Scanning tunnelling microscope imaging of nanoscale electron density gradients on the surface of GaAs

2003 J. Phys.: Condens. Matter 15 S3083

(http://iopscience.iop.org/0953-8984/15/42/007)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **15** (2003) S3083–S3093 PII: S0953-8984(03)63736-7

Scanning tunnelling microscope imaging of nanoscale electron density gradients on the surface of GaAs

B Hamilton[1], **J Jacobs**[1,3] **and M Missous**[2]

[1] Department of Physics, UMIST, Manchester M60 1QD, UK
[2] Department of Electrical Engineering and Electronics, UMIST, Manchester M60 1QD, UK

E-mail: bruce.hamilton@umist.ac.uk

Received 19 May 2003
Published 10 October 2003
Online at stacks.iop.org/JPhysCM/15/S3083

Abstract

This paper is concerned with the scanning tunnelling microscope tunnelling conditions needed to produce constant current images dominated *either* by surface topology *or* by electronic effects. A model experimental structure was produced by cleaving a GaAs multi δ-doped layer in UHV and so projecting a spatially varying electron gas density onto the (110) surface. This cross sectional electron density varies on a nanometre scale in the $\langle 100 \rangle$ growth direction. The electronic structure and tunnelling properties of this system were modelled, and the tunnelling conditions favouring sensitivity to the surface electron gas density determined.

1. Introduction

The scanning tunnelling microscope (STM) has revolutionized the imaging of surfaces and in particular the surfaces of those electronic materials which can be prepared in an ultra-clean manner. In z imaging mode, i.e. measurement at constant current and fixed applied bias, the z movement of the tip, under feedback control, ensures that the tunnelling gap changes by the correct amount to supply the required current at all points imaged. It is a well recognized feature of this mode that the image can contain both topological (mechanical) and electronic information: any interaction between the tip and surface which causes the tunnelling current to vary will result in a z movement of the tip and hence contribute to contrast.

The topological component is intuitively simple to appreciate; a change of surface height on the nanometre or even atomic scale will force the tip to approach or retract. For the case of semiconductors, the electronic contribution, although rather complex, is most frequently linked to the band structure of the solid. Thus local changes of the density of states in a band, e.g. a change in energy gap crossing a heterojunction, will influence the tunnelling current and move the tip. Usually it is the onset of electron tunnelling into an empty conduction band or out

[3] Present address: School of Pharmaceutical Sciences, The University of Nottingham, Nottingham NG7 2RD, UK.

0953-8984/03/423083+11$30.00

of a full valence band which is linked to the imaging of electronic information. We show here that, in addition to these DOS contributions, it is possible to image an electronic contribution from local variation in surface free electron density, provided the variation in density is large enough. This image contrast was observed on the unreconstructed (110) GaAs surface. Our data indicate that it is possible to establish tunnelling conditions that result in much greater sensitivity to either topology or to electron density. We also conclude that the almost pure electronic images observed in this work are linked to the electrostatic interaction between the tip and the electron gas in the surface; they are not strongly associated with the 3D GaAs DOS functions.

Every tunnelling barrier has a charge dipole associated with it. For a metal–metal barrier, the charges are located within a Sommerfeld screening length of each metal surface. In this case the applied potential drops entirely across the tunnelling gap. For a metal–semiconductor barrier (or any rather dilute electron gas system) the dipole field spans both the gap and a sub-surface depletion region in the semiconductor. The tunnelling gap potential *is no longer uniquely defined as the applied voltage* because the applied voltage must now partition between the gap and the semiconductor. The relative partitioning depends on the polarity of bias, the size of the tunnelling gap and the magnitude of the charge density in the semiconductor. The data discussed here were obtained with a negative bias on an n-type semiconductor i.e. in the 'forward bias' direction, which favours surface accumulation of electrons. Sub-surface charge and field distributions have been most frequently discussed for scanning tunnelling spectroscopy in the reverse bias case and referred to as 'tip-induced band bending' [1–4].

2. Semiconductor structure and images

Atomic plane (δ-doped) layers were used to obtain modulated electron densities in GaAs. The structures were grown in a solid source VG90 MBE system. Periodic Si δ layers were inserted into the growing layer using software-controlled shutters, and the Si atomic concentration was calibrated to be 5×10^{12} cm^{-2}; the total sheet electron density is also expected to be 5×10^{12} cm^{-2} at room temperature. A minimal As_4 to Ga ratio was used, and in order to suppress dopant migration the growth was carried out at a temperature of about 520 °C. A one micron buffer layer was grown before the δ doping commenced and growth was terminated with a 200 Å capping layer. The δ-doped layers were inserted at intervals of 100 nm. This relatively large spacing ensures that there is no electronic interaction between adjacent layers that might otherwise modify the potentials, wavefunctions and electron density distribution associated with the doping structure [5].

The doping modulation in these structures is in the ⟨100⟩ growth direction and is therefore not accessible to STM imaging on the (001) growth surface. By cleaving the crystal in a plane normal to the plane of the growth surface, namely the (110) plane, the doping modulation is projected completely onto the measured surface, i.e. it is fully accessible to the STM tip. This procedure was carried out *in situ* in the Omicron UHV-STM/AFM used for this work. Such a surface, created by fracture, is not reconstructed (unlike the growth surface which is heavily reconstructed). Measurement of such cross-sectional surfaces is recognized to provide non-surface-specific information such as defect and impurity properties [6].

Two STM z images obtained from exactly the same scan field on these cleaved structures are shown in figure 1. The inset to the figure illustrates the measurement configuration. These images were taken after establishing different set points; the primary effect of this is that the average tunnelling gap will differ significantly. In figure 1(a) the set point (−3 V, 0.5 nA) corresponds to a rather wide tunnelling gap, and in figure 1(b) with a set point of (−1.1 V, 0.3 nA) the tip was made to approach the surface for the image acquisition. Figure 1(a) shows

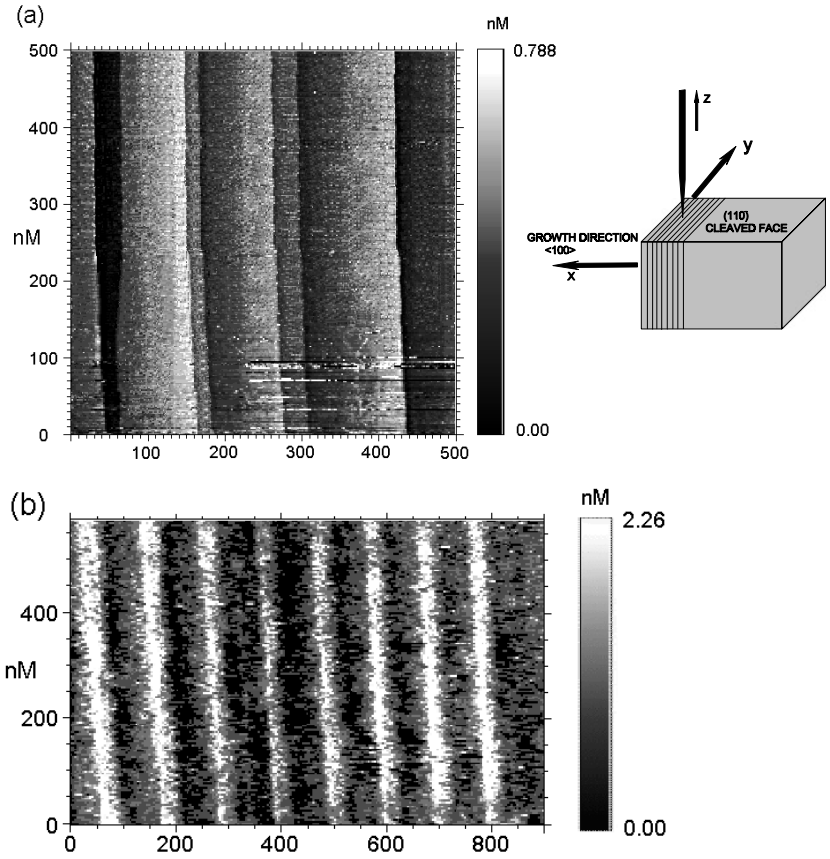

Figure 1. z images taken of the UHV-cleaved (110) face of Si δ-doped MBE GaAs. In (a) the set point was (-3 V, 0.5 nA); the image is dominated by the topology of cleavage steps with no apparent electronic contributions. Reducing the set point to (-1 V, 0.3 nA) for the image shown in (b) forces the tip to approach the surface. The δ layers are now clearly seen and the cleavage steps are missing. No topology is obvious in the image which is therefore dominated by electronic effects. The schematic diagram shows the measurement configuration.

the cleavage steps typically observed for the GaAs(100) surface cleaved in UHV, with steps of a few ångströms. During the formation of this image the STM tunnelling response was dominated by the surface topology; there is no evidence of δ-doped layers in this image.

The opposite is the case for the image taken after approaching the surface with the tip, which shows periodic contrast corresponding exactly to the known spacing of the δ-doped layers. The contrast, which corresponds to the tip moving away from the semiconductor as it passes over a δ layer, extends significantly (on the nanometre scale) either side of the doping plane. This gives a clue to the origin of the tunnelling component involved since the bound state electron wavefunctions associated with the δ-doping potential (discussed below) extend significantly beyond the doping plane in the growth direction, labelled x in figure 1. We take this image to be predominantly electronic since the cleavage steps are now absent. The electronic component, i.e. the mechanism for the strong increase in tunnelling gap is shown below to be the tunnelling of electrons associated with the doping plane into the tip under the negative sample bias.

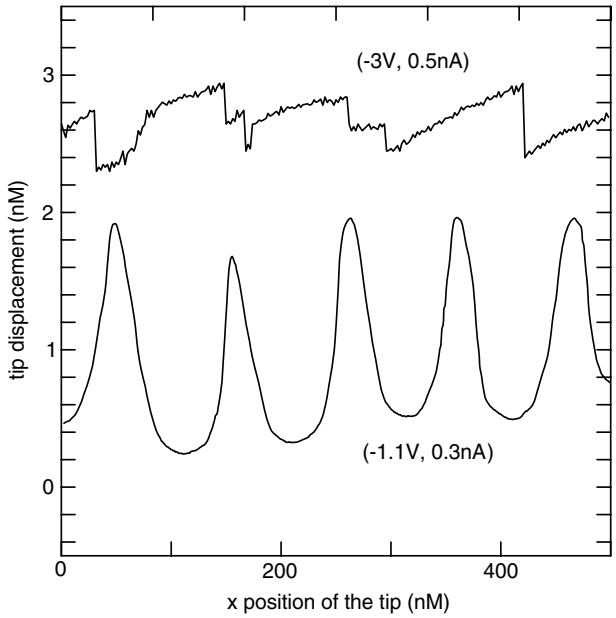

Figure 2. Tip displacements measured on adjacent line scans, using the two set points which resulted in topological (1(a)) and electronic images (1(b)). The average distance of the tip from the surface is clearly less for the low voltage set point (by about 15 Å), and the tip variation for this set point is large (about 14 Å) as it crosses the δ layers.

Since the basis of image formation is tip displacement, it is useful to consider z displacement line scans for the two set points used in image acquisition. Measured line scans are shown in figure 2. The line scans were taken within a few nanometres of each other and so are measuring at nearly the same location on the two images.

In order to model how the chosen tunnelling set point can result in such different tip displacements and hence images, it is necessary to have insight into the electron density and energy distributions of the 2-DEG and also the (essentially electrostatic) interactions between the tip and the 2-DEG.

3. Electronic structure of the δ-doped layers

By placing the silicon dopant atoms on a plane, the randomly distributed impurity charge of the uniform doping case is replaced by a coherent charge plane. Such a charge distribution gives rise to a well defined potential and electric field in the growth direction. In the absence of screening this would be a V-shaped potential, with a minimum located on the positive charge sheet. Such a potential binds the added electrons and is therefore screened and modified in shape. The electron localization is sufficiently strong that bound states are formed normal to the doping plane and so the spread of charge away from the doping plane must be calculated quantum mechanically. This spread of the bound state wavefunctions defines the lateral (x) electron density which is imaged in figure 1(b).

The electronic structure of δ-doped layers has been analysed by several authors [7–9]. A similar analysis was carried out here because numerical values for the δ-potential $V(x)$ and wavefunctions $\psi(x)$ were needed to calculate the tunnelling currents. The methodology will not be repeated in detail, but the basis of the calculation is described briefly.

The Poisson and Schrödinger equations were solved simultaneously and self-consistently under the Hartree approximation. The effective mass approximation of the Schrödinger equation was used, namely

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_0(x) + V_s(x)\right]\Psi_{nk}(r) = E_{nk}\Psi_{nk}(r) \tag{1}$$

where m is the electron effective mass, n is the sub-band index. The eigenvalues E and wavefunctions are also parametrized in terms of the in-plane wavevectors k.

In (1) V_0 is the bare potential of the ionized donors and V_s is the Hartree screening contribution. The bare potential was obtained from solution of Poisson's equation including charge from both the donors in the δ sheet, N_δ^+, and the ionized (low-doped) n-type background impurity density N_D^+:

$$\frac{d^2V_0}{dx^2} = \frac{4\pi}{\varepsilon_0\varepsilon_s}[N_\delta^+ + N_D^+]. \tag{2}$$

The Hartree potential was obtained using

$$\frac{d^2V_s}{dx^2} = \frac{4\pi}{\varepsilon_0\varepsilon_s}[N(x)] \tag{3}$$

with $N(x)$ being the free electron density. Since the potentials have only x dependence, the Schrödinger equation can be solved in the three orthogonal directions. In the y–z plane of the δ layers the single-particle wavefunctions are plane waves with corresponding wavevectors k_y and k_z; in the x direction the wavefunction represents the bound state. The total wavefunction, with the normalization coefficient C, can be written as

$$\Psi_{nk}(r) = \frac{1}{\sqrt{C}}\exp(ik_y y)\exp(ik_z z)\phi_n(x). \tag{4}$$

The wavefunction is obtained by self-consistently solving the simultaneous equations (1)–(3). The bound state wavefunction $\phi_n(x)$ contains the information needed to compute the electron density profile. The energy eigenvalues correspond to parabolic sub-bands with minima at E_n, i.e.

$$E_{nk} = E_n + \frac{\hbar^2}{2m}[k_y^2 + k_z^2]. \tag{5}$$

The spatial distribution of the total 2D electron density $N(x)$ in the system is

$$N(x) = \sum_n N_n^{2D}\{\phi_n(x)\}^2 \tag{6}$$

where the summation accounts for the total 2D electron density (N^{2D}) in each of the n occupied sub-bands which are spread according to the normalized wavefunction in x. The values of N^{2D} are defined by

$$N_n^{2D} = \int_{E_n}^{\infty} \rho^{2D}(E) f(E)\, dE. \tag{7}$$

The bulk effective mass for electrons in GaAs was used to calculate the 2D density of states function, ρ^{2D}. The appropriate boundary conditions (relating to parity) on each $\phi_n(x)$ and its derivative were employed, and the Fermi energy used in the Fermi–Dirac function $f(E)$ was forced to match the calculated value corresponding to the background electron density far away from the δ layer.

The solution for a single layer with the experimental value of 5×10^{12} cm^{-2} for the total sheet donor density is shown schematically in figure 3. The form of the screened

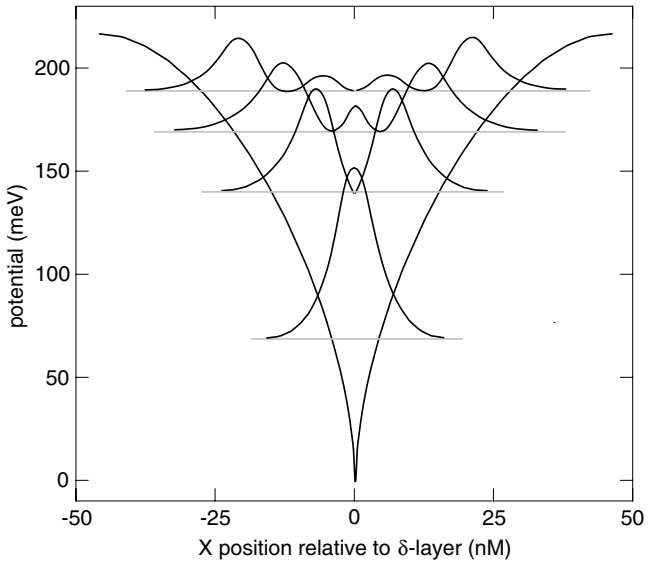

Figure 3. The forms of the potential and $\phi_n^2(x)$ resulting from the self-consistent solution of the Schrödinger and Poisson equations for a silicon donor sheet density of 5×10^{12} cm^{-2} δ layer. Four sub-bands are occupied at room temperature, and the sub-band minima are indicated.

potential shows the typical cusplike structure. The computed values of $\phi_n(x)^2$ normalized to the calculated sub-band occupancy are sketched. At the measurement temperature (300 K) the total potential depth, which is mirrored by the band-bending, is close to 200 meV. Four sub-bands i.e. $n = 1$–4 are occupied, though the Fermi energy is only a few tens of millielectronvolts above the bottom of the $n = 4$ sub-band, which therefore makes only a small contribution to the total electron density profile. It is also worth commenting that the shape of $V(x)$ lifts the lowest $n = 1$ sub-band minimum about 55 meV above the potential minimum. The sub-band spacing plays a significant role in the calculation of tunnelling currents.

Finally, the calculated result for the equivalent volume electron density (derived from (6)) is shown in figure 4. This, together with the energy distribution for these electrons, is central to the calculation of the z response of the tip due to the onset of tunnelling from the electron gas.

We note that many-body, electron–electron interactions should strictly be included in these calculations. However, it is well established that these effects result in fairly small rigid shifts of the bands. Density functional theory calculations for similar doping structures [8] yield shifts of between 15 and 20 meV. Such effects will make little impact on the issues discussed here.

4. Electron tunnelling and z-imaging near δ-doped layers

Our aim was to predict tip displacement for the high and low voltage set point regimes. For simplicity we performed a one-dimensional calculation using the WKB tunnelling approximation; this is a standard approach for phenomenological predictions. The electronic features of the δ-doped layers were included in the calculation and, although this involved further approximations, we believe that our model preserves the basic physics of the tunnelling processes and successfully explains the measured tip displacements.

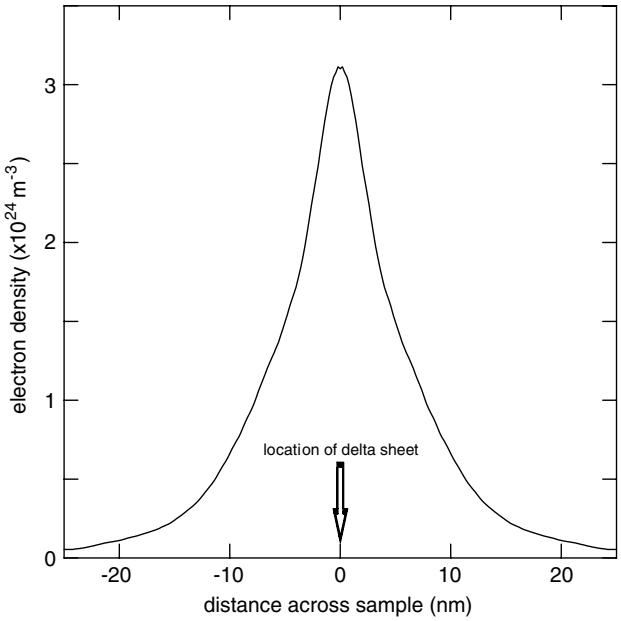

Figure 4. The volume electron density profile $N(x)$ calculated by summing the normalized $\phi_n^2(x)$ for the occupied two-dimensional density of states over all sub-bands and taking into account the wavefunction spread in x. This is the electron gas density that would be projected onto the (110) surface in the absence of any electric field gradients normal to the surface.

The general form of the tunnelling current can be written

$$J = \frac{4\pi e}{\hbar} \int f(E)_\delta (1 - f(E)_{\text{tip}}) \rho(E) D(E)\, \mathrm{d}E. \tag{8}$$

This expression includes Fermi functions for both tip and semiconductor, an appropriate DOS function $\rho(E)$ and the transmission coefficient $D(E)$. Remote from the δ-layers, $\rho(E)$ includes the usual valence and conduction band DOS functions. For all regions in which the 2-DEG density was greater than the background electron level the conduction band DOS function was replaced by $\rho(E)^{2\mathrm{D}}$.

For the lower voltage set point (-1.1 V), the dominant current component was found to be due to tunnelling of electrons out of the 2-DEG into the tip. Even when the tip was located close to the mid-point between δ sheets, the background dopant electron density accounted for the measured current. At this set point voltage no valence band tunnelling occurs, and we show below that a large tip movement towards the surface is required to sustain the current midway between the δ-layers. For the higher voltage set point (-3 V), the tip Fermi level is displaced below the valence band edge and valence band tunnelling dominates. The 2-DEG component is of course still active, but is small compared to the valence band component.

The WKB transmission coefficient in (8), $D(E)$, is given by

$$D(E, V) = \exp\left\{-2d\left[\frac{2m}{\hbar^2}(\langle\varphi\rangle - E) + \frac{eV}{2}\right]\right\} \tag{9}$$

where d is the tunnelling gap, $\langle\varphi\rangle$ the average work function of the tip and sample and E is the energy of the tunnelling electron referenced to the metal Fermi energy. The voltage term, V, in (9) is to be strictly interpreted as the *fraction* of the applied bias transferred across the gap. As mentioned in the introduction, this fraction is less than one for any tunnelling barrier

which involves sub-surface fixed charge. In the present case any bias-induced movement of the electron gas and associated change in the screening of the ionized donor sheet will change the relative partitioning of potential between the gap and the semiconductor. Furthermore, because of the strong electron density gradient in the x direction, this voltage partitioning and therefore the tunnelling barrier properties are also x dependent.

Since the electrostatic properties of the tunnelling barrier are an important component in the calculation of tip displacement, it is useful to review these. The simplest insight into the spatially varying form of the barrier can be gained by considering the equilibrium conditions that correspond to different tip locations on the surface. These are the barriers established simply by approaching the tip in the absence of bias. The potential barriers consistent with our calculation are illustrated schematically in figure 5.

Far away from a δ layer (figure 5(a)), the barrier is typical of n^- material. The contact potential V_c is divided between the gap, V_g, and depletion in the semiconductor, V_s. Clearly $V_c = V_g + V_s$. In figure 5(b) the tip is above the tail of the 2-DEG $N^{2D}(x)$ distribution and can sense the electron gas. At this new x location, the conduction band minimum at the edge of the barrier and beyond is lowered by an amount dictated by the form of $V(x)$. This represents an increase in V_c that is again distributed between the gap and the semiconductor. The magnitudes of V_s and the associated unscreened ionized donor charge density control the distance between the electron gas and the surface. Finally (figure 5(c)), when the tip is centred above the δ layer, the total barrier (contact) potential is greatest but the electron gas is much closer to the surface.

The data described above were obtained with two values of applied bias, V_a. In each case the total barrier potential obeys the rule $V_a + V_c(x) = V_g(x) + V_s(x)$, where the variation of V_c with tip location is included. For both data sets, the bias applied to the sample was negative; this 'forward' voltage drop reduces both V_g and V_s; in fact, for both values of V_a used to define the set points, the contact potential was completely offset, and the field in the gap reversed compared to the zero-bias case. However, the relative magnitudes of V_g responsible for the tunnelling current will vary at each x because the fraction of V_a 'consumed' in modifying the barrier shape varies.

These shape changes of the tunnelling barrier influence the value of the tunnelling current, and must be incorporated into any model for tip movement. However, for the set point corresponding to δ-layer imaging, the most important effect was the increase in the magnitude of the electron gas density translated into the surface by the bias. This electron reservoir was responsible for the tunnelling current.

The calculation procedure for the tip displacement was briefly as follows. For each tip location x (with associated V_c and N^{2D}), equilibrium barrier properties were computed assuming an initial value of tunnelling gap, d. The static charges in the barrier result from surface depletion of the 2-DEG which reduces the net screening of both donor impurities in the δ-sheet and of the background ionized donors. This was approximated in the calculation to an effective sub-surface volume charge. This approximation should be most accurate far away from a δ layer (where only background donors exist) and also directly on the δ sheet when the tunnelling current and the majority of ionized donors are in the same plane.

The set point voltage was then applied to the barrier, the static charges were adjusted and the tunnelling gap potential calculated. The tunnelling current was then computed, taking into account the barrier shape and sub-band populations. The gap was then adjusted and the process continued iteratively until the measured set point current was obtained. This value of d was stored as an element of the calculated z line scan. The calculation was repeated for x values corresponding to scanning the tip over the whole δ layer.

The results of this calculation are shown in figure 6. It is clear that our calculation reproduces the main features of the experimental line scans shown in figure 2. Since our

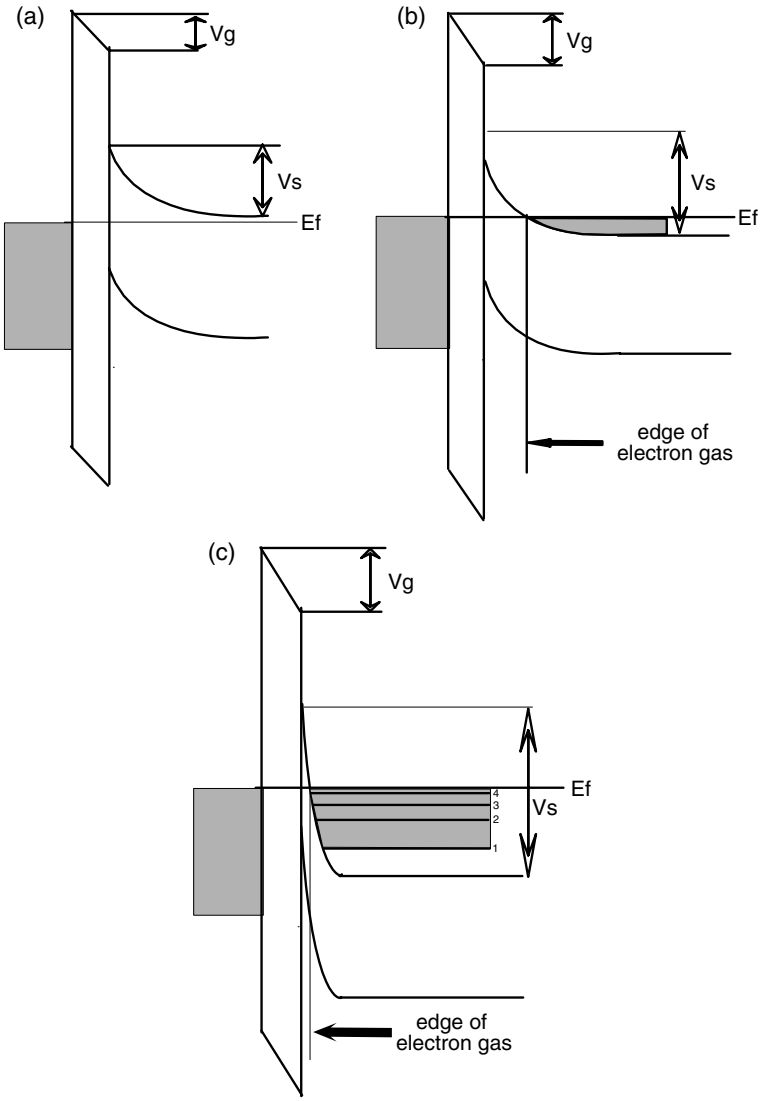

Figure 5. The zero-bias tunnelling barrier for different locations of the tip with respect to a δ plane. (a) Far away from the dopant sheet the barrier is typical of n material. (b) Approaching the δ layer, the self-consistent potential adds to the total potential variation in the semiconductor. (c) Centred on the δ layer, the potential variation is now a maximum and the electron gas is closest to the surface. The electrons which will become available for tunnelling are distributed in energy over four sub-bands.

analysis includes only electronic effects, the tip displacement for the wide gap (set point -3 V, 0.5 nA) is constant; we did not add artificial step edges to this calculation.

The tip displacement calculated for the smaller gap set point clearly shows the correct trend. In particular the total gap modulation is within 20% of the measured change. This figure corresponds to the on and off δ layer extremes where our calculation is most accurate. The calculated line profile is narrower than the measured one. We believe this is due both to approximations in the calculation and to spread of the δ sheet during growth. A contribution

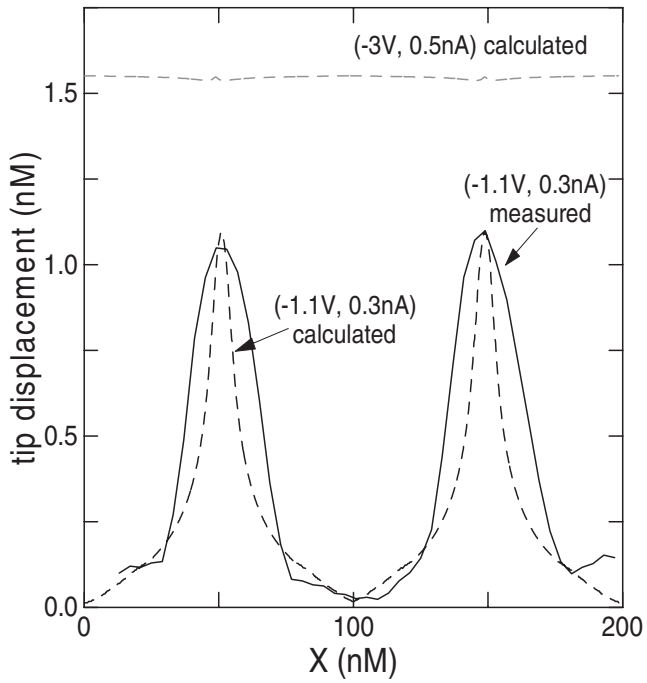

Figure 6. The calculated tip–sample displacement due to electronic effects for the two imaging set points. The set point corresponding to the topology image, (−3 V, 0.5 nA), with a large tunnelling gap, is predicted to show almost zero tip movement when crossing a δ sheet (upper curve). The tip movement for the δ layer imaging set point (−1.1 V, 0.3 nA) is calculated to be large and centred on the δ plane where the electron gas density is a maximum. The latter calculation is compared with experimental data.

from topology is present in the overall tip movement but is small compared to that from the 2-DEG interactions.

5. Conclusions

In this work, we have shown that it is possible to directly z-image variations in electron density on a nanometre scale on the UHV cleaved (110) GaAs surface. It emerged that in order to obtain such an image, specific tunnelling conditions, corresponding to a small tunnelling gap, must be met. Otherwise an image dominated by surface topology was obtained. It is possible to switch reversibly between the two regimes by appropriate choice of tunnelling set point.

The modulated electron density studied was obtained by cleaving MBE grown δ-doped layers, exposing the associated confined electron gas. The electronic structure of the 2-DEG and the tunnelling barriers formed by the tip and semiconductor were modelled. The model was used to calculate the z movement of the tip as it was scanned over a δ layer, and successfully predicted the trends measured. It should be noted that the absolute value of the tunnelling gap is never known in STM measurement; only tip displacement is measured. In the calculations described, the minimum value gap used (midway between δ sheets) was 3 Å.

The source of the tunnelling component responsible for the large tip displacements measured for the low voltage set point is unambiguously established as the 2-DEG electrons (and not valence band electrons) tunnelling into the tip. This form of tunnelling component is

not frequently observed, but has been detected as a 'dopant-induced component' in uniformly doped GaAs [1].

The fundamental reasons for obtaining the two very distinct images emerge naturally from the differing origins of the tunnelling current. At the low voltage set point, the bias is mainly consumed in flattening the band bending associated with the contact potential and hence translating the 2-DEG towards the surface. Although the gap voltage is position sensitive, the electron density gradient is the dominant influence on the tip displacement. This ensures that that the electron density gradient is imaged.

For the higher voltage set point, with the larger average tunnelling gap, valence band tunnelling dominates. The 2-DEG has little influence on the total measured current. Furthermore, at such high biases the energy dependence of the WKB transmission coefficient makes the tunnelling current rather insensitive to the valence band DOS functions [6]. It emerges then that the most sensitive remaining influence on the tunnelling current is the surface topography. The tip moves in response to surface height variation and in our case the cleavage steps are imaged.

Acknowledgments

The authors would like to thank EPSRC for the support required to make this work possible (JJ). Thanks are also due to Eric Whittaker for much technical support and to Richard Cobley and other staff of the Semiconductor Interface Laboratory at the University of Wales Swansea for valuable discussions on voltage-dependent STM imaging.

References

[1] Feenstra R M and Stroscio J A 1987 *J. Vac. Sci. Technol.* B **5** 923
[2] Weimer M, Kramer J and Baldeschweiler D 1989 *Phys. Rev.* B **39** 5572
[3] Ness H, Fisher A J and Briggs G A D 1996 *Surf. Sci. Lett.* **380** L479
[4] de Raad G J, Bruls D M, Koenraad P M and Wolter J 2002 *Phys. Rev.* B **66** 195306
[5] Ke M L, Rimmer J S, Hamilton B, Evans J, Missous M, Singer K E and Zalm P 1992 *Phys. Rev.* B **45** 227
[6] Feenstra R M 1999 *Physica* B **273/274** 796–802
[7] Zrenner A, Koch F and Ploog K 1988 *Surf. Sci.* **196** 671
[8] Ke M L and Hamilton B 1993 *Phys. Rev.* B **47** 4790–3
[9] Degni M H 1991 *Phys. Rev.* B **44** 5581