Each of these points corresponds to a saddle point between two middle adatoms, as one might expect from visual inspection of Fig 5-2.

Scan line (b) of Figure 5-4 shows the effect of the atomic relaxations discussed in Section 4.3.2. All the adatoms move upwards by about 0.5 Å with the FC adatom moving upwards more than the UC adatom by 0.04 Å, as indicated in Table 4-1. The relaxations have the effect of somewhat increasing the adatom topographic contrast. In the end the contrast is due partly to electronic structure and partly to atomic structure. From Table 4-1 we see that the CH relaxation is negligible, yet the relaxed scan line shows substantially more contrast, even dipping down to a slightly smaller AH value. The most likely explanation comes not from vertical relaxations but from lateral relaxations of the six adatoms surrounding the CH. Specifically, the UC shifts towards the CH by 0.13 Å while the FC shifts away by 0.018 Å. The net result is a decrease of adatom db charge density above the CH, so the tip has to move in closer to obtain the set-point current.

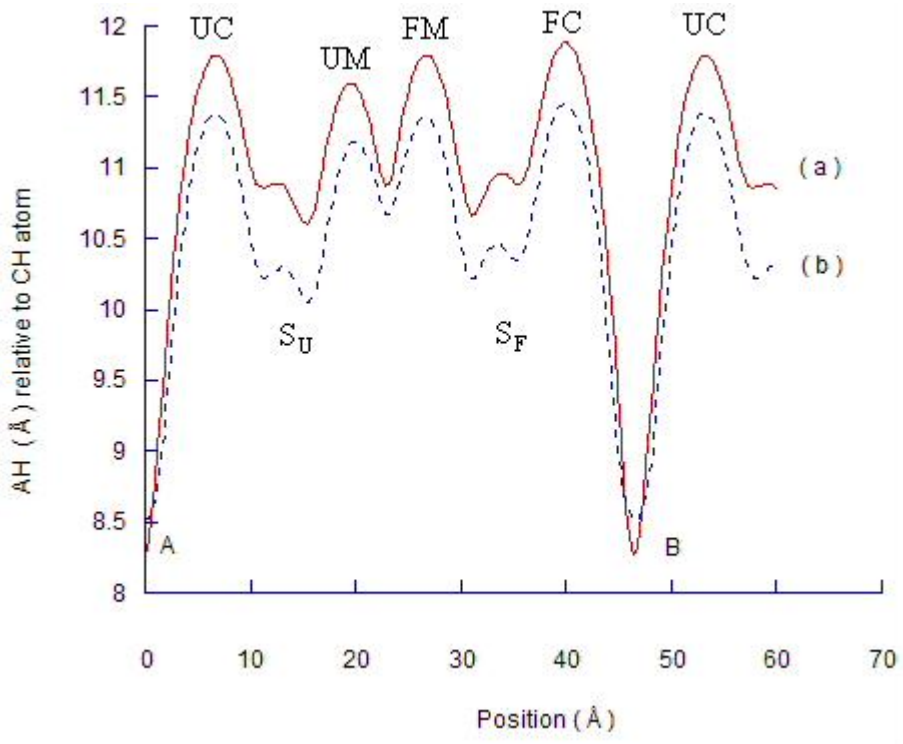


Figure 5-4: Comparison between two scan lines based on (a) relaxed atomic locations (solid line), (b) ideal atomic locations (dashed line). The sample bias is set to be -0.7 V.

In Figure 5.5 we show the effect of varying the THSP value. In experiments the STM operator can most easily adjust the THSP by changing the set-point current value. Setting a higher set-point current value effectively "pushes" the tip closer to the surface i.e. to a smaller THSP, usually with the goal in mind of increasing image resolution and/or contrast. Our simulations show that this is indeed the case: moving the tip by ~3 Å (compare a to c) roughly changes the contrast by a factor of two.

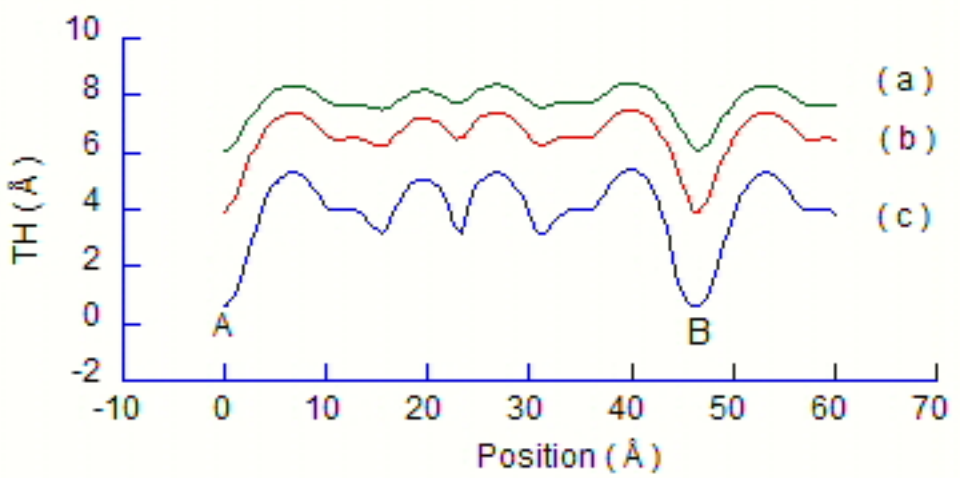


Figure 5-5: Scan lines with different THSP (Å) of (a) 8.36 (b) 7.39 (c) 5.24. The sample bias is -0.7 V.

We now make a direct comparison of our simulations to experimental data taken by Steven Horn with the University of Lethbridge STM system. Since measurements of TH or AH are not available experimentally we introduce subtracted height (SH) as the actual TH at a position minus the average TH over the scan line. This always ensures that the mean SH value is zero. In Figure 5-6, the two scan lines extracted from Figure 5-2 (a) and (b) and SH means subtracted height relative to the average height. The four peaks in each curve correspond to the four adatoms on the top layer shown in Figure 5-3. This simulated scan line was obtained from the least squares fit by adjusting the THSP to about 7.4 Å. The simulated scan line (solid line) is reasonably close to the experimental curve particularly in the range from 20 Å to around 50 Å

that corresponds to the unfaulted halves of a unit cell. The tip positions calculated at nearby the CHs are lower than the experiment results, which can be explained by the fact that the ideal "sharper" tip used in our calculation can detect more contrast and go deeper into the CH than the real tip. We calculate the deviation of our curve from the experimental data and obtain the rms deviation of 0.0385 Å . This good agreement makes us conclude that from a quantitative point of view, the simulation is very successful at a sample bias of -0.7 V. Two areas that could use improvement are the simulation of the US saddle point and also the adatom contrast is a little more pronounced in the experiment. At the moment we have not ruled out the distinct possibility that these small problems are due to experimental artefacts.

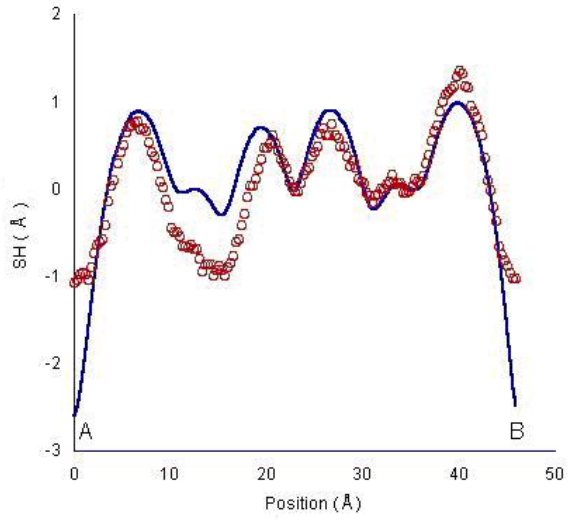


Figure 5-6: Plot of theoretical (solid line) and experimental (dot markers) STM depth vs. position across the long diagonal of the unit cell for occupied electronic states between A and B

The same method is used to examine the unoccupied images acquired at +1.5 V.

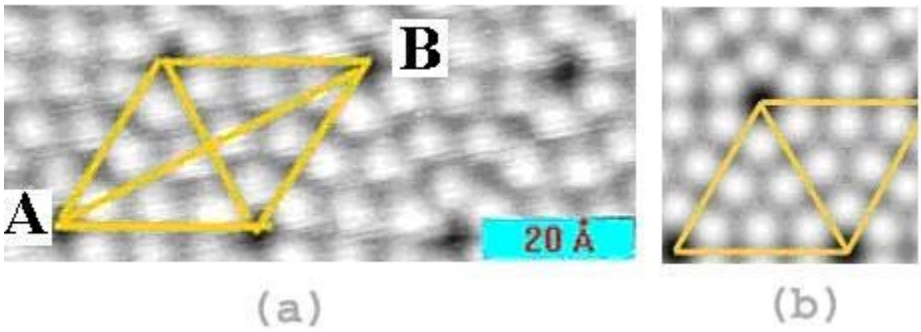


Figure 5-7: STM topographic image of the clean Si (111)-(7x7) surface, obtained at sample bias of +1.5 V. (a) experimental image, (b) simulated image.

In Figure 5-7, the 7x7 unit cell is outlined and the twelve adatoms are clearly visible and appear to be of even intensities. We take the scan line of the long diagonal from CH E to F and compare the experimental data to the calculated scan line. The comparison is shown in Figure 5-8.

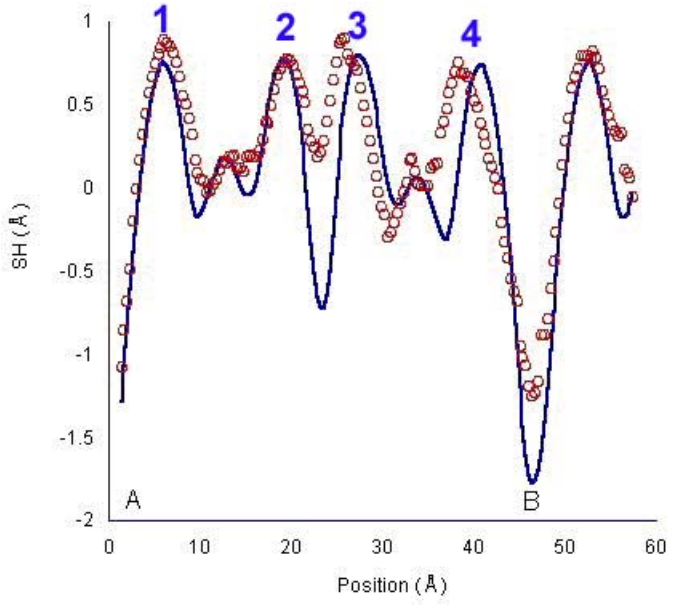


Figure 5-8: Plot of theoretical (solid line) and experimental (open circles) STM depth vs. position across the long diagonal of the unit cell for unoccupied electronic states between A and B. The four peaks that correspond to adatoms are numbered 1-4 from left to right.

We have shown in Figure 5-5 that the contrast in scan line is closely related to the THSP. We adjust the THSP to 4.0 Å so that the two scan lines can fit each other well. In Figure 5-8, the calculated tip positions over adatoms are almost the same, which agree with experimental observation of almost equal intensity among adatoms in the STM images at positive sample bias. The solid curve has matched the experimental results very well at the positions of the first and the second peaks, even the saddle points of middle adatoms. However, obvious shifts of experimental results toward the left happen at the third and the fourth peaks. These shifts are possibly because the effect of thermal drift on experimental images can not be compensated or corrected completely. In other words, it is really hard to achieve the perfect diamond-like unit cell even after the image compensation. Though more analysis must be done, at the moment, we believe that the simulation could not result in the positions of the faulted adatoms being so far off. The more likely explanation is an experimental artefact: perhaps an asymmetric tip shape or perhaps the s-orbital approximation is poor and an asymmetric p-orbital contributes strongly on the end of the tip. It appears that this simulation approach could be used as an effective tool to diagnose STM tips. Also, at the valley between the two middle adatoms (second and third peaks), a considerable deviation appears. The higher experimental tip position can be related to the decreased distance between the two middle adatoms caused by the shift. Compared to Figure 5-6, the deviation at the CH in this picture is smaller, which can possibly be explained by the improved tip condition such as a sharper tip. We calculate the deviation of our curve from the experimental data and obtain the rms deviation of around 0.0276 Å. All the evidence shows that the topography simulation is very successful even quantitatively. The next section will discuss the simulation on spectroscopy, in particular the $dI/dV * V/I$ spectrum obtained when the tip is placed above a specific atomic site.

5.4 Point spectroscopy

In Chapter 1, we introduce the quantity of dI/dV and explained briefly that this quantity is approximately proportional to the local density of states of tip position. At a specific spatial position, by varying the voltage, we can measure dI/dV spectrum as a function of energy. This spectrum is expected to reveal the features of the DOS function. However, the features of DOS can be possibly washed out when the voltage is considerably large, as can be seen from the energy dependence of the transmission coefficient in equation (1-2). This problem has been solved using a simple but effective solution proposed by Stroscio, Feenstra, and coworkers [3]. In their proposal, dI/dV is normalized by dividing it by I/V. This actually yields $d\ln I / d\ln V$, and so effectively cancels out the exponential dependence of $T(E,V)$ on V. See Appendix G for more detail.

In this section, the calculated $dI/dV * V/I$ spectra as a function of energy are presented and compared to the experimental results both qualitatively and quantitatively. We adopt the idea of the CITS technique that was used to obtain experimental spectra. The tip position over a specific feature site is first calculated when the "feedback" is activated. By holding the tip position fixed (the "feedback" has been deactivated), we can calculate the corresponding magnitude of $dI/dV * V/I$ given a value of voltage and thus plot $dI/dV * V/I$ as a function of voltage.

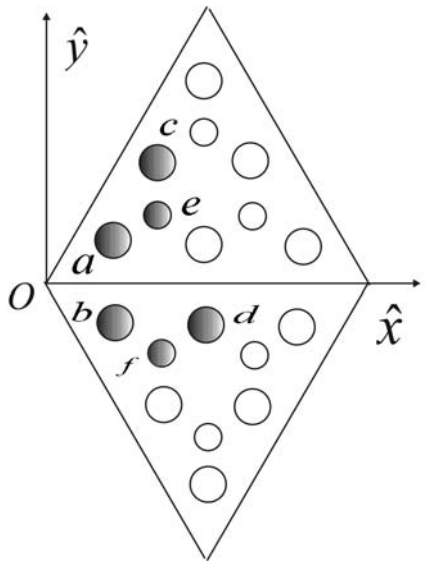
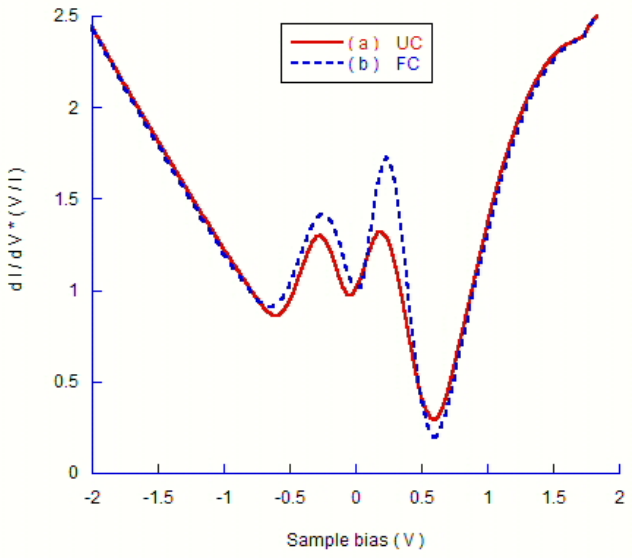
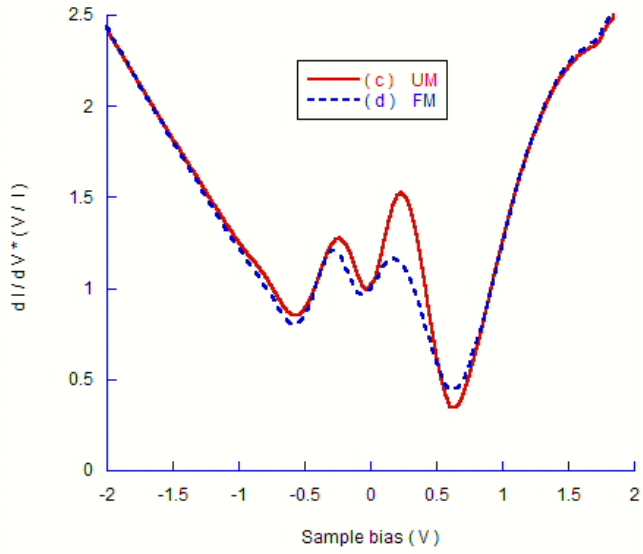


Figure 5-9: Top view of the DAS model for the Si (111)-(7x7) surface. One unit is sketched with six unique sites labelled by filled dots. The bigger dots represent adatoms while the smaller dots denote rest atoms.

In Figure 5-9, the shaded dots denote the feature sites over which the $dI/dV * V/I$ spectra are calculated. The six unique sites are : a (UC), b (FC), c (UM), d(FM), e (UR) and f (FR). Six spectra over the six unique sites are calculated and plotted in Figure 5-10.


(a) UC
(b) FC
dI/dV*(V/I)
Sample bias (V)


(c) UM
(d) FM
dI/dV*(V/I)
Sample bias (V)

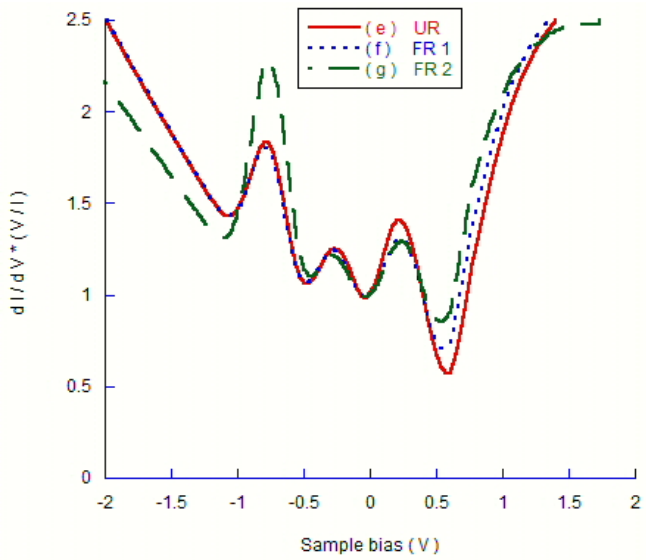


Figure 5-10: The $dI/dV*V/I$ spectra over (a) UC (b) FC (c) UM (d) FM (e) UR (f) FR 1 are calculated. The THSP is 8.46 Å. The spectrum over (g) FR 2 are calculated after we adjust the THSP to about 7.0 Å. All the spectra are calculated at the bias set point of -0.8 V.

In Figure 5-10, the plots (a)-(d) show the $dI/dV*V/I$ spectra over adatoms. Two peaks are observed between -0.5 V to +0.5 V in each spectrum. In this range of energy, the curves look similar to the DOS presented in Figure 4-13. The peaks are related to the two maxima of the two overlapping Gaussian functions, which have been used to describe the DOS of adatom dangling bonds. The reason for coming to this statement is given as follows. Once electrons start to flow through channels, increasing bias does not turn the channel off. On the other hand, the probability of tunneling is greatest for electrons near the Fermi surface; so, as the Fermi surface passes the appropriate channel, $dI/dV*V/I$ is maximized. At zero voltage, all the spectra pass through the point of (0, 1) in the two-dimensional plane as can be seen from equation (G-4) in Appendix D where the quantity of $dI/dV*V/I$ approaches to unity if the voltage goes to zero. This is a fairly general result, the only assumption that there be at least some DOS at the Fermi level.

Plots (e) to (g) show the $dI/dV*V/I$ spectra over rest atoms. The peak of greatest magnitude happens at a sample bias of -0.8 V, which corresponds to the energy level at which the DOS of rest atom dangling bonds is maximized. This confirms that the $dI/dV*V/I$ is approximately proportional to surface density of states. We can also observe two relatively small peaks in the range of -0.5 V to +0.5 V. After comparison to the spectra in plots (a) and (b), it is clear that the small peaks are the features of the adatoms. This appearance of an adatom-like feature in the spectrum of rest atom can be possibly explained by the fact that each rest atom is surrounded by three adatoms and the distance between rest atom and adatom is relatively short, about 4.5 Å. The tail of wave functions of adatoms can be possibly picked up by the tip that is held above the rest atom.

Our result is clear: the DOS peaks for the adatom and rest atom bands (see Figure 4-13) are plainly visible in the simulations of normalized conductance spectra. Measurements of normalized conductance spectra are commonly interpreted in terms of LDOS. In fact, if the barrier tunneling coefficient is assumed to have no dependence on applied bias or electron energy, then one can show that $dI/dV*V/I \propto LDOS$ exactly (see Appendix G). Realistically the varying tunnel barrier complicates the interpretation of normalized conductance curves. It is generally agreed that the complications introduced by the tunnel barrier variations only add a rather smoothly varying background to the spectrum i.e. the identification of peaks in normalized conductance spectra as LDOS features is acceptable. One can see from Figure 5-10 though that quantitative interpretation of peaks can be problematic: the rest atom peak at -0.8 eV sits on top of a fairly steeply rising background curve.

The next step is to compare our calculated spectra with experimental results obtained by R. Wolkow and Ph. Avouris [16] and Boland [25]. R. Wolkow and Ph. Avouris measured the

$dI/dV*V/I$ spectrum experimentally but unfortunately they did not provide the scale for y axis of the graphics or in other words it is impossible to extract the exact quantity of $dI/dV*V/I$ from their results. Therefore, $dI/dV*V/I$ spectrum can only be roughly estimated having it in mind that the spectrum has to pass through the point of (0,1). Boland's result only shows dI/dV spectrum as a function of sample bias, so it is necessary to take one more step from this spectrum to obtain the $dI/dV*V/I$ spectrum. The new spectrum derived from Boland's result is presented in Figure 5-11. The comparison between our calculated spectrum and the other two experimental results is shown in Figure 5-12.

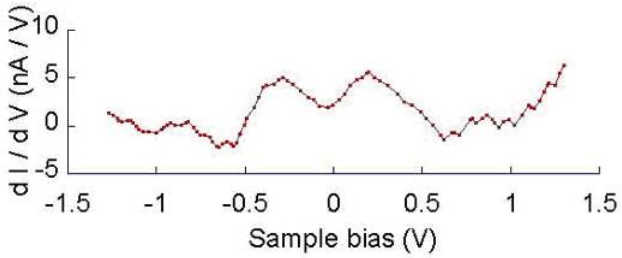


Figure 5-11: $dI/dV*V/I$ spectrum as a function of sample bias based on Boland's experimental measurements. The sample bias is +1.0 V, when feedback is active.

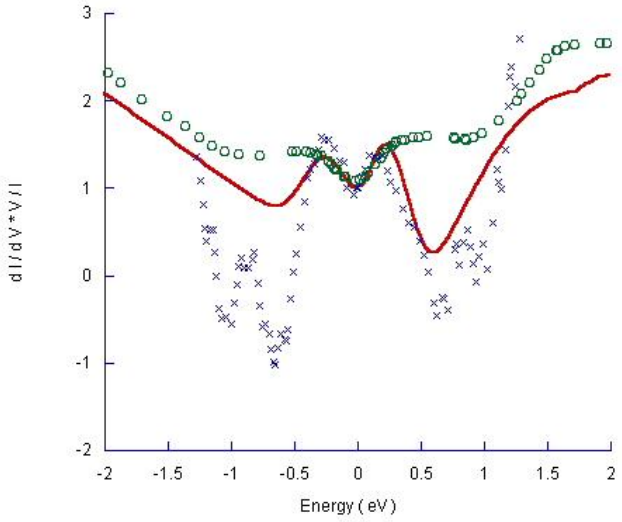


Figure 5-12: The comparison between our calculated spectrum (solid line) over adatoms and two experimental results by Wolkow and Avouris (circles) and by Boland (crosses) at sample bias of +1.0 V

In Figure 5-12, three $dI/dV*V/I$ spectra over adatom sites are presented. Our calculated spectrum is obtained by averaging the two spectra over UC and UM. At the vicinity of zero energy , the three spectra match each other very well. In Boland's spectrum, a feature (small peak) is observed at around -0.8 eV , which is not observed by Wolkow and Avouris and is also not present in our calculated results. This is perhaps because the tip Boland used was also detecting a rest atom during his measurement over an adatom. This is similar to our observation of the adatom's feature in our calculated spectrum over rest atoms. In the range of -1.0 V to +1.0 V, our calculated spectrum lies between the two experimental results. We can conclude that theoretical simulations yield a reasonably good agreement with the experimental results.

5.5 Image Spectroscopy

At a specific energy level, we can measure $dI/dV*V/I$ as a function of lateral positions by scanning the tip across the sample. Since the quantity of $dI/dV*V/I$ is approximately proportional to the DOS, the $dI/dV*V/I$ image spectroscopy would map the lateral distribution of the DOS of the sample at the specific energy eV with respect to the Fermi level. Therefore, using image spectroscopy, one can obtain a real-space imaging of surface electronic states [2]. The goal is to provide a clearer picture than obtained from a standard topograph which is a mixture of geometric and electronic structure information. This undesirable though interesting mixture often complicates the interpretation of the observed features.

In this section, our calculated $dI/dV*V/I$ images are compared to the CITS spectroscopic images obtained experimentally by R. J. Hamers [2] and the dI/dV image obtained by Steven Horn at sample bias of -0.8 V using our home-made STM in our laboratory. Even though the experimental images are obtained by a different technique, the comparison is

still reasonable because the basic principle of all kinds of STM spectroscopic techniques is to reveal the surface density of states in a real space. Our $dI/dV*V/I$ images are calculated using the method of Binnig, who first applied a high frequency (chosen above the response of the feedback system) modulation voltage to the constant dc bias between the sample and the tip [14]. We obtained the values of $dI/dV*V/I$ at the measurement points during the "raster scan" while the feedback loop was active. In the CITS technique, the measurements were taken after brief interruptions of the raster scan while the feedback loop was off. The differences between the two methods were discussed in the section 2.1.

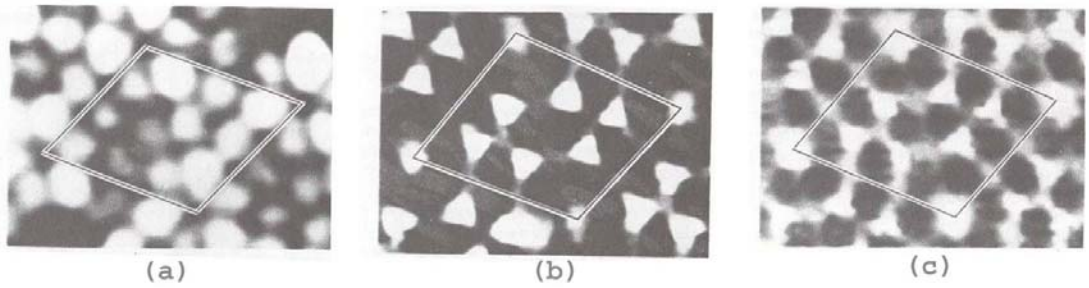


Figure 5-13: Current-imaging-tunneling spectroscopy (CITS) images [2] obtained at (a) -0.35 V (b) -0.8 V (c) -1.7 V

Figure 5-13 shows the experimental images obtained by R. J. Hamers [2]. The experimental technique used is current-imaging-tunneling spectroscopy (CITS). In CITS images, regions of high current appear bright and regions of low current are dark. The image (-0.35 V) reflects the spatial distribution of the adatom dangling bonds states. The uneven intensity of current at adatom sites is observed among the adatom regions. At -0.8 V, six bright sites in one unit are surface states due to rest atom dangling bonds. A bright dot is observed in each corner hole. At -1.7 V, some backbond states are observed and the corner holes still appear to be bright.

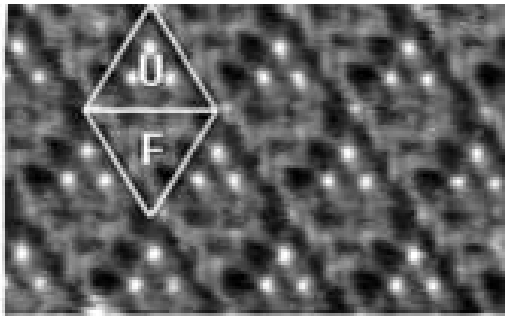


Figure 5-14: the dI/dV image obtained by Steven Horn at sample bias of -0.8 V using our home-made STM in our laboratory. Three bright dots are observed in the unfaulted half cells.

Figure 5-14 shows the similarity to Figure 5-13 (b): adatoms are "filtered" out effectively and rest atoms appear. The rest atoms in the faulted halves appear dimmer than those in the unfaulted halves, which is not observed in Figure 5-13 (b)

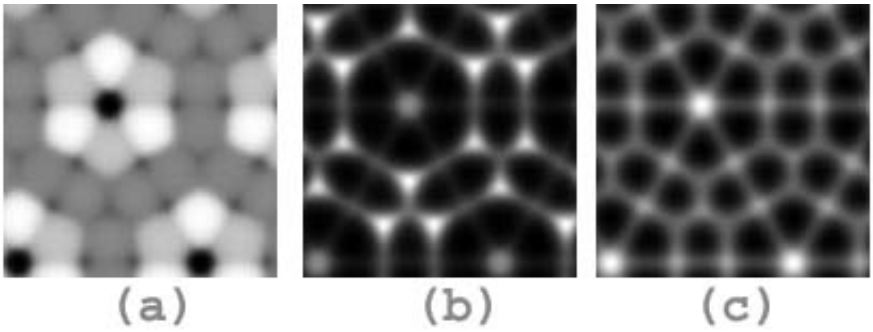


Figure 5-15: Simulated $dI/dV*V/I$ images of Si (111)-(7x 7) surface obtained at (a) -0.31 V (b) -0.8 V (c)-1.7 V. The images are displayed in the top view.

In Figure 5-15, we present our $dI/dV*V/I$ images calculated at sample bias of -0.31 V,

-0.8 V and -1.7 V. Having the higher DOS displayed as light color in greyscale image, the bright regions correspond to high DOS while the dark regions are referred to as low DOS. In Figure 5-15 (a), there are twelve regions of the uneven brightness in each unit cell. The regions

correspond to the adatom sites according to the DAS model. The uneven intensity agrees with Figure 5-13 (a), and corresponds to the uneven DOS of different adatom sites. More specifically, if we extract the four quantities of $dI/dV*V/I$ at sample bias of -0.31 V from the four adatom spectra in Figure 5-10, the four data would be slightly different.

In Figure 5-15 (b), a unit cell has six points of high DOS centered between the adatoms while the rest of the unit cell appears to be rather uniform. The six high DOS regions are the spatial locations of rest atoms on the sample surface according to the DAS model. This calculated image looks closer to Figure 5-13 (b) than Figure 5-14. More excitingly, the triangular shape of the high DOS (due to the adatom "mask" effect) appears almost the same. The reason for the disappearance of the adatom features and the appearance of the rest atom features can be found in Figure 5-10 where at sample bias of -0.8 V, the quantity of $dI/dV*V/I$ is around 1.2 in the adatom spectra , which is relatively small compared to 1.8 extracted from the rest atom spectrum. Note that the corner hole is a high DOS region.

In Figure 5-15 (c), high DOS surrounding the adatoms is observed. At -1.7 V, a silicon adatom backbond state is expected to dominate the images [37]. Geometrically, each adatom has three adatom backbonds. However, the striking feature in Figure 5-13 (c) of rest atom back bonds can not be observed in our calculated images. The possible reason is that in our model, the rest atom backbonds are treated exactly the same as those bonds inside the silicon bulk. The rest atom backbonds are likely more strained than we have considered in our model.

5.6 Modeling H/Si

It has been of great importance from technological and scientific points of view to understand the interactions of atoms and molecules with the Si (111)-(7x7) surfaces. It represents

one of the most challenging problems to theory, because of the complexity of the systems involved and the wide variety of possible outcomes.

The Si (111)-(7x7) reconstruction surface has a rich and complex chemistry because in a unit cell there are nineteen dangling bonds that define nineteen different reactive sites, each on a silicon atom. Seven of them are unique, with the other twelve related by symmetry. The interaction of different organic molecules with the Si (111)-(7x7) surface has been extensively studied experimentally and theoretically, i.e. ethylene [42-44], benzene [45-48], acetylene [44,49-55]. The simplest chemical probe of this surface is atomic hydrogen, and the simplest molecule that reacts with this surface is the hydrogen molecule [25,26]. Studies by STM have shown silicon adatoms become substantially dimmer (lower in topograph) after bonding with a hydrogen atom.

5.6.1 STM topographic, spectroscopic image of one hydrogen atom on Si (111)-(7x7) surface

The mechanism of hydrogen adsorption and desorption has been studied experimentally. It is believed that one hydrogen atom is most likely bonded to adatom sites or rest atom sites [25,27]. The adsorption process of one hydrogen atom onto the surface can be simply interpreted as the replacements of adatom or rest atom dangling bonds with Si-H sigma bonds.

The bonding molecular orbital corresponding to a Si-H sigma bond is believed to be far below the sample Fermi-level and the anti-bonding molecular orbital of a Si-H sigma bond is about 3.5 eV above the Fermi level [25]. Because of the transmission probability's exponential dependence on the potential barrier width, the states are of minor importance in contributing to the tunneling current except for bias voltages around +3.0 volts and higher. In this work, the adsorption of hydrogen is modeled simply by removing the silicon atom dangling bond from the

calculation. This provides us with an easy way to simulate the adsorption of hydrogen to Si (111)-(7x7) surface.

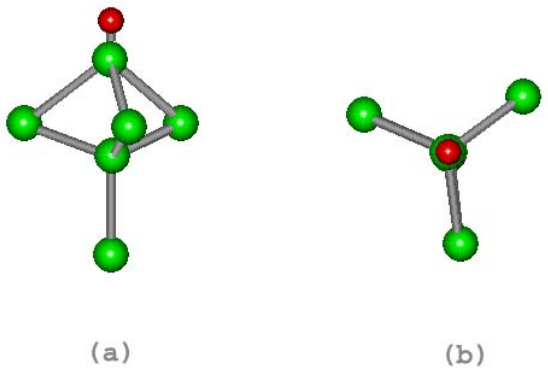


Figure 5-16: The optimized geometry corresponding to the chemisorption of one hydrogen atom onto an adatom (a) side view (b) top view [27]. Large (green) circles denote silicon atoms and hydrogen is represented by the smaller (red) circle.

We first discuss the adsorption of one hydrogen atom onto an adatom's site. The hydrogen atom sits directly above the adatom site [27], as shown in Figure 5-16. The adatom dangling bond is replaced by a Si-H sigma bond. In the modeling, we take a simulated topograph with one of the adatom dangling bonds disabled and compare it to the "intact" topography. The comparison is shown in Figure 5-17.

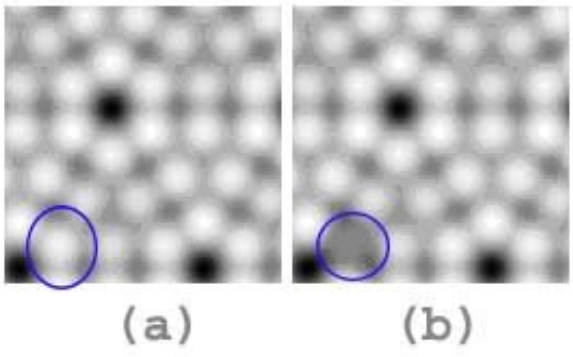


Figure 5-17: Simulated topographic STM image of (a) clean Si(111)-(7x7) surface (b) adsorption of one hydrogen atom onto the adatom site (denoted by circle) . Both images are obtained at sample bias of -2.0 V.

In Figure 5-17, picture (a) is saturated with the adatom dangling bonds as can be seen from the twelve dots in each unit cell. In picture (b), one adatom site (at the left bottom corner) appears to be dark because its dangling bond is disabled after combining with a hydrogen atom into a strong Si-H bond. Similar to our treatment of clean Si(111)-(7x7) surfaces, we take a scan line over the longest diagonal of one unit cell in both of the pictures and plot the tip height as a function of position across the long diagonal of the unit cell. All scan lines in this study are chosen to be across the long diagonal of the unit cell so that they pass through atomic sites in the order of "UC", "UR", "UM", "FM", "FR" and "FC" that are under spectroscopic investigation as shown in Figure 5-9.

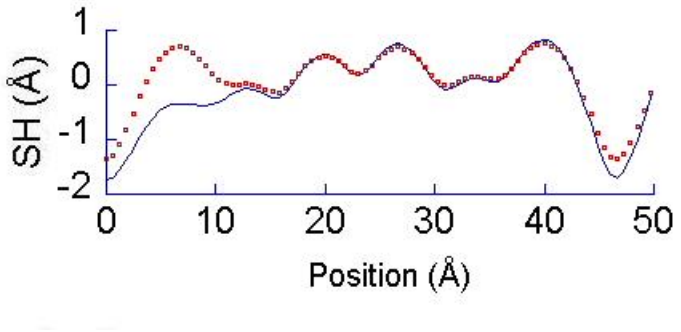


Figure 5-18: Comparison between two scan lines over the longest diagonal of the unit cell. The solid curve represents the scan line obtained from picture (b) in Figure 5-17 and the dots denote the scan line acquired from picture (a) in Figure 5-17. The sample bias is -2.0 V and THSP is 7.86 Å.

An obvious decrease by about 1.0 Å in tip height at an adatom position (0 Å -10Å) can be seen in Figure 5-18 due to the loss of the adatom dangling bond at the site. The other noticeable difference happens at the CH (about 47 Å). The tip goes deeper into the CH because the absence of the UC dangling bond resulted in a diluted DOS in CH region. The next step is to simulate the

adsorption of one hydrogen atom onto the rest atom site E by using exactly the same method. The hydrogen atom sits directly above a rest atom site [27] as shown in Figure 5-19.

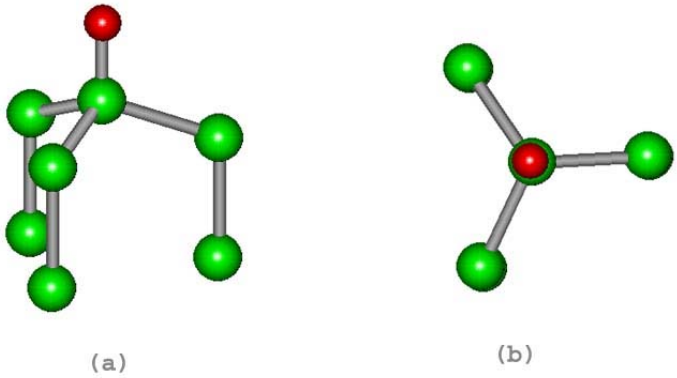


Figure 5-19: The optimized geometry corresponding to the chemisorption of one hydrogen atom onto a rest atom (a) side view; (b) top view [27].

Rest atoms can hardly be seen in the STM topography due to the image's saturation with adatoms as shown in experimental and simulated images reported everywhere. Therefore we will not observe an obvious difference by directly comparing the features of the two topographs acquired with the dangling bond of a rest atom enabled and disabled. A proper approach to this problem is to examine the scan lines.

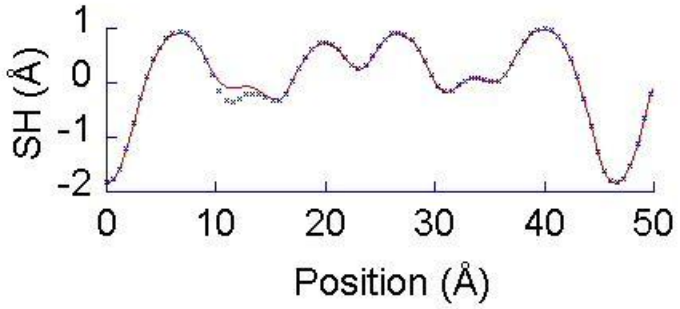


Figure 5-20: Comparison between two scan lines over the longest diagonal of the unit cell. The solid curve represents the scan line obtained in the "unreacted" case and the "cross" dots denote the scan line acquired in the "reacted" case.

A slight drop in tip height at around the saddle point of the middle adatoms (at ~ 13 Å) can be seen in Figure 5-20 due to the loss of the rest atom dangling bond. The tip height stays the same at the CH because the rest atom site is far enough from the CH region and thus the DOS at the CH can hardly be affected by the rest atom. We can conclude that it is very hard to detect hydrogen on rest atoms using topography alone.

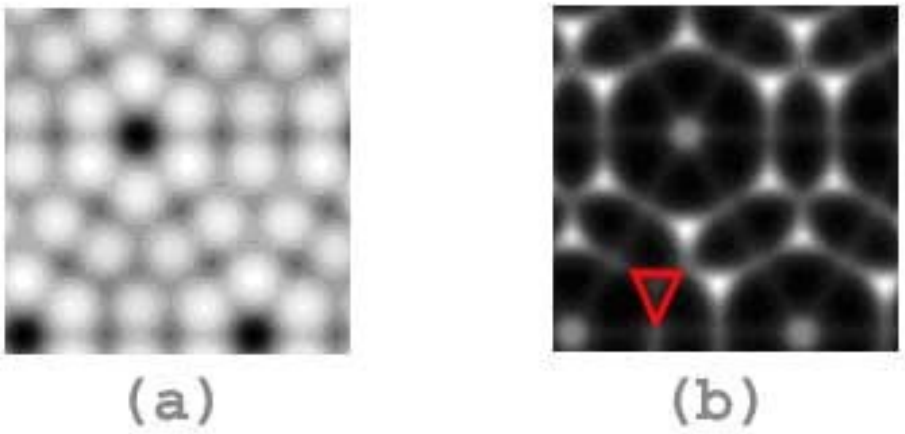


Figure 5-21: Topographic (a) and spectroscopic (b) images of the adsorption of one hydrogen atom onto a rest atom site in Si (111)-(7x7) surface. The triangle surrounds the rest atom adsorption site. The images are simulated at a sample bias of -0.8 V and THSP is about 7.0 Å.

Another way to investigate the adsorption is to take the spectroscopic image at a sample bias of around -0.8 V at which the features of rest atom dangling bonds can be resolved. We have obtained in one unit cell six high DOS regions that correspond to the dangling bonds state of rest atoms in Figure 5-15. The explanation is that the energy state for the rest atom dangling bonds is 0.8 eV below the Fermi level according to our previous discussion in Chapter 4. In Figure 5-21 (b), we observe that one rest atom site marked with red appears dark compared to Figure 5-15 (b). This observation confirms a disabled dangling bond of rest atom. It is also hard to tell the difference between Figure 5-15 (a) and Figure 5-21 (a), even though the rest atom dangling bond was disabled in the latter picture of topograph.

5.6.2 Simulated normalized conductance spectrum over feature sites

The use of the $dI/dV*V/I$ spectra has been very successful for the study and the understanding of surface chemistry, i.e. the $dI/dV*V/I$ spectra over both reacted and un-reacted sites have been studied when the Si (111)-(7x7) surface is exposed to the molecules of NH_3 [16] or H_2.[25] In this section, we calculate the $dI/dV*V/I$ spectra over features in the DAS model for Si (111)-(7x7) surface. The features include: adatom sites, reacted or unreacted and rest atom sites, reacted or unreacted. At the set point , the sample bias is -2.0 V. The spectra over reacted (dash line) and unreacted (solid line) rest atoms have been calculated. The comparison is shown in Figure 5-22.

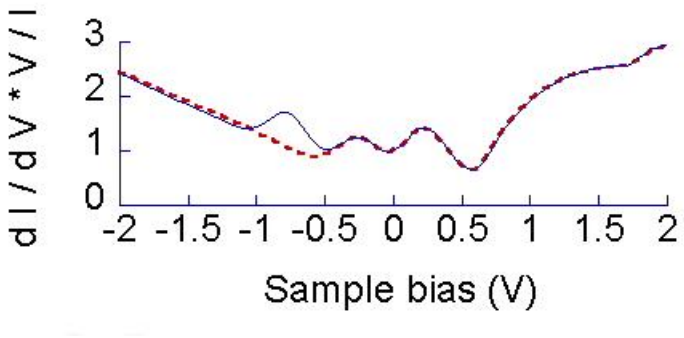


Figure 5-22: Comparison of two normalized conductance curves obtained with the tip over a reacted (dashed line) and an unreacted (solid line) UR site "e" shown in Figure 5-9.

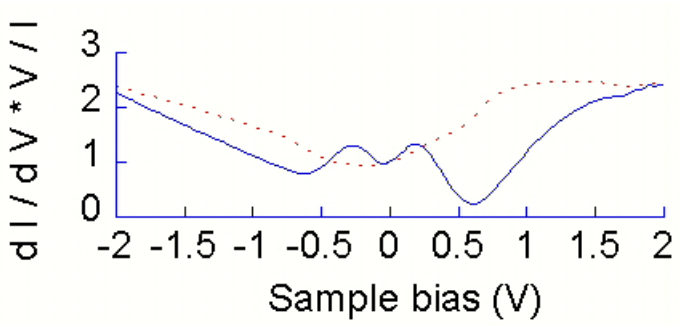


Figure 5-23: Comparison of two normalized conductance curves obtained with the tip over a reacted (dashed line) and unreacted (solid line) UC site "a" shown in Figure 5-9.

In Figure 5-22, a peak at -0.8 V that corresponds to the rest atom dangling bond state can not be seen in the dashed line but can be seen in the solid curve because the rest atom dangling bond state does not exist on the reacted rest atom site any more. In Figure 5-23, with the same set point voltage, the spectra over reacted and unreacted adatoms are obtained and compared. The two peaks that correspond to the adatom atom dangling bond state can not be seen in the dashed line but can be seen in the solid curve because the adatom bond state does not exist on the reacted adatom atom site.

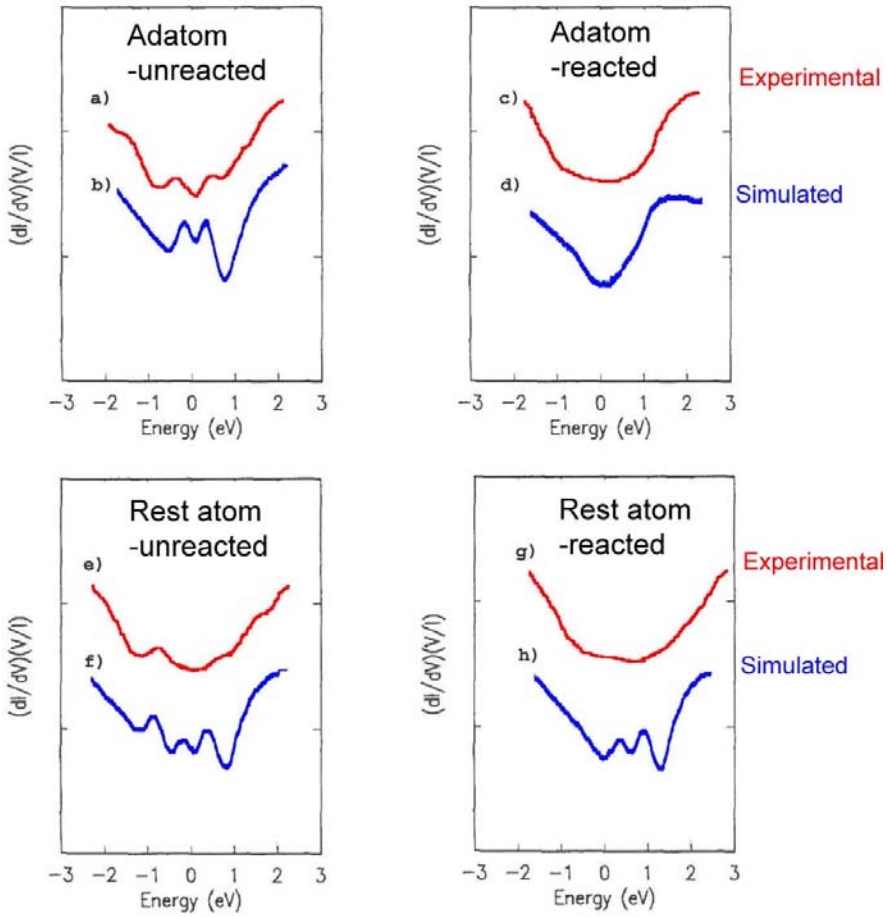


Figure 5-24: The qualitative comparison between experimental spectra and calculated results. Curves (a) and (b) refer to an un-reacted adatom site. Curves (c) and (d) refer to a reacted adatom site. Curves (e) and (f) refer to an un-reacted rest atom site. Curves (g) and (h) show a reacted rest atom site. In each set of the two curves, the setup of the figure always makes the experimental curve above the simulated curve, which means, for example, curve a is experimental and curve b is simulated . All the experimental curves are from Reference [25].

In Figure 5-24, curves (a) and (b) resemble each other in the sense of the two peaks at the vicinity of zero energy. In both curves (c) and (d) , we could not observe the two adatom feature peaks any more because of the dangling bonds are removed from the reacted adatom sites. Curves (e) and (f) both have a peak at the energy around -0.8 eV , which has been identified as a

feature that comes from the rest atom dangling bond state. The two relatively small peaks in curve (f) are the features from the adatom dangling bond state, which are not observed in curve (e). The explanation is that the ideal tip used in our calculation must have picked up the tails of adatom wave functions while the real tip is not sharp enough to detect that information . Finally, in both curves (g) and (h), the peak of high magnitude disappeared because the rest atom site under investigation is reacted and thus its dangling bond is removed.

5.7 Summary

The calculated topography at positive sample bias show similar corrugation to experiment: all the adatom sites are of almost the same height. At negative sample bias, the calculated topography shows contrast between corner and middle adatoms as well as between faulted and unfaulted halves of the unit cell, which agrees with experiment. As expected, the simulated $dI/dV*V/I$ or spectroscopic images at -0.8 V show rest atoms. The $dI/dV*V/I$ spectra over adatom sites at +1.0 V are similar in a certain extent to experimental results obtained by Boland and R. Wolkow.

STM imaging and tunneling spectroscopies of a single hydrogen atom on the Si(111)-(7x7) surface are studied. Simulated results show that the tip heights over reacted sites are depressed due to the loss of the dangling bonds. A depressed region in topography is observed as expected. In the simulated $dI/dV*V/I$ spectra, the disappearance of dangling bond features is in accordance with experimental results.

Chapter 6
Conclusion and future work

This thesis addresses the applicability of computer simulation of STM and STS imaging to the STM study of the clean or molecule-covered Si (111)-(7x7) surfaces. In this work, simple atomic orbital models for the clean Si (111)-(7x7) surface were developed. Our simulated topography at positive sample bias show similar corrugation to experiment: all the adatom sites are of the same height. At negative sample bias, the simulated topography shows contrast between corner and middle adatoms as well as between faulted and unfaulted halves of the unit cell, which also agrees with experiment. The simulated dI/dV or spectroscopic images at -0.7 V shows rest atoms as expected. Simulated normalized conductance spectra over adatom sites at +1.0 V are similar to experimental results obtained by Boland.

STM imaging and tunneling spectroscopies of a single hydrogen atom on the Si (111)-(7x7) surface are studied. Simulated results show that the tip heights over reacted sites are depressed due to the loss of the dangling bonds. A depressed region in topography is observed as expected. In the simulated normalized conductance spectra, the disappearance of dangling bond features is in accordance with experimental results.

The result is a model that uses essential results from total-energy pseudopotential density-functional theory and yet is easy to implement and is not significantly slower than standard topographic imaging. The model simulates well a very complicated physical system, i.e. the clean Si (111)-(7x7) surface, and is highly adaptable for adding adsorbates for detailed study.

6.1 Future work

The main work for the future includes the simulation of the STM and STS imaging of molecules adsorbed onto the Si (111)-(7x7) surface. More complicated molecules such as benzene, acetylene and ethylene are possible options for the study. Different possible binding geometries will be modeled and the resultant STM and STS imaging will be compared to experimental results. The comparison will possibly help us understand the adsorption or desorption process that happens on the Si (111)-(7x7) surface. This will definitely increase the complexity of program coding and require more computational effort. Therefore, the optimization of algorithms should also be considered to reduce the execution time of the simulation program.

Without sacrificing too much running time, the hydrogen-like wave functions might be replaced with the more exact wave functions of silicon atomic orbitals. Thus, more accuracy in tunneling current could be expected as the current is highly dependent on the wave functions. Actually we can also include exchange and correlation effects to get things right.

Bibliography

1. G.Binnig, H. Rohrer, Ch. Gerber and E. Weibel. "7x7 reconstruction on Si (111) resolved in real space" *Phys.Rev.Lett.* **50**, 120-123 (1983).

2. R.J.Hamers , R. M. Tromp and J. E. Demuth. "Surface electronic structures of Si (111)-(7x7) resolved in real space" *Phys.Rev.Lett* **56**, 1972-1975 (1986).

3. R.M.Feenstra, J. A. Stroscio and A. P. Fein. "Tunneling spectroscopy of the Si (111) 2x1 surface" *Surf.Sci.* **181**, 295-306 (1987).

4. D.R.Vij. "Handbook of Applied Solid State Spectroscopy". Springer-Verlag: New York, NY (2006).

5. J.Tersoff and D.R.Hamann. "Theory of the scanning tunneling microscope" *Phys.Rev.Lett* **31**, 805-813 (1985).

6. K.Takayanagi, Y. Tanishiro, M. Takahashi and S. Takahashi. "Structural analysis of Si(111)-7x7 by UHV-transmission electron diffraction and microscopy" *J.Vac.Sci.Technol.A* **3**, 1502-1506 (1985).

7. A.P.French and E.F.Taylor. "An Introduction to Quantum Physics".W·W· Norton&Company·INC: New York, NY (1978).

8. R.J.Colton. "Nanoscale measurements and manipulation (Review article)" *J.Vac.Sci.Technol.B* **22**, 1609-1632 (2004).

9. http://www.infoplease.com/ce6/sci/A0839004.html

10. J.A.Stroscio and W.J.Kaiser. "Scanning Tunneling Microscopy, Volume 27,Method of Experimental Physics". Academic Press: Boston, MA (1993).

11. http://www.almaden.ibm.com/vis/stm/atomo.html

12. J.Bardeen. "Tunneling from a many-particle point of view" *Phys.Rev.Lett* **6**, 57-59 (1961).

13. N.D.Lang. "Spectroscopy of single atoms in the scanning tunneling microscope" *Phys.Rev.B* **34**, 5947-5950 (1986).

14. G.Binnig, K. H. Frank, H. Fuchs, N. Garcia, B. Reihl, H. Rohrer, F. Salvan and A. R. Williams. "Tunneling spectroscopy and inverse photoemission: image and field states" *Phys.Rev.Lett.* **55**, 991-994 (1985).

15. Ph.Avouris and R.Wolkow. "Atom-resolved surface chemistry studied by scanning tunneling microscopy and spectroscopy" *Phys.Rev.B* **39**, 5091-5100 (1989).

16. R.Wolkow and Ph.Avouris. "Atom-resolved surface chemistry using scanning tunneling microscopy" *Phys.Rev.Lett* **60**, 1049-1052 (1988).

17. W.A.Hofer, A. S. Foster and A. L. Shluger. "Theories of scanning probe microscopes at the atomic scale" *Rev.Mod.Phys.* **75**, 1287-1331 (2003).

18. R.M.Tromp, R. J. Hamers and J. E. Demuth. "Atomic and electronic contributions to Si (111)-(7x7) scanning-tunneling-microscopy images" *Phys.Rev.Lett.* **34**, 1388-1390 (1986).

19. A.Selloni , P. Carnevali, E. Tosatti and C. D. Chen. "Voltage-dependent scanning-tunneling microscopy of a crystal surface: Graphite" *Phys.Rev.B* **31**, 2602-2605 (1985).

20. J.Pollmann, P. Krüger and A. Mazur. "Self-consistent electronic structure of semi-infinite Si(001) (2x1) and Ge(001) (2x1) with model calculations for scanning tunneling microscopy" *J.Vac.Sci.Technol.B* **5**, 945-952 (1987).

21. J.H.Wilson , D. A. McInnes, J. Knall, A. P. Sutton and J. B. Pethica. "Quantitative voltage-dependent STM image simulations for semiconductors" *Ultramicroscopy* **42-44**, 801-808 (1992).

22. J.A.Stroscio, R. M. Feenstra, D. M. Newns and A. P. Fein. "Voltage-dependent scanning tunneling microscopy imaging of semiconductor surfaces" *J.Vac.Sci.Technol.A* **6**, 499-507 (1988).

23. K.D.Brommer , B. E. Larson, M. Needels and J. D. Joannopoulos. "Modeling large surface reconstructions on the connection machine" *Jpn.J.Appl.Phys.* **32**, 1360-1367 (1993).

24. C.L.Wu , J. L. Hsieh, H. Huseh and S. Gwo. "Thermal nitridation of the Si (111)-7x7 surface studied by scanning tunneling microscopy and spectroscopy" *Phys.Rev.B* **65**, 045309 1-6 (2002).

25. J.J.Boland. "The importance of structure and bonding in semiconductor surface chemistry: hydrogen on the Si (111)-7x7 surface" *Surf.Sci.* **IS244**, 1-14 (1991).

26. L.Bolotov, N. Uchida and T. Kanayama. "Electronic properties of assembled islands of hydrogen-saturated silicon clusters on Si(111)-(7x7) surfaces by scanning tunneling spectroscopy" *Eur.Phys.J.D* **16**, 271-274 (2001).

27. J.Que , M. W. Radny and P. V. Smith. "High exposure hydrogen chemisorption on the Si (111) 7x7 surface: a semiempirical cluster study" *J.Phys.: Condens.Matter* **8**, 4205-4222 (1996).

28. K.D.Brommer, M. Galvan, A. D. Pino Jr. and J. D. Joannopoulos. "Theory of adsorption of atoms and molecules on Si (111)-(7x7)" *Surf.Sci.* **314**, 57-70 (1994).

29. R.S.Becker, J. A. Golovchenko, D. R. Hamann and B. S. Swartzentruber. "Real-Space Observation of Surface States on Si (111) 7x7 with the Tunneling Microscope" *Phys.Rev.Lett.* **55**, 2032-2034 (1985).

30. D.Wortmann. "Interpretation of Scanning Tunneling Microscopy and Spectroscopy of Magnetic Metal surfaces by Electron Theory". Thesis for Diplom-Physi, Department of Physics at the University of Dortmund (2000).

31. S.Heinze, R. Abt and S. Blügel. "Scanning tunneling microscopy images of transition-metal structures buried below noble-metal surfaces" *Phys.Rev.Lett* **83**, 4808-4811 (1999).

32. V.P.LaBella, H. Yang, D. W. Bullock and P. M. Thibado. "Atomic structure of the GaAs(001)-(2x4) surface resolved using scanning tunneling microscopy and first-principles theory" *Phys.Rev.Lett* **83**, 2989-2992 (1999).

33. N.Pertaya. "The a-Ga (010) surface investigated by room and low temperature scanning tunneling microscopy and helium atom scattering". Dissertation for The degree of Doctor of Natural Science, Department of Physics, Free University Berlin. (2004).

34. K.D.Brommer, M. Needels, B. E. Larson and J. D. Joannopoulos. "Ab initio theory of the Si (111)-(7x7) surface reconstruction: a challenge for massively parallel computation" *Phys.Rev.Lett* **68**, 1355-1358 (1992).

35. G.X.Qian and D.J.Chadi. "Si(111)-7×7 surface: Energy-minimization calculation for the dimer–adatom–stacking-fault model" *Phys.Rev.B* **35**, 1288-1293 (1987).

36. C.E.Housecroft and A.G.Sharpe. "Inorganic Chemistry". Prentice Hall: Toronto, ON (2001).

37. J.E.Northrup. "Origin of Surface States on Si(111) (7x7)" *Phys.Rev.Lett.* **57**, 154-157 (1986).

38. H.Kim, K. Cho, I. Park and J. D. Joannopoulos. "Ab initio study of hydrogen adsorption on the Si (111)-(7x7) surface" *Phys.Rev.B* **52**, 17 231-17 237 (1995).

39. D.A.Papaconstantopoulos. "Handbook of the band structure of elemental solids". Plenum Press: New York, NY (1986).

40. R.M.Tromp, E. J. van Loenen, J. E. Demuth and N. D. Lang. "Tip electronic structure in scanning tunneling microscopy" *Phys.Rev.B* **37**, 9042-9045 (1988).

41. C.J.Chen. "Tunneling matrix elements in three-dimensional space: the derivative rule and the sum rule." *Phys.Rev.B* **42**, 8841-8857 (1990).

42. M.N.Piancastelli, N. Motta, A. Sgarlata, A. Balzarotti and M. D. Crescenzi. "Topographic and spectroscopic analysis of ethylene adsorption on Si (111) 7x7 by STM and STS" *Phys.Rev.B* **48**, 17892-17896 (1993).

43. F.Rochet, F. Jolly, F. Bournel, G. Dufour, F. Sirotti and J. Cantin. "Ethylene on Si (001)-2x1 and Si (111)-7x7: X- ray photoemission spectroscopy with synchrotron radiation" *Phys.Rev.B* **58**, 11029-11042 (1998).

44. M.Carbone, R. Zanoni, M. N. Piancastelli, G. Comtet, G. Dujardin, L. Hellner and A. Mayne. "Adsorption of ethylene on Si (111) 7x7 by synchrotron radiation photoemission" *J.Electron Spectrosc.* **76**, 271-276 (1995).

45. Y.Taguchi, M. Fujisawa and M. Nishijima. "Adsorbed state of benzene on the Si (111) (7x7) surface" *Chem.Phys.Lett.* **178**, 363-368 (1991).

46. I.D.Petsalakis , J. C. Polanyi and G. Theodorakopoulos. "Theoretical study of the induced attachment of benzene to Si (111) 7x7" *Surf.Sci.* **544**, 162-169 (2003).

47. G.Jiang, J. C. Polanyi and D. Rogers. "Electron and photon irradiation of benzene and chlorobenzene on Si (111) 7x7" *Surf.Sci* **544**, 147-161 (2003).

48. H.Tomimoto, T. Sekitani, R. Sumii, E. O. Sako, S. Wada and K. Tanaka. "Study of adsorption structure of benzene and toluene on Si (111) 7x7 surfaces" *Surf.Sci.* **566**, 664-670 (2004).

49. B.Weiner, C. S. Carmer and M. Frenklach. "Acetylene reaction with the Si (111) surface : A semiempirical quantum chemical study" *Phys.Rev.B* **43**, 1678-1684 (1991).

50. F.Rochet, G. Dufour, P. Prieto, F. Sirotti and F. C. Stedile. "Electronic structure of acetylene on Si (111)-7x7: X-ray photoelectron and x-ray absorption spectroscopy" *Phys.Rev.B* **57**, 6738-6748 (1998).

51. J.Yoshinobu, H. Tsuda, M. Onchi and M. Nishijima. "Rehybridization of acetylene on the Si (111) (7x7) surface - A vibrational study" *Chem.Phys.Lett.* **130**, 170-174 (1986).

52. X.Lu, X. Wang, Q. Yuan and Q. Zhang. "Diradical mechanisms for the cycloaddition reactions of 1,3-butadiene, benzene, thiophene, ethylene, and acetylene on a Si(111)-7x7 surface" *J.Am.Chem.Soc.* **125**, 7923-7929 (2003).

53. C.Sbraccia, C. A. Pignedoli, A. Catellani, R. D. Felice, P. L. Silvestrelli, F. Toigo, F. Ancilotto and C. M. Bertoni. "Chemisorption sites and reaction pathways for acetylene on Si(111)-(7x7)" *Surf.Sci.* **557**, 80-90 (2004).

54. C.Sbraccia, C. A. Pignedoli, A. Catellani, R. D. Felice, P. L. Silvestrelli, F. Toigo, F. Ancilotto and C. M. Bertoni. "Acetylene on Si (111) from computer simulations" *Computer Physics Communications* **169**, 32-35 (2005).

55. S.Chu and A.B.Anderson. "Acetylene adsorption on Si (111): Molecular orbital theory" *Surf.Sci.* **194**, 55-62 (1988).

56. L.E.Kinsler. " Fundamentals of Acoustics".Wiley: New York, NY (1999).

Appendix A

Feedback response to a variety of signals

Transient response to a delta function

We consider $b(t) = B_0\delta(t)$ as a delta function signal. Its frequency domain function is:

$$B(\omega) = \int_{-\infty}^{\infty} b(t)e^{-j\omega t}dt = B_0 \qquad (A\text{-}1)$$

where B_0 has a unit of voltage.

Then equation (2-33) can be rewritten

$$E(\omega) = -j\omega\tau R_{pa}GB_0 \big/ \left[2\pi(1+j\omega\tau)\right] \qquad (A\text{-}2)$$

Its counterpart in the time domain is:

$$e(t) = -j\,R_{pa}GB_0 \big/ (2\pi) \int_{-\infty}^{\infty} \omega\tau \big/ (1+j\omega\tau) e^{j\omega t}d\omega \qquad (A\text{-}3)$$

This integral in equation (A- 3) can be evaluated using residue theory. In Figure A-1, if $R \to \infty$, the integral over the path C_R is zero since $e^{-\frac{t}{\tau}}$ is very small for points far above the real axis. We can replace the integral $\int_{-\infty}^{\infty} \omega e^{j\omega t}d\omega \big/ (1+j\omega\tau)$ with a contour integral:

$$\int_{-\infty}^{\infty} \omega e^{j\omega t}d\omega \big/ (1+j\omega\tau) = (-j/\tau)(2\pi j)\,\mathrm{Res}\left[je^{-t/\tau}/\tau\right] = 2\pi e^{-t/\tau}/\tau^2 \qquad (A\text{-}4)$$

$$e(t) = -j\,B_0R_{pa}G\tau \big/ (2\pi) \int_{-\infty}^{\infty} \omega e^{j\omega t}d\omega \big/ (1+j\omega\tau) = -j\,B_0R_{pa}Ge^{-t/\tau}/\tau \qquad (A\text{-}5)$$

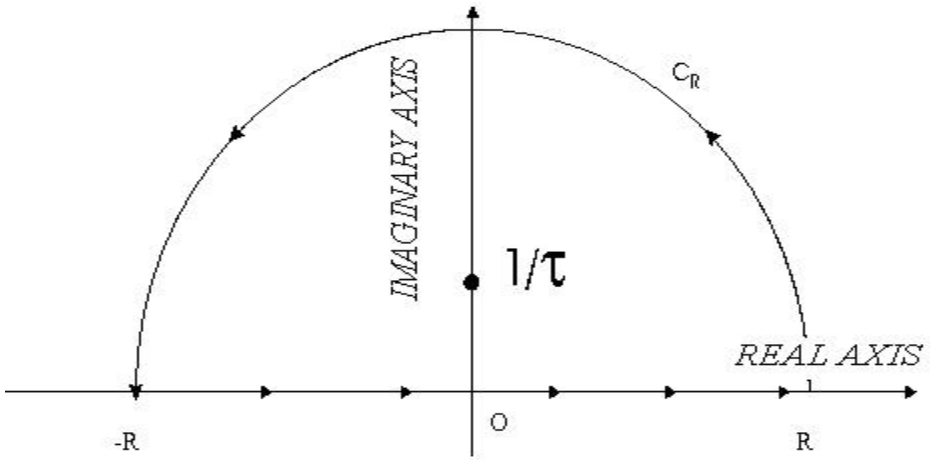


Figure A-1: Path C_R chosen to solve an integral by residue theory

After several time constants, the error signal goes to zero. One can also calculate the transient response to a step input.

Transient response to a step function

The ac signal is described by:

$b(t) = B_0^{'}\theta(t)$,

where $\theta(t) = \begin{cases} 1 & t > 0 \\ 0 & t \le 0 \end{cases}$, is defined as a step function.

Its first derivative is the delta function. $\frac{d\theta(t)}{dt} = \delta(t)$. Its counterpart in the frequency domain is:

$$B(\omega) = \int_{-\infty}^{+\infty} B_0^{'}\theta(t)e^{-j\omega t}dt = B_0^{'}\theta(t)\frac{e^{-j\omega t}}{-j\omega}\Bigg|_{-\infty}^{\infty} - B_0^{'}\int_{-\infty}^{+\infty}\delta(t)\frac{e^{-j\omega t}}{-j\omega}dt = \frac{-jB_0^{'}}{\omega} \qquad (A\text{-}6)$$

Equation (2-33) is rewritten after the introduction of equation (A- 6):

$$E(\omega) = jB_0^{'}R_{pa}G/(\omega - j/\tau) \qquad (A\text{-}7)$$

and correspondingly in the time domain

$$e(t) = 1/(2\pi)\int_{-\infty}^{\infty} jB_0^{'}R_{pa}Ge^{j\omega t}d\omega/(\omega - j/\tau) = jB_0^{'}R_{pa}Ge^{-t/\tau} \qquad (A\text{-}8)$$

The residue theorem has been used to evaluate the integral in equation (A- 8). The equation shows us the transient response of the feedback loop to a step in sample bias. After several time constants, the feedback loop settles down.

Feedback response to a sine wave input (frequency response)

$$B(\omega) = b_0\delta(\omega - \omega_0) \qquad (A\text{-}9)$$

(The ac voltage signal in the time domain is**:** $b(t) = 1/(2\pi)\int_{-\infty}^{\infty} B(\omega)e^{j\omega t}d\omega = b_0e^{j\omega_0 t}/(2\pi)$**).**

Combining equation (2-33) and equation (A- 9) gives**:**

$$E(\omega) = -j\omega\tau R_{pa}Gb_0\delta(\omega - \omega_0)/(1 + j\omega\tau) \qquad (A\text{-}10)$$

The error signal in the time domain is then given by

$$e(t) = -j\omega_0\tau R_{pa}Gb_0e^{j\omega_0 t}/[2\pi(1 + j\omega_0\tau)] \qquad (A\text{-}11)$$

$$|e(t)| = \omega_0 \tau R_{pa} G b_0 \Big/ \left[2\pi \left(\sqrt{1 + (\omega_0 \tau)^2} \right) \right] \qquad \text{(A- 12)}$$

We plot in Figure A-2 $|e(t)| / R_{pa} G b_0$ as a function of $\omega_0 \tau$ and the two variables are both dimensionless.

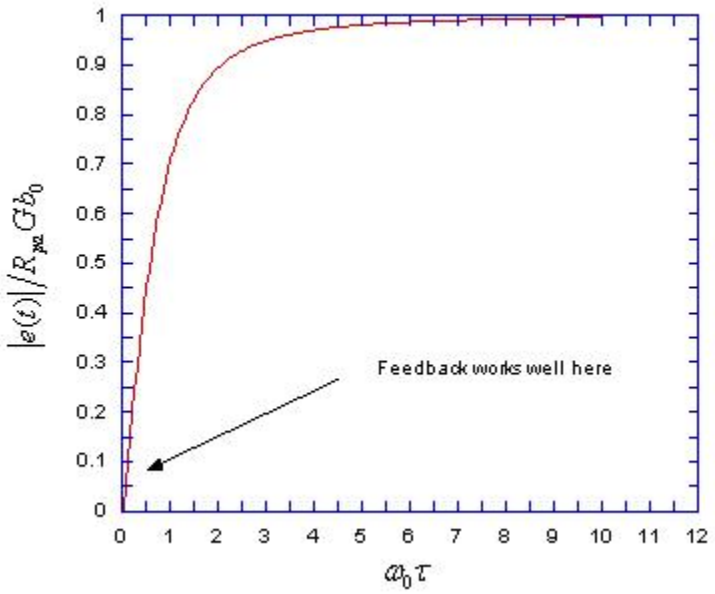


Figure A-2: The feedback response to a sine wave input

Figure A-2 shows that when ω_0 is small, the feedback works well. This means that when the ac signal frequency is low, the feedback can catch up to the change and responds by nulling the error signal.

Introducing equations (2-23) to (2-25) , equation (A-10) and equation (2-17) into equation (2-26) leads to

$$I(\omega) = \frac{b_{DC}}{R_J}\delta(\omega) - \frac{Gb_0\delta(\omega-\omega_0)}{1+j\omega\tau} + Gb_0\delta(\omega-\omega_0) \qquad (A\text{-}13)$$

The second term on the right side of this equation comes from the feedback loop. The third term is the driving ac signal. So the negative sign in the second term is very important, which comes from the integrator in the feedback loop. Now we see that within the bandwidth of the feedback loop $\omega\tau << 1$ the feedback (second term) cancels off the effect of the external signal (third term) to keep the current constant.

Now we start our discussion on equation (2-41). For the purpose of factorization, the denominator in equation (2-41) is set to be zero.

$$\omega^3 - j\,\omega^2\omega_p/Q - \omega\omega_p^2 + j\,\omega_p^2/\tau = 0 \qquad (A\text{-}14)$$

We find that the last term on the left side of the equation.(A- 14) is zero when τ is very large.

$$\omega^3 - j\,\omega^2\omega_p/Q - \omega\omega_p^2 = 0 \qquad (A\text{-}15)$$

If $\omega = 0$ is assumed to be a solution in equation (A- 15), the actual solution would be $\omega = \frac{j}{\tau}$ from equation (A- 14) where two high order terms are ignored since $\omega << 1$. The two other roots of ω can be obtained by solving

$$\omega^2 - j\,\omega\omega_p/Q - \omega_p{}^2 = 0 \qquad (A\text{-}16)$$

The two roots for the equation (A- 16) are:

$$\omega = j\,\omega_p/(2Q) \pm \omega_p\sqrt{1-1/(4Q)^2} \qquad (A\text{-}17)$$

Rewriting equation (A- 17):

$$\omega/\omega_p = j/(2Q) \pm \sqrt{1 - 1/(4Q)^2} \qquad \text{(A- 18)}$$

Equation (A- 18) implies that ω/ω_p is a circle with a radius of 1.0 in a complex plane.

For $B(\omega) = b_0\delta(\omega - \omega_0)$ (response to sine wave ac signal), consider what happens when

$\omega_0 = \omega_p$

Equation (2-41) becomes equation (A- 19) and then is transformed into the new domain as shown in equation (A- 20):

$$E(\omega) = \frac{\omega(\omega_p^2 - \omega^2 + j\,\omega\omega_p/Q)R_{pa}Gb_0\delta(\omega - \omega_p)}{\omega^3 - j\,\omega^2\omega_p/Q - \omega\omega_p^2 + j\,\omega_p^2/\tau} \qquad \text{(A- 19)}$$

$$e(t) == \frac{-R_{pa}Gb_0e^{j\omega_p t}}{2\pi(1 - Q/\omega_p\tau)} \qquad \text{(A- 20)}$$

For large τ (small gain), $\omega_p\tau >> Q$

Then $e(t) = -R_{pa}Gb_0e^{j\omega_p t}/2\pi$, which is the same as equation (A- 11) if $\omega_0\tau >> 1$. As we decrease τ, $E(\omega)$ in equation (A- 19) gets larger and can actually diverge

Now, we look for the general solution of the equation (A- 14):

1. When $Q >> 1$ and $Q >> \omega_p\tau$,consider a sub-case

   1) $\omega_p\tau >> 1$ (Low gain)

   The solutions are given in (E-13) to (E-15): (Refer to Appendix E).

$\omega_1 \approx \omega_p - j/(2\tau)$, $\omega_2 \approx -\omega_p - j/(2\tau)$, and $\omega_3 \approx j/\tau$.

For a delta function input, we bring equation (A- 6) into equation (2-41) and the latter becomes

$$E(\omega) = \frac{-j\left(\omega_p^2 - \omega^2 + j\,\omega\omega_p/Q\right)R_{pa}GB_0^{'}}{\omega^3 - j\,\omega^2\omega_p/Q - \omega\omega_p^2 + j\,\omega_p^2/\tau} \qquad \text{(A- 21)}$$

Since ω_3 is the only pole above the real axis, we need evaluate only one residue at $\omega = \omega_3$

$$\begin{aligned} e(t) &= 1/(2\pi)\int_{-\infty}^{\infty} E(\omega)e^{j\omega t}d\omega \\ &= jR_{pa}GB_0^{'}\mathrm{Res}\left[\frac{-j\left(\omega_p{}^2 - \omega_3{}^2 + j\,\omega_3\omega_p/Q\right)e^{j\omega_3 t}}{(\omega_3 - \omega_1)(\omega_3 - \omega_2)}\right] \end{aligned} \qquad \text{(A- 22)}$$

Since $\omega_3 = j/\tau << 1$, $Q >> 1$, $\omega_3 - \omega_2 \approx \omega_p$ and $\omega_3 - \omega_1 \approx -\omega_p$, equation (A- 22) is simplified to be:

$$e(t) \approx jR_{pa}GB_0^{'}e^{-t/\tau} \qquad \text{(A- 23)}$$

This result is exactly the one we have discussed without a consideration of tube resonance. The following diagram explains why the results are the same.

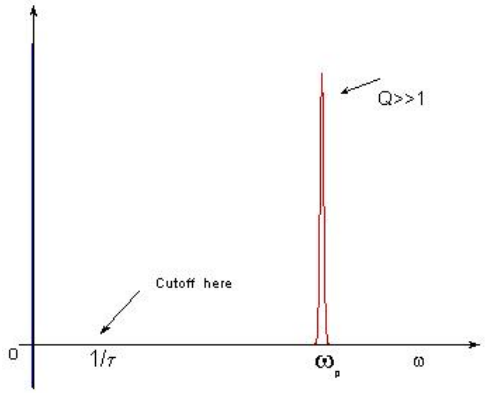


Figure A-3: The piezo resonant frequency is far greater than the cut off frequency

We cannot see the tube resonance because the tube resonant frequency is far greater than the cut off frequency shown in Figure A-3.

2) $Q >> 1 >> \omega_p \tau$ **(**High gain**)**

The solutions are (E-16) to (E-18): (Refer to Appendix E)

$$\omega_1 \approx \left(\omega_p^2/\tau\right)^{1/3} e^{-j\frac{\pi}{6}}\ ,\ \omega_2 \approx \left(\omega_p^2/\tau\right)^{1/3} e^{-j\frac{5\pi}{6}} \text{ and } \omega_3 \approx \left(\omega_p^2/\tau\right)^{1/3}\ j$$

The inverse Fourier transform of equation (A- 21) is

$$\begin{aligned} e(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} E(\omega) e^{j\omega t} d\omega \\ &= jR_{pa} GB_0^{'}\ \mathrm{Re}\,s\left[\frac{\omega_3\left(\omega_p{}^2 - \omega_3{}^2 + j\,\omega_3\omega_p/Q\right)e^{j\omega_3 t}}{(\omega_3-\omega_1)(\omega_3-\omega_2)}\right] \end{aligned} \qquad \text{(A- 24)}$$

The residue theorem is used since ω_3 is the only pole above the real axis.

Since $\omega_p{}^2 << 1$, $\omega_3\omega_p/Q << 1$, $\omega_3 - \omega_1 = \sqrt{3}e^{j\pi/6}\omega_3$ and $\omega_3 - \omega_2 = \sqrt{3}e^{-j\pi/6}\omega_3$,

$$e(t) = jR_{pa} GB_0^{'}\left(\omega_p^2/\tau\right)^{1/3} e^{-\left(\omega_p^2/\tau\right)^{1/3} t}\Big/3 \qquad \text{(A- 25)}$$

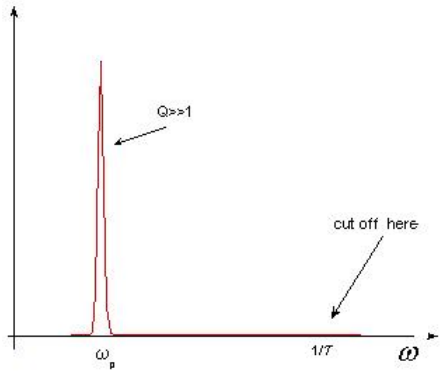


Figure A-4: The piezo tube frequency is smaller than the cut off frequency

Even though the resonance frequency is smaller than the cut-off frequency, the expression for $e(t)$ tells us that the feedback has a very fast response.

2. when $\omega_p \tau >> Q >> 1$

The solutions are presented in equation (E-21) to (E-23):

$$\omega_1 = \omega_p \left(1 + j/(3Q)\right),\ \omega_2 = \omega_p \left(-1 + j/(3Q)\right) \text{ and } \omega_3 = j\omega_p /(3Q).$$

All the poles are above the real axis. Using the residue theorem to solve the contour integral above the real axis, $e(t)$ will have 3 terms, all decaying like $e^{-\omega_p t/3Q}$ and with oscillation at ω_p inside this envelope.

Appendix B

The Laplace transformation

The Laplace transformation of a function $h(t)$ is defined in equation (B- 1):

$$H(s) = \int_0^\infty h(t) e^{-st} dt \qquad \text{(B- 1)}$$

The Laplace transform of the nth derivative $\frac{d^n h(t)}{dt^n}$ is:

$$s^n H(s) - s^{n-1} h(0) - s^{n-2} \frac{dh(0)}{dt} - \cdots - \frac{d^{n-1} h(0)}{dt^{n-1}} \qquad \text{(B- 2)}$$

The Laplace transform of nth-order integral $\frac{d^{-n} h(t)}{dt^{-n}}$ is

$$\frac{H(s)}{s^n} + \frac{\frac{d^{-1} h(0)}{dt^{-1}}}{s^n} + \cdots + \frac{\frac{d^{-n} h(0)}{dt^{-n}}}{s} \qquad \text{(B- 3)}$$

The table B-1 shows the counterpart in the s domain (frequency domain) of some important time-dependent functions in the t domain. Here we list them for we may use them in our following discussion.

Table B-1: Laplace transformation of time-dependent functions

Domain (t)	Domain(s)
$\delta(t)$	1
A	A/s
t	$1/s^2$
t^2	$2/s^3$

t^n	$n!/s^{(n+1)}$
e^{-at}	$1/(s+a)$
$\sin(\omega t)$	$\omega/(s^2+\omega^2)$
$\cos(\omega t)$	$s/(s^2+\omega^2)$
te^{-at}	$1/(s+a)^2$

We will use Laplace transformation to solve equation (B- 4):

$$e(t) = r(t) - \frac{b_{DC}}{R_J} R_{pa} \left(1 + 2\alpha (l/T) d_{31} g_v p K_i \int_{-\infty}^{t} e(t')dt' \right) \qquad \text{(B- 4)}$$

According to equation (B- 3), the Laplace transformation of the first-order integral $\int_{-\infty}^{t} e(t')dt'$ is:

$$L\left(\int_{-\infty}^{t} e(t')dt' \right) = \frac{E(s)}{s} + \frac{\int_{-\infty}^{0} e\left(t'\right) dt'}{s} \qquad \text{(B- 5)}$$

Since $\int_{-\infty}^{0} e(t)dt = 0$, and $b_{DC} R_{pa} / R_j$ is a constant, after performing Laplace transformation on both sides of equation (B- 4), we have the following equation in the s domain:

$$E\left(s\right) = R(s) - \frac{b_{DC}}{R_J s} R_{pa} - \frac{E\left(s\right)}{s\tau} \qquad \text{(B- 6)}$$

where $\tau = \left(2\alpha d_{31} g_v p K_i \, l/T\right)^{-1}$, the signal $r(t)$ is the reference voltage signal and it is defined by

$$r\left(t\right) = r_1 \quad \text{if } t > 0 \quad \text{and} \quad r\left(t\right) = b_{DC} R_{pa} / R_J \quad \text{if } t < 0$$

Therefore, the Laplace transformation of $r(t)$ is:

$$R(s) = \int_0^\infty r(t) e^{-st} dt = \frac{r_1}{s} \qquad \text{(B- 7)}$$

Therefore, equation (B- 6) becomes:

$$E(s) = \frac{r_1 - \frac{b_{DC}}{R_J} R_{pa}}{s + \frac{1}{\tau}} \qquad \text{(B- 8)}$$

By looking up the table B-1, we can quickly write the corresponding t domain function of the frequency domain function in equation (B- 8).

$$e(t) = \left(r_1 - \frac{b_{DC} R_{pa}}{R_J} \right) e^{-(t/\tau)} \qquad \text{(B- 9)}$$

Appendix C

Spherical Potential Well for the Tersoff and Hamann theory

In Tersoff-Hamann's model, the potential is assumed to be a spherical well potential. This means that the potential is all the same at any point inside the sphere and zero outside the sphere**.** In Figure C-1 the potential difference is assumed to V_0 and energy diagram are only sketched in one dimension of three-dimensional space.

$$V(r)=\begin{cases}-V_0 & r<a_0\\ 0 & r>a_0\end{cases} \qquad (C-1)$$

The symbols of a_0 and E are the radius of the sphere and the energy of electrons respectively.

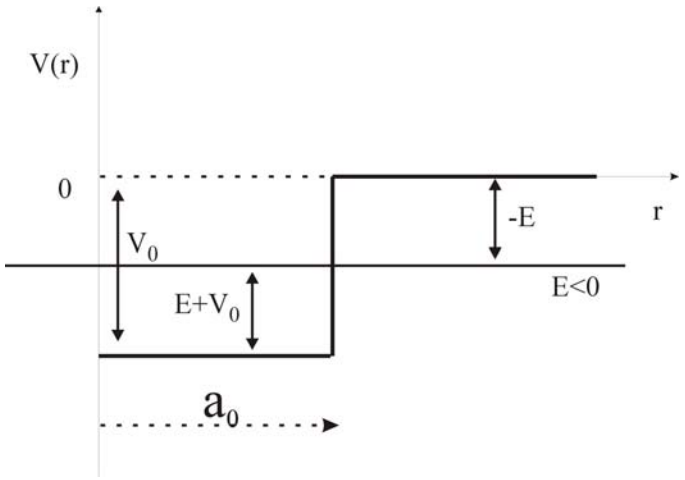


Figure C-1: The potential in Tersoff-Hamann's model on tip

According to basic quantum mechanics:

$$\left(-\frac{\hbar^2}{2m}\nabla^2+V\right)\psi=E\psi \qquad (C-2)$$

$$\psi = R_{nl}(r)Y_{lm}(\theta,\varphi) \qquad (C\text{-}3)$$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR_{nl}}{dr}\right)+\left[\frac{2m}{\hbar^2}\left(E-V(r)\right)-\frac{l(l+1)}{r^2}\right]R_{nl}=0 \qquad (C\text{-}4)$$

Two parameters are defined as follows:

$$\alpha=\sqrt{\frac{2m(E+V_0)}{\hbar^2}} \quad \text{and} \quad \beta=\sqrt{\frac{-2mE}{\hbar^2}}$$

For $l=0$ (s wave function), equation (C- 4) becomes if $r<a_0$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR_{n0}}{dr}\right)+\alpha^2R_{n0}=0 \qquad (C\text{-}5)$$

or if $r>a_0$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR_{n0}}{dr}\right)+\alpha^2R_{n0}=0 \qquad (C\text{-}6)$$

The solution of equation (C- 5) is :

$$R_{n0}(r)=A\frac{\sin(\alpha r)}{\alpha r} \qquad (C\text{-}7)$$

It can be verified by simply bringing equation (C- 7) into (C- 5) and checking the equivalence of both sides of equation (C- 5):

$$\begin{aligned}\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{dR_{n0}}{dr}\right)&=\frac{A}{\alpha r^2}\frac{d}{dr}\left(\alpha r\cos(\alpha r)-\sin(\alpha r)\right)\\&=-\frac{A\alpha\sin(\alpha r)}{r}=-\alpha^2R_{n0}\end{aligned} \qquad (C\text{-}8)$$

The solution of equation (C- 6) is:

$$R_{n0}(r) = C\frac{\exp(-\beta r)}{\beta r} \qquad (C\text{-}9)$$

which can be brought back to check in equation (C- 6).

The wave function is continuous at the boundary surface, which requires the equivalence of equations (C- 7) and (C- 9) at $r = a_0$:

$$C\frac{\exp(-\beta a_0)}{\beta a_0} = A\frac{\sin(\alpha a_0)}{\alpha a_0} \qquad (C\text{-}10)$$

The model shown in Figure C-2 assumes that the potential barrier is not large but that the tip radius of a_0 is large. If $\alpha \approx \beta$ and $\sin(\alpha a_0) \approx 1$, from equation (C- 10), we have

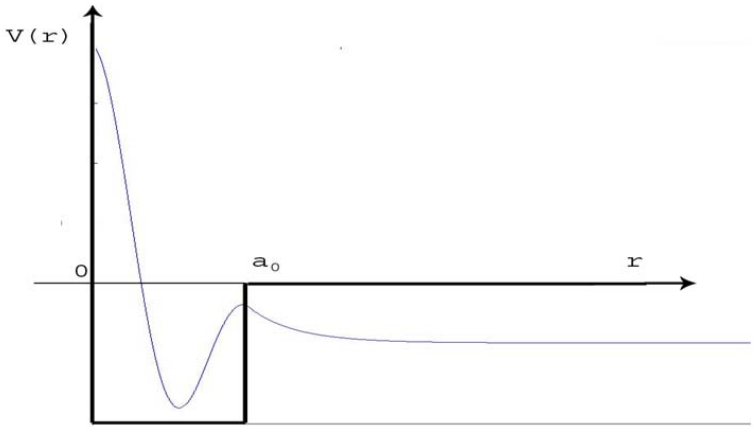


Figure C-2**:** Behaviour of the wave function in a spherical square well

:

$$A \approx \exp(-\beta a_0)C \text{ except sometimes } \exp(-\beta a_0)C << A$$

The normalization of the wave function over all space is considered :

$$1 = 4\pi A^2 \int_0^{a_0} r^2 dr \frac{\sin^2(\alpha r)}{(\alpha r)^2} + 4\pi C^2 \int_{a_0}^{\infty} r^2 dr \frac{\exp(-2\beta r)}{\beta^2 r^2}$$
$$= \frac{2\pi A^2}{\alpha^2} a_0 + \frac{2\pi C^2 \exp(-2\beta a_0)}{\beta^3} \qquad \text{(C- 11)}$$

Since $\alpha \approx \beta$ and $a_0 >> \frac{1}{\alpha}$ i.e. $a_0 \alpha >> 1$,

$$1 \approx \frac{2\pi A^2}{\alpha^2} a_0 \quad \text{or} \quad A^2 \approx \frac{\alpha^2}{2\pi a_0} \qquad \text{(C- 12)}$$

$$\psi = \begin{cases} \frac{\alpha}{\sqrt{2\pi a_0}} \frac{\sin(\alpha r)}{\alpha r} & for \quad r < a_0 \\ \frac{\alpha}{\sqrt{2\pi a_0}} \exp(\beta a_0) \frac{\exp(-\beta r)}{\beta r} & for \quad r > a_0 \end{cases} \qquad \text{(C- 13)}$$

We can relate the wave function to the volume of the tip modeled as a sphere, since:

$$V_t = \frac{4}{3} \pi a_0^{\,3}$$

The wave function is then given by equation (C- 14) , and if we take $\frac{\alpha}{\beta}\sqrt{2/3} = c_t$, equation (C- 14) becomes (C- 15)

$$\psi = \frac{\alpha a_0}{\sqrt{V_t}} \exp(\beta a_0) \frac{\exp(-\beta r)}{\beta r} \sqrt{\frac{2}{3}} \qquad \text{(C- 14)}$$

$$\psi = \frac{a_0 c_t}{\sqrt{V_t}} \exp(\beta a_0) \frac{\exp(-\beta r)}{r} \qquad \text{(C- 15)}$$

Appendix D

Theoretical calculation of piezo tube resonance frequencies in longitudinal and bending modes

1) Longitudinal mode of the scan piezo tube

We can start from the basic wave equation [56]:

$$\frac{\partial^2 \xi}{\partial x^2} = \frac{1}{c^2}\frac{\partial^2 \xi}{\partial t^2} \qquad \textbf{(} \text{D- 1} \textbf{)}$$

where $\xi \equiv \xi(x,t)$ is the longitudinal displacement and $c = \sqrt{Y/\rho}$ is the phase speed. Y is Young's modulus of the piezo tube, and ρ is the piezo tube density.

The complex harmonic solution of equation (D- 1) is:

$$\xi = Ae^{j(\omega t - kx)} + Be^{j(\omega t + kx)} \qquad (\text{D- 2})$$

where k is the wave vector and ω is the angular frequency.

We assume the scan piezo tube to be clamped at one end labelled by $x = 0$ and loaded with a concentrated mass of the tip holder at the other end $x = L$. The boundary conditions take into account the fact that no displacement happens at the clamped end while at the mass loaded end, Newton's second law can be applied. Therefore, the following two equations **(** D- 3 **)** and **(** D- 4 **)** rise from the two boundaries:

$$\text{At } x = 0 \text{ (clamped end)} \qquad \xi = 0 \qquad (\text{D- 3})$$

$$\text{At } x = L \text{ (mass loaded end)} \qquad (f)_{x=L} = m_{load}\left(\frac{\partial^2 \xi}{\partial t^2}\right)_{x=L} \qquad (\text{D- 4})$$

L is the length of the piezo tube. Solving equation (D- 2) together with boundary condition equations gives us:

$$\cot(kL) = \frac{m_{load}}{m_b} kL \qquad \text{(D- 5)}$$

Table D-1: The parameters of the scan piezo tube we use

Young's Modulus $\left(10^{10} N/m^2\right)$	Density $\left(g/cm^3\right)$	Inner Diameter (cm)	Outer Diameter (cm)	Length (cm)	Mass (g)
6.1	7.65	0.5	0.625	1.27	1.17

The mass of tip holder m_{load} is estimated to be 0.656 (g). Therefore, the resonant frequencies of scan piezo tube in longitudinal mode can be estimated using equation (D- 5).

Table D-2: The resonant frequencies of scan piezo tube in longitudinal mode

Series	1st	2nd	3rd	4th
Frequency(Hz)	35928.9	126373.0	231034.0	339496.0

2) The bending mode

The general equation of motion in bending mode is given by:

$$\frac{\partial^2 y}{\partial t^2} = -\kappa^2 c^2 \frac{\partial^4 y}{\partial x^4} \qquad \textbf{(} \text{D- 6} \textbf{)}$$

where $c = \sqrt{Y/\rho}$ is the phase speed as for longitudinal waves. The constant κ can be thought of as the radius of gyration of a solid. For a hollow cylindrical solid bar with inner radius r_{in} and outer radius r_{out}, its radius of gyration is given by: $\kappa = \sqrt{\left(r_{in}^{\ 2} + r_{out}^{\ 2}\right)/2}$

The complete solution to equation (D- 6) is:

$$y = e^{j\omega t}\left(A^{'}e^{\omega x/v} + B^{'}e^{-(\omega x/v)} + C^{'}e^{j\omega x/v} + D^{'}e^{-j\omega x/v} \right) \qquad \textbf{(} \text{D- 7} \textbf{)}$$

where $v = \sqrt{\omega c \kappa}$. The actual solution is [56]:

$$y = \cos(\omega t + \phi)\left[A\cosh(\omega x/v) + B\sinh(\omega x/v) + C\cos(\omega x/v) + D\sin(\omega x/v) \right] \qquad \textbf{(} \text{D- 8} \textbf{)}$$

i) Clamped at one end, and free at the other end

If the end ($x = 0$) of the bar is rigidly clamped, both the displacement and the slope must be zero at this end at all times. The boundary conditions are therefore

$$y = 0 \text{ and } \frac{\partial y}{\partial x} = 0 \qquad \textbf{(} \text{D- 9} \textbf{)}$$

At a free end ($x = L$), there can be neither an externally applied torque nor a shearing force, and hence both are zero, at $x = L$.

$$\frac{\partial^2 y}{\partial x^2} = 0 \text{ and } \frac{\partial^3 y}{\partial x^3} = 0 \qquad \textbf{(} \text{D- 10} \textbf{)}$$

The solution of (D- 8) turns out to be:

$$\cot\left(\frac{\omega L}{2v} \right) = \pm \tanh\left(\frac{\omega L}{2v} \right) \qquad \textbf{(} \text{D- 11} \textbf{)}$$

The results of resonant frequencies of scan piezo tube in bending mode determined from equation.(D- 11) are:

Table D-3: The resonant frequencies of scan piezo tube in bending mode

Series	1st	2nd	3rd	4th
Frequency (Hz)	27723.9	173743.0	486485.0	953316.0

ii) Clamped at one end, and mass loaded at the other end.

The boundary condition at the clamped end ($x = 0$) is still the same as shown in equation (D- 9). At the mass-loaded end ($x = L$), both torque and force are not zero any more.

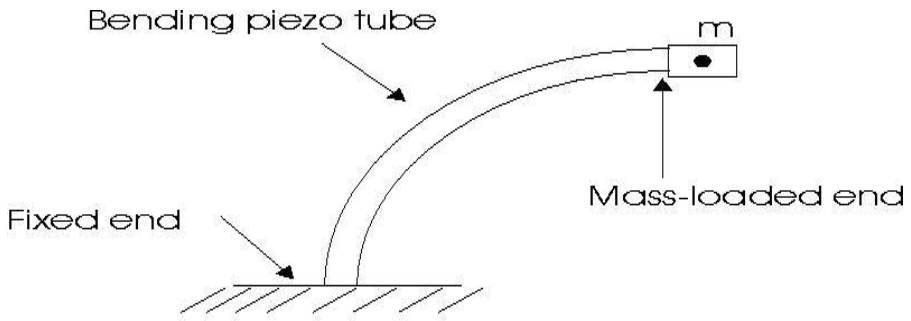


Figure D-1: The piezo tube is bended with mass load**.**

In the picture above, the torque is equal to the multiplication of moment of inertia and angular acceleration. We simply assume the moment of inertia can be represented by the multiplication of mass and the square of a distance a. The angle is related to both $\frac{\partial y}{\partial x}$ and angular acceleration. The force is related to $\frac{\partial^2 y}{\partial x^2}$, according to Newton's second law.

Two of newly made tip-holders plus tip together weigh 0.355(g). So the tip holder weighs (6-1 about 0.178 (g).

Table D-4: The resonant frequencies of scan piezo tube and coarse approach piezo tube in the longitudinal mode**.**

Longitudinal Mode		1st	2nd	3rd	4th
Scan Piezo Tube	Tip-off	55586.6 (Hz)	166760 (Hz)	277933 (Hz)	389107 (Hz)
	Tip-on	47788.1 (Hz)	145570 (Hz)	247519 (Hz)	352767 (Hz)
Coarse Approach Piezo Tube	Tip-off	10381.2 (Hz)	118591 (Hz)	235786 (Hz)	353288 (Hz)
	Tip-on	10206.3 (Hz)	118559 (Hz)	235770 (Hz)	353278 (Hz)

Table D-5**:** The resonant frequencies of scan piezo tube and coarse approach piezo tube in the bending mode**.**

Bending Mode		1st	2nd	3rd	4th
Scan Piezo Tube	Tip-off	27723.9 (Hz)	173743 (Hz)	486485 (Hz)	953316 (Hz)
	Tip-on	21224.4 (Hz)	128029 (Hz)	307933 (Hz)	589399 (Hz)
Coarse Approach Piezo Tube	Tip-off	1580.30 (Hz)	3044.19 (Hz)	6766.95 (Hz)	835796 (Hz)
	Tip-on	1566.48 (Hz)	3016.60 (Hz)	6708.20 (Hz)	83566 (Hz)

For coarse approach piezo tube, a is estimated to be 4.16(mm). For scan piezo tube, a is estimated to be 2.0(mm).

Appendix E
Solving the cubic equation

The main goal is to solve the equation (E-1)**:**

$$\omega^3 - j\frac{\omega^2\omega_p}{Q} - \omega{\omega_p}^2 + \frac{j\omega_p^2}{\tau} = 0 \qquad \text{(E-1)}$$

To solve the equation, we first rewrite it in a format as follows:

$$x^3 + px + q = 0 \qquad \text{(E-2)}$$

This means that we need to avoid any second order terms in the equation. It can be realized by introducing a new variable x and defining $\omega = x + j\frac{\omega_p}{3Q}$. A new equation with variable x is obtained:

$$x^3 - \omega_p^2\left(1 - \frac{1}{3Q^2}\right)x + j\omega_p^3\left[\frac{1}{\omega_p\tau} + \frac{2}{27Q^3} - \frac{1}{3Q}\right] = 0 \qquad \text{(E-3)}$$

As expected, equation (E-3)does have second order term. Comparing equation (E-3) to (E-2) gives us:

$$p = -\omega_p^2\left(1 - \frac{1}{3Q^2}\right) \text{ and } q = j\omega_p^3\left[\frac{1}{\omega_p\tau} + \frac{2}{27Q^3} - \frac{1}{3Q}\right] \qquad \text{(E-4)}$$

The problem becomes now to find the general solution to equation (E-2).

We define $u + v = x$ and $uv = -\frac{p}{3}$, it is easy to get:

$$u^3 + v^3 = (u+v)^3 - 3uv(u+v) = x^3 + px = -q \qquad \text{(E-5)}$$

$$(uv)^3 = -\frac{p^3}{27} \qquad \text{(E-6)}$$

Equations (E-5) and (E-6) can lead to:

$$(t-u^3)(t-v^3) = t^2 - t(u^3+v^3) + (uv)^3 = t^2 + qt - \frac{p^3}{27} \qquad \text{(E-7)}$$

Apparently u^3 and v^3 are the roots of $(t-u^3)(t-v^3) = t^2 + qt - \frac{p^3}{27} = 0$, which gives two roots of

$$t_{\pm} = -\frac{q}{2} \pm \sqrt{\frac{q^2}{4} + \frac{p^3}{27}} \qquad \text{(E-8)}$$

Particularly in our case, Q is not very small, and so according to equation (E-4), $q^2 < 0$ and $p^3 < 0$. Thus, in equation (E-8), the radical is negative and therefore $t_{\pm}$ are imaginary

Call $u^3 = t_+ = |t_+| e^{j\frac{\pi}{2}}$ and $v^3 = t_- = |t_-| e^{-j\frac{\pi}{2}}$

Define $u_1 = |t_+|^{\frac{1}{3}} e^{j\frac{\pi}{6}}$ $u_2 = u_1 e^{j\frac{2\pi}{3}}$ and $u_3 = u_1 e^{-j\frac{2\pi}{3}}$

$v_1 = |t_-|^{\frac{1}{3}} e^{-j\frac{\pi}{6}}$ $v_2 = v_1 e^{-j\frac{2\pi}{3}}$ $v_3 = v_1 e^{j\frac{2\pi}{3}}$

To satisfy the condition of $u_1 v_1 = u_2 v_2 = u_3 v_3 = -\frac{p}{3}$. The vectors are shown in the complex plane in Figure E-1.

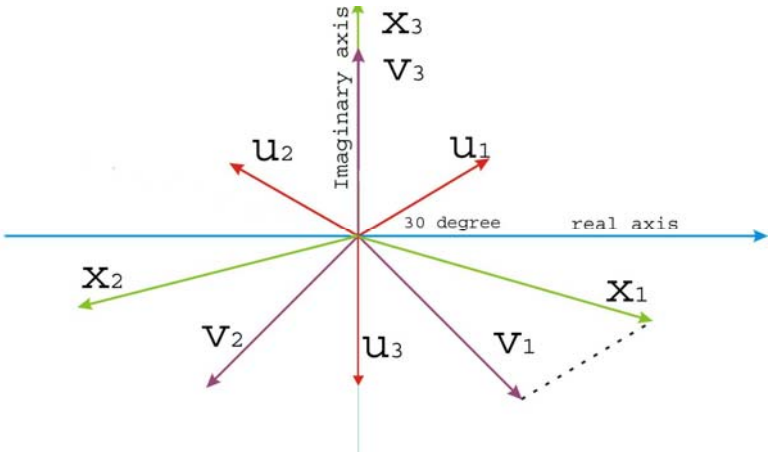


Figure E-1: The vectors of $\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{v}_1, \vec{v}_2, \vec{v}_3$ are in the complex plane**.**

(1) when $Q >> 1$ and $Q >> \omega_p \tau$

$$\omega_1 = x_1 + j\frac{\omega_p}{3Q} \approx x_1$$

Then equation (E-4) becomes:

$$p = -\omega_p^2 \text{ and } q = j\frac{\omega_p^2}{\tau} \qquad \text{(E-9)}$$

Therefore, the radical in equation (E-8) is:

$$\frac{q^2}{4} + \frac{p^3}{27} = -\omega_p^6 \left[\frac{1}{\left(2\omega_p \tau\right)^2} + \frac{1}{27} \right] = -\alpha^2 \omega_p^6 \qquad \text{(E-10)}$$

where $\alpha^2 = \dfrac{1}{\left(2\omega_p \tau\right)^2} + \dfrac{1}{27} > \dfrac{1}{\left(2\omega_p \tau\right)^2} \rightarrow |a| > \dfrac{1}{2\omega_p \tau}$

so equation (E-8) turns into :

$$t_{\pm} = -\frac{j\omega_p^3}{2\omega_p \tau} \pm j\alpha\omega_p^3 \qquad \text{(E-11)}$$

Now consider the subcase

i) $\omega_p\tau >> 1$ (low gain)

we have $\alpha^2 = \dfrac{1}{\left(2\omega_p\tau\right)^2} + \dfrac{1}{27} > \dfrac{1}{\left(2\omega_p\tau\right)^2} \rightarrow |a| > \dfrac{1}{2\omega_p\tau}$, $\alpha \approx \dfrac{1}{\sqrt{27}}$, and

$$|t_{\pm}| = \alpha\omega_p^3 \mp \frac{\omega_p^3}{2\omega_p\tau} = \left(\frac{\omega_p}{\sqrt{3}}\right)^3 \left[1 - \frac{3^{3/2}}{2\omega_p\tau}\right] \qquad \text{(E-12)}$$

Therefore, all the values of the variables are found:

$$u^3 = t_+ = |t_+| e^{j\frac{\pi}{2}} \Rightarrow u_1 = \frac{\omega_p}{\sqrt{3}} \left[1 - \frac{3^{3/2}}{3\left(2\omega_p\tau\right)}\right] e^{j\frac{\pi}{6}}$$

$$v^3 = t_- = |t_-| e^{-j\frac{\pi}{2}} \Rightarrow v_1 = \frac{\omega_p}{\sqrt{3}} \left[1 + \frac{3^{3/2}}{3\left(2\omega_p\tau\right)}\right] e^{-j\frac{\pi}{6}}$$

The solutions of equation (E-1) are approximately:

$$\omega_1 \approx x_1 = u_1 + v_1 = \omega_p \left(1 - \frac{j}{2\omega_p\tau}\right) \qquad \text{(E-13)}$$

$$\omega_2 \approx x_2 = u_2 + v_2 = u_1 e^{j\frac{2\pi}{3}} + v_1 e^{-j\frac{2\pi}{3}} = -\omega_p - \frac{j\omega_p}{2\omega_p\tau} \qquad \text{(E-14)}$$

$$\omega_3 \approx x_3 = u_3 + v_3 = u_1 e^{-j\frac{2\pi}{3}} + v_1 e^{j\frac{2\pi}{3}} = j\frac{\omega_p}{\omega_p\tau} \qquad \text{(E-15)}$$

This vector is very small, with no real component, and is above the x-axis. This is the only pole above the real axis.

ii) $Q >> 1 >> \omega_p\tau$ (High gain)

$$\alpha^2 = \frac{1}{\left(2\omega_p\tau\right)^2} + \frac{1}{27} = \frac{1}{\left(2\omega_p\tau\right)^2}\left[1 + \frac{\left(2\omega_p\tau\right)^2}{27}\right] \text{ and }, \alpha \approx \frac{1}{2\omega_p\tau}\left[1 + \frac{2\left(\omega_p\tau\right)^2}{27}\right]$$

Thus, $t_+ = -\frac{j\omega_p^3}{2\omega_p\tau} + j\omega_p^3 \frac{1}{2\omega_p\tau}\left[1+\frac{2\left(\omega_p\tau\right)^2}{27}\right] = j\omega_p^3\frac{\omega_p\tau}{27}$ (A very small positive imaginary)

$t_- = -\frac{j\omega_p^3}{2\omega_p\tau} - j\omega_p^3 \frac{1}{2\omega_p\tau}\left[1+\frac{2\left(\omega_p\tau\right)^2}{27}\right] \approx -\frac{j\omega_p^3}{\omega_p\tau}$ (Large, negative imaginary)

$u_1 = \frac{\omega_p}{3}\left(\omega_p\tau\right)^{1/3} e^{j\frac{\pi}{6}}$ and $v_1 \approx \frac{\omega_p}{\left(\omega_p\tau\right)^{1/3}} e^{-j\frac{\pi}{6}}$

The solutions to equation (E-1) are :

$$\omega_1 \approx x_1 \approx v_1 = \frac{\omega_p}{\left(\omega_p\tau\right)^{1/3}} e^{-j\frac{\pi}{6}} \qquad \text{(E-16)}$$

$$\omega_2 \approx x_2 \approx v_2 = \frac{\omega_p}{\left(\omega_p\tau\right)^{1/3}} e^{-j\frac{\pi}{6}} e^{-j\frac{2\pi}{3}} = \frac{\omega_p}{\left(\omega_p\tau\right)^{1/3}}\left(-\frac{\sqrt{3}}{2} - j\frac{1}{2}\right) \qquad \text{(E-17)}$$

$$\omega_3 \approx x_3 \approx v_3 = \frac{\omega_p}{\left(\omega_p\tau\right)^{1/3}} j \qquad \text{(E-18)}$$

This vector is above the real axis.

(2) $\omega_p\tau >> Q >> 1$

Equation (E-4) becomes:

$$p = -\omega_p^2\left(1-\frac{1}{3Q^2}\right) \Rightarrow p \approx -\omega_p^2$$
$$q = j\omega_p^3\left[\frac{1}{\omega_p\tau} + \frac{2}{27Q^3} - \frac{1}{3Q}\right] \approx -\frac{j\omega_p^3}{3Q} \qquad \text{(E-19)}$$

We can rewrite equation (E-8):

$$t_\pm = -\frac{q}{2} \pm \sqrt{\frac{q^2}{4}+\frac{p^3}{27}} = -j\frac{\omega_p^3}{6Q} \pm \sqrt{-\frac{\omega_p^6}{9Q^2} - \frac{\omega_p^6}{27}} \approx \pm j\frac{\omega_p^3}{\sqrt{27}} \qquad \text{(E-20)}$$

Call $u^3 = t_+ = |t_+| e^{j\frac{\pi}{2}}$ and $v^3 = t_- = |t_-| e^{-j\frac{\pi}{2}}$

Define $u_1 = |t_+|^{\frac{1}{3}} e^{j\frac{\pi}{6}} = \frac{\omega_p}{\sqrt{3}} e^{j\frac{\pi}{6}}$ $u_2 = u_1 e^{j\frac{2\pi}{3}}$ and $u_3 = u_1 e^{-j\frac{2\pi}{3}}$

$$v_1 = |t_-|^{\frac{1}{3}} e^{-j\frac{\pi}{6}} = \frac{\omega_p}{\sqrt{3}} e^{-j\frac{\pi}{6}} \quad v_2 = v_1 e^{-j\frac{2\pi}{3}} \quad v_3 = v_1 e^{j\frac{2\pi}{3}}$$

The solutions of equation (E-1) are :

$$\omega_1 = x_1 + j\frac{\omega_p}{3Q} = \omega_p\left(1 + \frac{j}{3Q}\right) \qquad \text{(E-21)}$$

$$\omega_2 = x_2 + j\frac{\omega_p}{3Q} = \omega_p\left(-1 + \frac{j}{3Q}\right) \qquad \text{(E-22)}$$

$$\omega_3 = x_3 + j\frac{\omega_p}{3Q} = \frac{j\omega_p}{3Q} \qquad \text{(E-23)}$$

Appendix F
Optimizing the feedback model

The following discussion on the modeling of the circuit is divided into two parts. In the first, we consider the condition that the tunneling current does not flow between the sample and the tip when the tip-sample separation is not sufficiently small. As the separation shrinks, the tunneling current flows and a resistance of the tip-sample junction has to be considered in the circuit as will be discussed in the second part.

1) no tunneling current

Capacitive coupling C_j between tip and sample and stray capacitance C_p over the pre-amp resistance are considered as shown in Figure F-1. Since there is no tunneling current, the resistance of the tip-sample junction is infinity and is not included in the model. We assume the magnitude and frequency of the ac signal b(t) to be b_0 and ω_0 respectively. In time-domain, $b(t) = b_0 e^{j\omega_0 t}$ and in frequency domain, $B(\omega) = b_0 \delta(\omega - \omega_0)$.

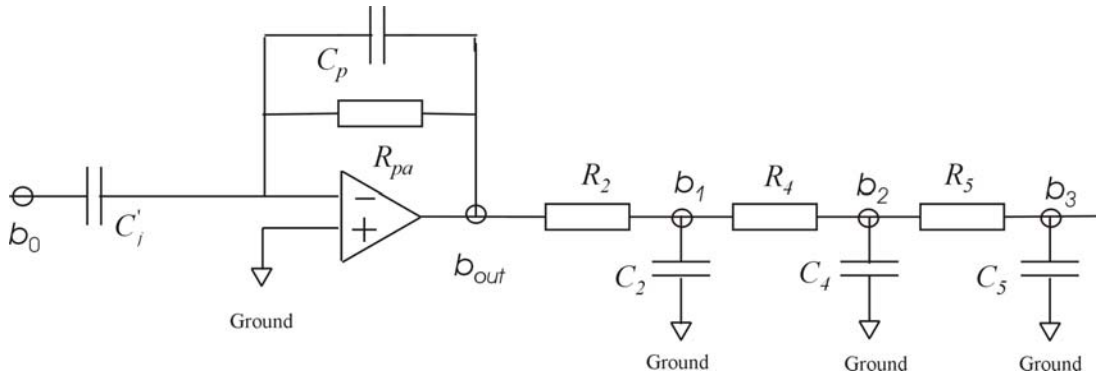


Figure F-1: The modeling of the circuit involving in spectroscopic imaging. The capacitive coupling C_j is the only interaction between the tip and the sample under the condition of no tunneling between them

In the frequency domain, we calculate in each stage the ratio of the output voltage signal to the input according to the analog circuit:

$$\frac{b_{out}}{b_0} = -\frac{j\omega_0 C_j R_{pa}}{(1 + j\omega_0 R_{pa} C_p)} \qquad (F-1)$$

$$\frac{b_1}{b_{out}} = \frac{1}{1 + j\omega_0 R_2 C_2} \qquad (F-2)$$

$$\frac{b_2}{b_1} = \frac{1}{1 + j\omega_0 R_4 C_4} \qquad (F-3)$$

$$\frac{b_3}{b_2} = \frac{1}{1 + j\omega_0 R_5 C_5} \qquad (F-4)$$

From equation (F- 1) to (F- 4), we can have

$$b_3 = \frac{-jb_0\omega_0 R_{pa} C_j}{(1 + j\omega_0 R_j C_p)(1 + j\omega_0 R_2 C_2)(1 + j\omega_0 R_4 C_4)(1 + j\omega_0 R_5 C_5)} \qquad (F-5)$$

In practice, the lock-in amplifier is capable of measuring the modulus of b_3 as well as the phase difference between b_3 and b_0 through the R output channel and θ output channel respectively. In our measurements, the magnitude b_0 is set to be 10 mV.

$$\mathrm{Re}(b_3) = R\cos\theta \qquad (F-6)$$

$$\mathrm{Im}(b_3) = R\sin\theta \qquad (F-7)$$

According to equation (F- 5), the real part and imaginary part of b_3 are given theoretically by:

$$\mathrm{Re}(b_3) = \frac{-b_0\omega_0^2 R_{pa} C_j \left[(1-\omega_0^2 R_{pa}R_2C_2C_p)(R_4C_4+R_5C_5)+(R_{pa}C_p+R_2C_2)(1-\omega_0^2 R_4R_5C_4C_5)\right]}{(1+\omega_0^2 R_5^{\,2}C_5^{\,2})(1+\omega_0^2 R_4^{\,2}C_4^{\,2})(1+\omega_0^2 R_{pa}^{\,2}C_p^{\,2})(1+\omega_0^2 R_2^{\,2}C_2^{\,2})} \qquad (\text{F-}8)$$

$$\mathrm{Im}(b_3) = \frac{-b_0\omega_0 R_{pa} C_j \left[(1-\omega_0^2 R_{pa}R_2C_2C_p)(1-\omega_0^2 R_4R_5C_4C_5)-\omega_0^2(R_4C_4+R_5C_5)(R_{pa}C_p+R_2C_2)\right]}{(1+\omega_0^2 R_5^{\,2}C_5^{\,2})(1+\omega_0^2 R_4^{\,2}C_4^{\,2})(1+\omega_0^2 R_{pa}^{\,2}C_p^{\,2})(1+\omega_0^2 R_2^{\,2}C_2^{\,2})} \qquad (\text{F-}9)$$

Now we may use the equations (F- 8) and (F- 9) to make curve fittings on experimental data of the real part (X) as well as the imaginary part (Y) respectively to get the stray capacitance C_p and tip-sample capacitive coupling C_j.

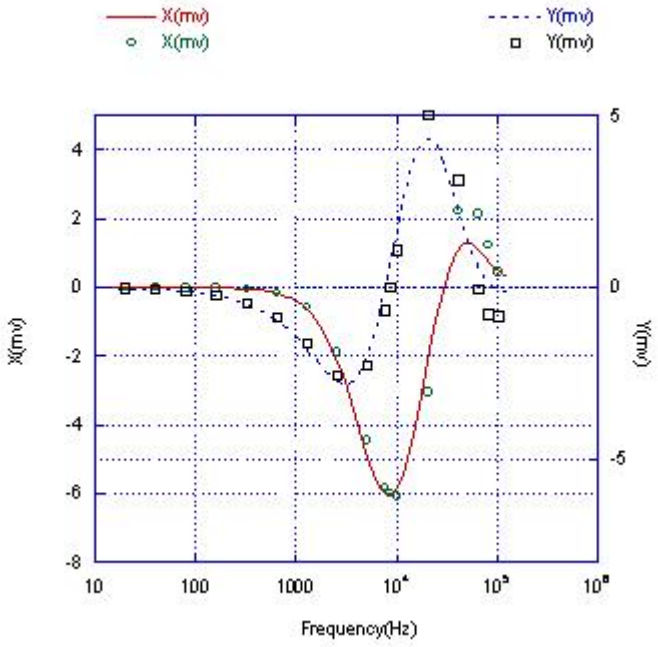


Figure F-2: Curve fittings (solid line for X and dashed line for Y) on experimental data (dots for X, square for Y) obtained at 0 nA tunneling current to get the capacitance between the tip and the sample. The solid line matches very well the dots and the dashed line matches very well with the squares

The curve fitting in Figure F-2 was very successful. In a broad range of frequency, the two curves match reasonably well with the experimental data. Each curve

gives a set of values for C_j and C_p. The average value of C_j is $0.242\,pF$,and $0.06\,pF$ for C_p.

2) tunneling

We also model a capacitive and resistive coupling between tip and sample for the condition of tunneling. The tip-sample separation is reduced and the capacitive coupling between the tip and the sample is expected to be stronger. The value for C_j is expected to be greater.

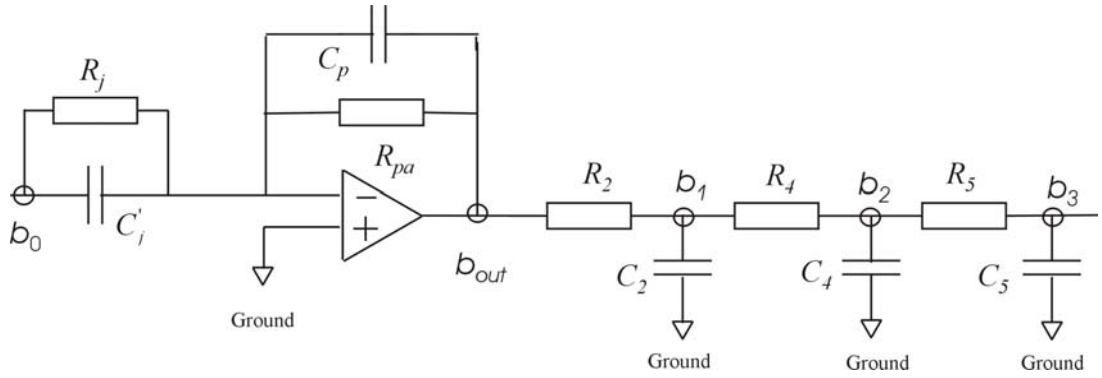


Figure F-3: The modeling of the circuit involving in spectroscopic imaging. Besides the capacitive coupling C_j, the resistance R_J is not infinity any more when the tunneling current flows between the tip and the sample and has to be considered.

Equation (F- 5) becomes:

$$b_3 = -\frac{b_0 R_{pa}(1+j\omega_0 R_J C_j')}{R_J(1+j\omega_0 R_5 C_5)(1+j\omega_0 R_4 C_4)(1+j\omega_0 R_2 C_2)(1+j\omega_0 R_{pa} C_p)} \qquad (\text{F- } 10)$$

The modulus R of b_3 is :

$$R = \frac{b_0 R_{pa}\sqrt{1+\omega_0^2 R_J{}^2 C_j'{}^2}}{R_j\sqrt{1+\omega_0^2 R_5{}^2 C_5{}^2}\sqrt{1+\omega_0^2 R_4{}^2 C_4{}^2}\sqrt{1+\omega_0^2 R_2{}^2 C_2{}^2}\sqrt{1+\omega_0^2 R_{pa}{}^2 C_p{}^2}} \qquad (\text{F- } 11)$$

This quantity can be measured through the R output channel of the lock-in amplifier. The resistance of the tunneling junction is R_J =7.5 GΩ since the sample bias and tunneling current is set to be –1.5(V) and 0.2(nA) respectively. We use equation (F- 11) to make a curve fit over experimental data to obtain the C_j

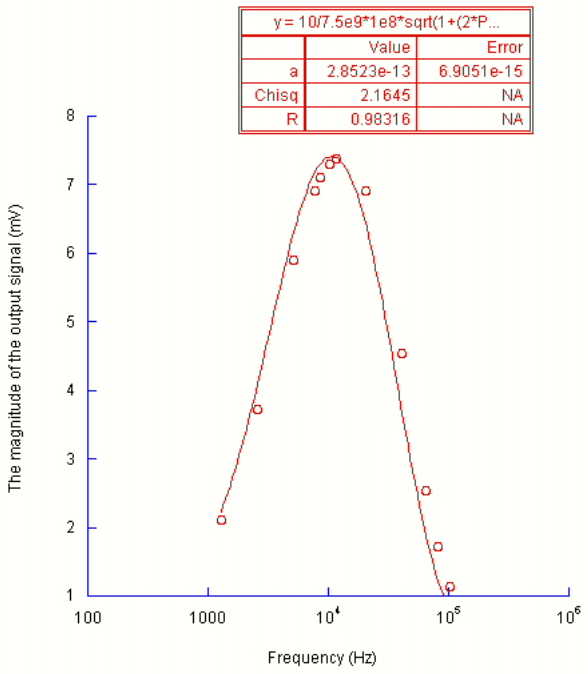


Figure F-4: Curve fitting (solid line) on experimental data (dot markers) measured with a tunneling current 0.2 nA at sample bias of -1.5 V to get the capacitance between tip and sample.

The coupling capacitance $C_j^{'}$ is $0.285\,pF$,1.18 times greater than $0.242\,pF$ obtained under the condition of when no tunneling current happens. This result

agrees with the fact that the capacitive coupling between tip and sample is stronger as the distance is smaller when the tunneling phenomena occurs.

Appendix G

Discussion of $dI/dV*V/I$

In this section, we first illustrate step by step in detail how the $dI/dV*V/I$ and LDOS are related particularly in this work. We start from the tunneling expression in equation (3-31). For illustrative purposes, we write an approximate expression for equation (3-31).

$$I \propto \int_0^{eV} g(E)\,T'(E,eV)\,dE \qquad \text{(G- 1)}$$

where $g(E)$ is the surface DOS of the sample, and T' is the transmission probability of the electron.

T' is brought into the equation to account for the effect of the electric field on tails of the wave functions $\phi_i(\vec{r}_0 - \vec{R})$. This effect has been discussed by using the WKB method in Chapter 4. In the simple model, the transmission probability T' is a function of the energy of the electron and the voltage applied across the tunneling junction.

$$T' = T'(E,eV) = \exp\left(\frac{2d_0}{3m\hbar V}(T_1 - T_2)\right) \qquad \text{(G- 2)}$$

where $T_1 = \left[2m(V-E)\right]^{3/2}$ and $T_2 = \left[2m(-E)\right]^{3/2}$, and d_0 is the distance between the tip and the origin of the atomic orbital described by the wave function. Taking a derivative on both sides of equation (G- 1) leads us to :

$$dI/dV \propto eg(eV)T'(eV,eV) + e\int_0^{eV} g(E)\frac{d}{d(eV)}\left[T'(E,eV)\right]dE \qquad \text{(G- 3)}$$

Combining equation (G- 1) and (G- 3) gives us:

$$\frac{dI/dV}{I/V} = \frac{g(eV) + \int_0^{eV} \frac{g(E)}{T'(eV,eV)} \frac{d}{d(eV)} \left[T'(E,eV) \right] dE}{\frac{1}{eV} \int_0^{eV} \frac{g(E) T'(E,eV)}{T'(eV,eV)} dE} \qquad \text{(G- 4)}$$

The first term $g(eV)$ in the numerator is the surface DOS of the sample at an energy of eV, which is relative to the Fermi level. The denominator can be thought of as providing a "normalization" of the surface DOS. In the expression, $T'(E,eV)$ and $T'(eV,eV)$ tend to cancel out each other. This prevents the spectra from diverging exponentially both in voltage and in separation.