Comparison of electronic and mechanical contrast in scanning tunneling microscopy images of semiconductor heterojunctions

Randall M. Feenstra

Carnegie Mellon University, feenstra@andrew.cmu.edu

Follow this and additional works at: http://repository.cmu.edu/physics

Published In
Physica B: Condensed Matter, 273-274, 796-802.
Comparison of Electronic and Mechanical Contrast in Scanning Tunneling Microscopy Images of Semiconductor Heterojunctions

R. M. Feenstra
Department of Physics, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213 USA

Abstract
The use of cross-sectional scanning tunneling microscopy (STM) to study strain in semiconductor heterostructures is discussed. In particular, intermixing between constituent heterostructure layers leads to internal strains in the heterostructure, and these strained regions are evident by displacement of the cleavage surface formed in the STM study. A theoretical analysis is made of the magnitude of electronic compared to mechanical contributions to the contrast of STM images, from which it is found that the former are relatively small, on the order of 0.1 Å, for typical In$_x$Ga$_{1-x}$As$_y$P$_{1-y}$ heterostructures imaged with sufficiently large, positive sample bias.

I Introduction
Within the area of semiconductor film growth, scanning tunneling microscopy (STM) has been used over the past fifteen years to reveal many important aspects of atomic-scale structures and growth processes. A distinction can be made between cross-sectional imaging and plan-view imaging: the former is performed on (110) surfaces prepared by in situ cleavage and which are not reconstructed, whereas the latter is performed typically on (001) growth surfaces which are heavily reconstructed. Thus, in plan-view imaging, the information obtained pertains largely to the surface reconstruction itself, whereas for cross-sectional imaging one can study properties of the material which are not surface specific, e.g. concentration and types of defects in the semiconductor. For the case of heterostructures, cross-sectional imaging is useful for probing the structure of the interfaces between neighboring layers, as well as for studying the detailed properties of the individual layers themselves [1-4].

An example of point defect studies using cross-sectional STM (xSTM) is shown in Fig. 1 [5]. Those images were obtained from In$_{0.53}$Ga$_{0.47}$As layers grown at a relatively low growth temperature of 240° C. Such material is known to contain excess arsenic, which predominantly forms antisite defects – arsenic on a gallium or indium site. The constant-current image of Fig. 1(a) shows one such defect, located near the center of the image. The same defect is also seen in the conductance image of Fig. 1(b), acquired simultaneously with Fig. 1(a). Detailed studies of As antisite defects in GaAs reveals that the antisite defects on different layers relative to the (110) cleavage plane can be distinguished, and spectroscopic studies reveal the gap states introduced by the antisites [6]. For the case of the In$_{0.53}$Ga$_{0.47}$As investigation shown in Fig. 1, observation of the antisite defects in material of different doping concentrations and with different post-growth annealing conditions permitted an understanding of the electron-hole recombination dynamics on the ps time-scale [5]. Many other point defect studies have been performed in recent years by Ebert and co-workers [7].
In addition to studies of point defects, xSTM has been very useful in probing properties of semiconductor heterostructures. An example is given in Fig. 2, which shows an In_{0.53}Ga_{0.47}As quantum well imbedded between InP barrier layers. Figure 2(a) shows the as-grown heterostructure, and Fig. 2(b) shows the structure after ion implantation and subsequent annealing [8]. The ion implantation causes intermixing between the quantum well and the barrier layers. This intermixing, in turn, produces a blue shift in the optical transition energy of the quantum well, and thus the quantum well intermixing process can be used for tuning the emission wavelength of lasers and for other photonic integrated device applications [9]. The intermixing between quantum well and barrier layers is clearly evident in Fig. 2(b): we see some white (higher tip height) bands forming at the well/barrier interfaces, and the quantum well itself is darker (lower tip height) in Fig. 2(b) compared to Fig. 2(a). For the particular choice of implantation parameters used in this case [8] we have apparently produced a thin region near the well/barrier interface with new chemical composition which gives the white contrast in the STM images, along with modifying the overall quantum well composition thereby producing its darker contrast.

Given STM data of the type shown in Fig. 2, it is desirable to obtain a quantitative evaluation of chemical composition as a function of position in the heterostructure. Such an evaluation is, of course, quite difficult in general since the contrast in STM images contains significant contributions from electronic effects in the tunnel current and such effects are not easily quantified. However, in a series of studies of strained heterostructures, we have recently observed a second significant contribution to STM images, arising from displacement of the (110) cleavage face due to strain in the underlying material [10-12]. In cases where electronic effects can be shown to be small, this mechanical or elastic contribution to the STM contrast can be relatively simply evaluated using finite-element solutions of the elasticity equations, thus providing a means of determining strain and associated chemical composition in the heterostructure.

Figure 3 illustrates these electronic and mechanical contributions to the STM contrast for the case of a semiconductor superlattice with compressively strained barriers and tensilely strained quantum well layers. In this particular case, the respective contributions have opposite sign, although for compressive wells with tensile barriers the electronic and mechanical effects would have the same sign. Let us consider the magnitude of the mechanical effect. As noted in the caption of Fig. 3, the peak-to-peak corrugation amplitude is approximately equal to $2\varepsilon L$ where $L$ is the width of barrier and well, and their strains (in-plane, diagonal component) are $\pm\varepsilon$. For example, with 5 nm thick well and barrier, having strains of $\pm0.01$, the amplitude would be 1 Å. Offhand, this result is of the same order as typical electronic effects in the tunnel current between different materials, so it may seem difficult to distinguish between electronic and mechanical effects.

The main purpose of this paper is to provide a theoretical analysis of the tunnel current for In_{x}Ga_{1-x}As_{y}P_{1-y} heterostructures, with the goal of placing upper limits on the magnitude of electronic effects which may occur in the STM measurements. We demonstrate that for appropriate tunneling conditions (i.e. large, positive sample bias), the electronic effects are in fact remarkably small, on the order of 0.1 Å for typical heterostructures. Hence, it is possible to interpret many xSTM images of strained In_{x}Ga_{1-x}As_{y}P_{1-y} heterostructures directly in terms of the elastic displacement of the cleavage surface, thereby yielding a relatively direct measurement of the strain vs. position in the heterostructure.
II Results

The situation for which we compute the tunnel current is illustrated in Fig. 4. We consider two regions of neighboring semiconductor materials which are separated by an ideal interface. The Fermi levels in the two materials are aligned. An important starting point in our analysis is the use of the “electron affinity rule” which states that the conduction band offset between neighboring materials is determined by the difference in their electron affinities (distance from vacuum level to conduction band minimum). Dipoles forming at the interface may produce deviations in such offsets [13], but these effects appear to be small at least for the case of InGaAs on InP [14]. For the materials GaAs, InP, InAs, and GaP, the electron affinities are given by 4.13, 4.50, 5.06, and 3.75 eV respectively, with the latter value referring to the X-valley conduction band minimum of GaP [15]. Thus, offsets of the conduction band of GaAs, InP, and GaP relative to InAs are 0.93, 0.56, and 1.31 eV respectively. In a heterostructure, the vacuum levels of the neighboring materials will be aligned as pictured in Fig. 4. The separation between vacuum level and Fermi level is the work function of the sample, \( \phi \), which is identical for all the materials we consider here due to our use of the electron affinity rule. We take the Fermi level to be located at the InAs CB minimum, so that the work function for all materials is 5.06 eV.

The goal of our study is to evaluate the material dependence of the tunnel current in In\(_x\)Ga\(_{1-x}\)As\(_y\)P\(_{1-y}\) alloys. In particular, if one moves across a heterointerface from one material to the next, by how much will the tunnel current change? We are interested in tunnel current variations in the immediate vicinity of the interface, so that band bending effects in the semiconductor (e.g. due to charge transfer across the interface in doped materials and the resulting formation of space charge regions) will not be included, although we return to this point at the end of Section III.

Let us now consider the tunnel currents in the materials pictured in Fig. 4. The separation between conduction band minimum and tip Fermi level is larger for the material with smaller band gap, which would imply a greater number of states available for tunneling and hence a larger tunnel current in that case compared to the larger band gap material. However, this conclusion is valid only for relatively small bias voltages between tip and semiconductor. For larger voltages, typically \( \geq 2 \) V, one must explicitly include consideration of the voltage-dependence of the transmission term for tunneling through the vacuum region, as pictured by the function \( D(E,V) \) in Fig. 4. Assuming a simple trapezoidal barrier, this function is given by

\[
D(E,V) = \exp \left\{ -2s \left[ (2m/\hbar^2) \left( \bar{\phi} - E + (eV/2) \right) + k_{||}^2 \right]^{1/2} \right\}
\]

where \( \bar{\phi} \) is the average work function between tip and sample, \( s \) is the tip-sample separation, \( V \) is the bias voltage applied to the sample relative to the tip, and \( E \) is the energy of a state in the sample relative to the sample Fermi-level, with \( E \) varying between 0 and \( eV \). This formula also includes the dependence of the transmission term on the parallel wavevector of the tunneling electron, \( k_{||} \). It is well known that the energy and voltage dependence of \( D(E,V) \) is such that the highest lying states are favored in the tunneling process [16]. For \( V > 0 \) the maximum of \( D(E,V) \) occurs at \( E=V \) and for \( V < 0 \) the maximum is at \( E=0 \). The former case is illustrated in Fig. 4. The rate of decrease of \( D(E,V) \) with decreasing \( E \) depends on the work function and tip-sample separation. For typical values of \( \bar{\phi} = 4 \) eV and \( s=6 \) Å, \( D(E,V) \) changes by a factor of 20 as \( E \) varies over a 2 eV interval. Thus, for tunneling voltages larger than about 2 V, increasing the voltage does not lead to a large difference in the number of states available for tunneling, and so an increase in the tunnel current
is not expected. However, the $eV/2$ term in the exponent of Eq. (1), arising from the decrease in the average barrier height with increasing voltage, will still produce an overall increase in the current with increasing voltage.

Given an expression for the transmission term, the tunnel current between tip and sample can be computed according to [17]

$$J = \frac{2e}{(2\pi)^{\frac{3}{2}} \hbar} \int dk_\perp d^2 k_\parallel \left[ f(E - eV) - f(E) \right] D(E, k_\parallel) \frac{\partial E}{\partial k_\perp} \vartheta \left( \frac{\partial E}{\partial k_\perp} \right)$$

where $f(E - eV)$ and $f(E)$ are Fermi-Dirac occupation factors in tip and sample respectively and $k_\perp$ is the component of the wavevector perpendicular to the surface. The step-function $\vartheta ( \partial E / \partial k_\perp )$ in this equation, which restricts the integral to positive values of $\partial E / \partial k_\perp$, arises from the assumption that waves traveling only in one direction can tunnel out of the material. (For the case of free electrons, this same factor is implicitly included in the integral by restricting it to run over only positive values of $k_\perp$ [17]). The bulk band structure is implicitly included in the formula by the dependence of $E$ on $k_\parallel$ and $k_\perp$, although, as is well known from early studies of tunneling, we expect the band structure effects to be small since the group velocity term $\partial E / \partial k_\perp$ in Eq. (2) is largely cancelled by a density of states term which appears when the integral over $k_\perp$ is performed [17]. In the computations below, we explicitly evaluate the dependence of the tunnel current on band structure, by computing it for GaAs, InAs, InP, and GaP materials.

To evaluate the tunnel current, band structure computations are performed using empirical nonlocal pseudo-potentials. Spin-orbit coupling is not included. We focus on the materials GaAs, InAs, InP, and GaP which form the endpoints of the $\text{In}_x\text{Ga}_{1-x}\text{As}_y\text{P}_{1-y}$ alloy system, since these alloys are most relevant to our prior xSTM measurements [5,8]. Results for the band structure and density of states of these materials are very similar to those given previously by Fischetti [18], and will not be repeated here. The tunnel current is computed by numerically evaluating the integral in Eq. (2). We use a wavevector spacing in reciprocal space of 0.01 $(2\pi / a_0)$, yielding 89076 points in the irreducible wedge of the first Brillouin zone. Computations are performed for a tip-sample separation of 6 Å, a temperature of 0 K, and a tip work function of 4 eV.

Results for tunneling into CB states are shown in Fig. 5. As seen there, at low voltages the tunnel current does vary considerably between the various materials. However, at higher voltage $> 3$ V, the current is practically the same amongst the different materials. We find less than a factor of 1.5 variation in the tunnel current, corresponding to a tip height variation of $< 0.20$ Å. The observed variation in tunnel current for voltages $> 3$ V arises in part from the differing lattice constant, $a_0$, of the materials. This dependence arises in Eq. (2) because the volume of the Brillouin zone is proportional to $a_0^{-3}$. Physically, a material with small lattice constant, like GaP, will have a greater state-density than a material with large lattice constant, like InAs. However, this $a_0$ dependence is also partially compensated by the $k_\parallel$ term in Eq. (1), which also depends on $a_0$. The resulting tunnel current varies approximately like $a_0^{-n}$ where $n$ is slightly greater than 1 [19]. If we set all the lattice constants to have the value for GaAs in the computation, then the resulting variation in tunnel current for voltages $> 3$ V is less than a factor of 1.35; this amount represents the effect of the differing band structures on the tunnel current.

We note that the zero of voltage in Fig. 5 is taken to be the CB minimum of InAs. In a typ-
ical experiment it would be at a different point, *i.e.* the lowest-lying conduction band minimum of the relevant alloy material, typically 0.5 V higher than that shown in Fig. 5. In that case, the voltage range for which the tunnel currents converge would be $> 2.5$ V. Furthermore, we consider experiments in which the \( \text{InGaAsP} \) alloy compositions used are in a rather restricted range, which is quite far from GaP \([8,12]\). In that case, the expected variation in tip height variation would be less, typically $\leq 0.07$ Å, and the minimum voltage needed to achieve this convergence of the tunnel currents is also less, typically $\geq 2.0$ V. This voltage range is well within the parameters of usual xSTM experiments on \( \text{InGaAsP} \) alloys, which typically employ sample voltages with magnitude in the range $2.0 - 2.5$ V. Values significantly less than this are generally avoided since they can lead to tip-sample contact, and values much larger than this are also dangerous since they may lead to material transfer between tip and sample.

A close inspection of the current vs. voltage curves in Fig. 5 reveals some weak inflection points, seen most clearly for InAs and InP and indicated by arrows. These features correspond to the onset of tunneling into the $L$-valley conduction band, located at $0.98$ and $0.85$ eV above the $\Gamma$-valley minimum for InAs and InP respectively. No such onset is seen for GaAs in Fig. 5; the $L$-valley is only $0.33$ eV above the $\Gamma$-valley minimum in that case, which apparently prevents any clear discrimination of its onset. Experimentally, tunneling spectroscopy reveals clear $L$-valley onsets for InAs and InP, but no such feature is seen for GaAs \([20]\), in good agreement with the results of Fig. 5. For the case of GaP, it is an indirect material with conduction band minimum at the $X$-point and with $L$-valley and $\Gamma$-valley minima located $0.33$ and $0.42$ respectively above that. In Fig. 5, a very weak inflection point is seen for GaP approximately $0.5$ V above the onset of the current, which corresponds to tunneling into these higher lying bands.

Figure 6 shows results for tunneling into the valence band states. At voltage with low magnitude, the tunnel current again varies greatly between the various materials. At higher voltages, the tunnel currents tend to approach each, although their difference is significantly greater than that seen for the CB states. The states most responsible for the tunnel current originate from the top of the VB, and their barrier height for tunneling is the sum of the work function and band gap. Those values vary considerably between the various materials, giving rise to the spread in current values seen in Fig. 6.

III Discussion

The main result of the paper has been presented in the previous section, namely, that at large, positive sample voltages the tunnel current shows a relatively small variation between the various materials GaAs, InP, InAs, and GaP. In this section we examine some of the approximations and limitations of our theoretical treatment, in an effort to evaluate the reliability of our result.

One issue to be considered is our use of the electron affinity rule which led to equal work functions for all the materials. The band offsets derived above ($0.93$, $0.56$, and $1.31$ eV respectively for GaAs, InP, and GaP relative to InAs) differ slightly from those obtained from the model-solid theory of Van de Walle: $0.84$, $0.55$, and $1.11$ eV respectively \([14]\). These differences may be within the combined uncertainty of the experimental and theoretical determinations \([14,15]\), but they may also reflect some real limitation of the electron affinity rule. In particular, the existence of a dipole layer on a surface will affect the electron affinity, and differing dipole layers for materials across a heterointerface will lead to differing work functions for those materials. In an effort to test the
sensitivity of the computation on this effect, we vary the work functions of the materials. We use the band offsets from the model-theory solid, and add to them the experimental electron affinities to arrive at the work functions: 4.97, 5.05, 5.06 and 4.86 eV for GaAs, InP, InAs, and GaP respectively (again, placing the Fermi-level at the CB minimum of InAs). Using these values, the main effect on the tunnel current at large, positive voltage is to increase the GaP result (due to its reduced work function) by a factor of 1.2, resulting in a total variation of a factor of 1.8 in the tunnel current between materials. Our previous upper limit of < 0.2 Å for the tip height variation would then increase to < 0.3 Å, although considering the restricted range of alloy compositions mentioned in Section II leads to an expected upper limit of ≤ 0.1 Å.

Another important issue to consider is the role of strain in our results. Typically, In$_x$Ga$_{1-x}$As$_y$P$_{1-y}$ alloys are grown coherently on InP, and the resulting biaxial strain in the alloys will shift the energies of the electronic states. This effect will, of course, shift the onsets of the current vs. voltage curves in Figs. 5 and 6. Nevertheless, it is important to realize that these shifts do not significantly change the results for the tunnel current at large, positive voltages. The reason for this insensitivity is that shifting the energy bands does not greatly change the number of states available for tunneling. For example, if we strain all the materials onto InP, then the strain-induced shifts for the conduction band minima are -0.29, +0.27, and +0.14 eV for GaAs, InAs, and GaP respectively [14]. Applying these shifts to the entire band (i.e. neglecting the variation in deformation potential with energy), and recomputing the CB tunnel currents, we find for voltages > 3 V that the current varies by less than 15% from the results presented in Fig. 5.

The most significant approximation made in our computations is the neglect of surface state effects in the tunnel current. It is well known that the current is proportional to the local density of states at the surface, so that surface states will certainly play a role. However, for the (110) face of the III-V semiconductors, tunneling spectroscopy results demonstrate that, at energies ≥ 2 eV into the conduction band, the surface states make only a very small contribution to the tunnel current [20]. At these energies the surface states are all strongly mixed (resonant) with the bulk states, so that we expect Eq. (2) to provide a reasonably good estimate of the tunnel current. Certainly it would be desirable to include surface states in the computation, but we expect that the conclusions of this work will remain unchanged even in that case.

Finally, we return to one aspect of xSTM measurements which we mentioned briefly in Section II, namely, the effect on the tunnel current of band bending (space charge regions) in the semiconductor, as encountered e.g. in pn-junctions or around heterointerfaces in doped material. Such band bending produces significant effects in the tunnel current [21], although those effects are not important in the theory above since we focus only on the region near the heterointerface. However, our result that the tunnel current is insensitive to the material into which tunneling occurs can be used to develop a simple way for including band bending effects. Considering the case for large, positive voltages, we find above that the tunnel current can be expressed in the form $C \exp \left\{ -2 s \sqrt{\left( 2m \left( \Phi - eV/2 \right) / \hbar^2 \right)} \right\}$ where the constant $C$ is nearly independent of $V$ and material parameters. In this case, a shift in the Fermi level of the semiconductor can be easily accounted for by taking the average work function $\Phi$ to be a function of position across the heterostructure (i.e. it would contain an additive term of $-e/2$ times the electrostatic potential). Computations of this sort could account for observed variations in current due to intrinsic electrostatic variations in the semiconductor material, although the effect of tip-induced band bending and/or band bending due
to surface Fermi-level pinning are more complicated and would require additional considerations [21,22].

IV Conclusions

Based on the computations reported here, we find for large, positive voltages that the tunnel current expected from $\text{In}_x\text{Ga}_{1-x}\text{As}_y\text{P}_{1-y}$ alloys of any composition will be nearly the same. Less than a factor of 2 variation is found between GaAs, InAs, InP, and GaP materials, corresponding to a tip height variation in the STM of $<0.3$ Å. Typically in xSTM experiment one encounters alloy variations which span only a considerably restricted subset of the total range between GaAs, InAs, InP, and GaP. In that case, the expected variation in tip height due to electronic effects would be smaller, typically $\lesssim0.1$ Å. As discussed above, the relatively small size of this electronic contribution to the tunnel current then enables the possibility of quantitatively determining strain variations in the heterostructures from the observed strain-induced displacement of the cleavage face.

V Acknowledgments

We are most grateful to M. Fischetti for supplying the band structure code used in the computations reported here. We also thank Huajie Chen for many useful discussions during the course of this work, and Chris Van de Walle for a critical reading of the manuscript. We acknowledge Sung Ha Park and Li Zhang for assistance with the computer computations. This work was supported by the National Science Foundation under grant DMR–9615647.

2251 (1999).


[19] This dependence of the current on lattice constant will, of course, change slightly when one considers layers of In$_x$Ga$_{1-x}$As$_y$P$_{1-y}$ biaxially strained on a particular substrate (such as InP).


---

Figure 1  (a) Constant-current image and (b) associated conductance image of the (110) cleaved surface of low-temperature-grown In$_{0.53}$Ga$_{0.47}$As, acquired with 0.1 nA tunnel current, at a sample bias of −1.45 V. A point defect with two satellite features is observed in both images. The grey scale range in (a) is 0.08 nm. From Ref. [5].
Figure 2  Cross-sectional STM images of In$_{0.53}$Ga$_{0.47}$As/InP heterostructures. (a) as-grown sample; (b) implanted and annealed sample. Images were acquired at sample voltages of +2.5 and +2.0 V respectively, and grey scale ranges are 0.05 and 0.06 nm respectively. An average of the topographic line scans is shown in (c) and (d). Layer growth direction is from right to left. Arrows indicate approximate width of the quantum wells. From Ref. [8].
Figure 3 Illustration of STM contrast mechanisms for a strained semiconductor superlattice containing quantum wells in tension and barrier layers in compression. The dashed line shows the constant-current contour followed by the STM probe tip considering (a) only electronic, and (b) only mechanical effects. For case (a), the barrier has a larger band gap than the quantum well, so for a given tip-sample voltage there are fewer states available for tunneling to the barrier. Thus, a lower current is produced for a fixed tip-sample separation, so that the tip moves towards the sample to maintain a constant tunnel current. For case (b), relaxation of the strain produces an undulating surface morphology across the superlattice. For a strain of $\pm \varepsilon$ in the layers and width of both barrier and well of $L$, the peak-to-peak amplitude $2h$ of the undulations is computed by finite elements to be $h / \varepsilon L \approx 1.0$ for Poisson ratio of 0.35. From Ref. [12].
Figure 4 Energy level diagram for the case of tunneling into two neighboring semiconductor materials, with aligned Fermi levels, $E_{F,s}$. The probe tip Fermi level is denoted by $E_{F,t}$, and is separated from the sample Fermi level by $eV$ where $V$ is the applied voltage. The vacuum level of the sample is denoted by $V_{AC}$ and electron affinity by $\xi$. The tunneling transmission term is indicated by $D(E,V)$.

Figure 5 Computed tunnel current for CB states of GaAs, InAs, InP, and GaP. Arrows indicate the location of inflection points seen for InAs and InP, and seen weakly for GaP.
Figure 6  Computed tunnel current for VB states of GaAs, InAs, InP, and GaP.