1 Introduction

1.1 The Dominance of Silicon

The electronics industry is currently worth over $100 billion with over 98% of this belonging to the silicon market.[1] Since the birth of semiconductors in the 1950's the silicon market has shown an exponential increase that is expected to continue for many years (see Figure 1.1). The dominance of silicon can be attributed to several factors including the abundance (and therefore low cost) of silicon raw materials, the ease of processing of silicon electronic devices and the properties of the readily formed oxide (SiO_2). This oxide forms a convenient gate dielectric for the fabrication of metal oxide semiconductor (MOS) devices and can be used for their passivation and electrical isolation.

In order to increase the speed of silicon devices, to keep up with consumer demands, it is necessary to continually decrease their size. This device scaling is

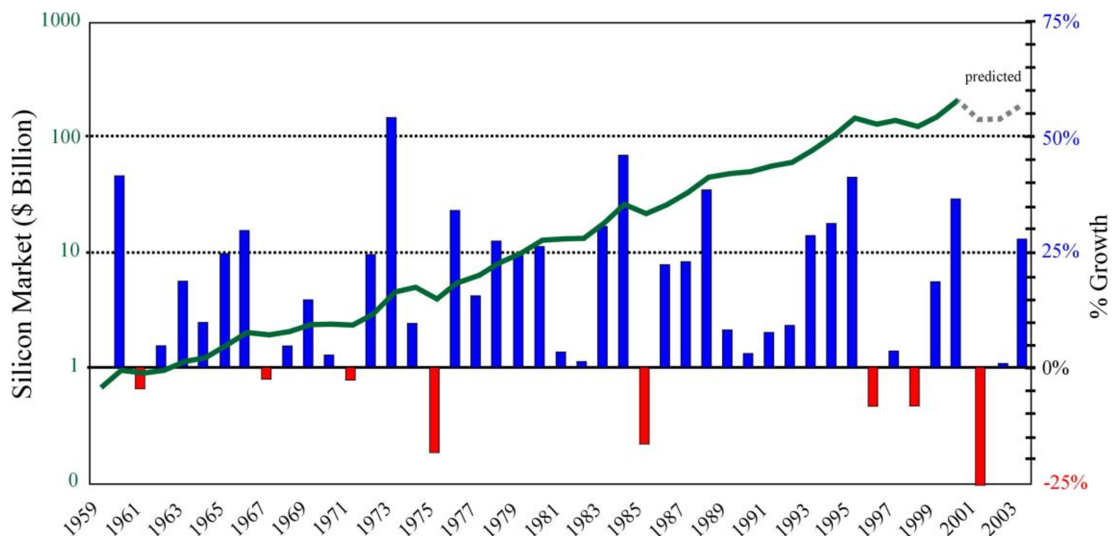


Figure 1.1 Graph showing the value of the silicon market and semi-annual percentage growth since 1959. [Source WSTS.TS]

often described by Moores law, an empirical rule that shows that the size of commercially available devices halves every 18 months. Currently, production is at 0.18 μm gate lengths with 0.13 μm gate lengths expected in early 2003. However, the fundamental limits to device scaling will soon be met and it is necessary for new materials to be found if the increase in device performance is to continue. The processing of silicon devices is now a mature technology with billions of dollars of investment, for this reason the industry is unwilling to change to materials which are incompatible with existing fabrication processes. Since the 1980's silicon germanium (SiGe) has been studied in order that it may extend the capabilities of the silicon industry. With only small modifications, this SiGe technology can easily be incorporated into existing silicon fabrication labs.

1.2 The Advantages of Silicon-Germanium

Germanium is completely miscible in silicon and readily forms the random alloy $Si_{1-x}Ge_x$ were x represents the fractional germanium composition and has a value $0 < x < 1$. The lattice parameter of $Si_{1-x}Ge_x$ varies approximately linearly with x throughout the compositional range (Vegard's Law). If a thin layer of SiGe is grown on a silicon substrate, using one of the epitaxial layers to be discussed in chapter 2, the SiGe extends the crystal of the underlying substrate with each atom of the SiGe layer keeping in complete registry with the atomic positions of the substrate, this is known as *pseudomorphic* growth. In order that this may occur, the SiGe layer has to be compressively strained in the plane of growth. This biaxial strain profoundly affects the electronic bandstructure and leads to enhancement of the electronic properties.

	Electron mobility ($cm^2V^{-1}s^{-1}$)	Hole mobility ($cm^2V^{-1}s^{-1}$)
Silicon	1450	505
Germanium	3900	1800

Table 1.1 Electron and hole mobilities in bulk silicon and germanium.[2]

A useful parameter in characterising an electronic material is the carriers drift mobility (μ) which is defined by[3]

$$\mu = \frac{v}{E} \tag{1.1}$$

where v is the drift velocity and E is the electrical field. This mobility can be shown to be given by

$$\mu = \frac{e\tau}{m^*} \tag{1.2}$$

where e is the electronic charge (1.6022×10^{-19} C), τ is the mean scattering time and m^* is the effective mass of the carriers. The mobility of bulk silicon and germanium is given in Table 1.1. The mobility is generally reduced in a MOS device due to scattering at the Si/SiO_2 interface which reduces τ. Typical values of the hole and electron mobilities in MOS devices are[4] 100 and 400 $cm^2V^{-1}s^{-1}$ respectively. This discrepancy between the hole and electron mobilities leads to a disparity between the size of p and n-channels of CMOS devices and leads to a reduction of the packing density.

The minima of the conduction band of unstrained silicon and $Si_{1-x}Ge_x$ ($x \leq 0.85$) lie along the six <001> directions, with the addition of biaxial strain the out of plane valleys move further away from the zone centre and the in-plane valleys move closer. This splitting reduces the inter-valley scattering which leads to an increase in electron mobility. The valence band of unstrained SiGe which lies at the zone centre (Γ point) is degenerate containing the light hole (LH) and heavy hole

(HH) bands. The effect of the strain is to lift this degeneracy by raising the energy of the HH bands and reducing the energy of the LH bands with a corresponding reduction of the HH mass leading to a “mass inversion”. This splitting leads to the suppression of interband scattering (increasing τ) and the reduction of the effective mass of the HH band, both increasing the hole mobility. A schematic diagram of these band modifications is shown in Figure 1.2

Another important difference between unstrained silicon and bi-axially strained SiGe is their bandgap. As the germanium composition increases, the bandgap becomes smaller, with most of this reduction being accommodated by the valence band, Figure 1.3 shows the bandgaps of strained and unstrained $Si_{1-x}Ge_x$ as a

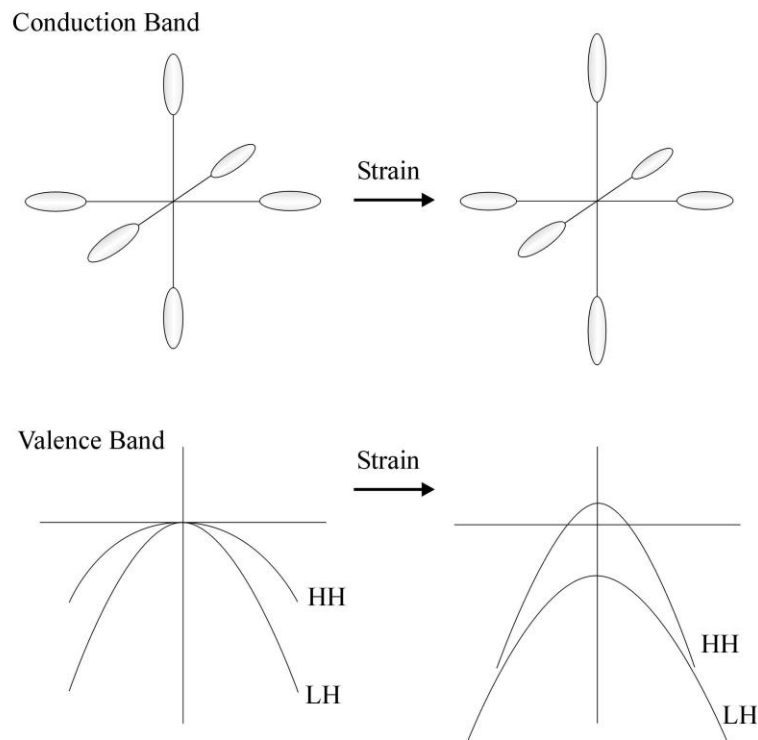


Figure 1.2 Schematic diagram of the elongation in the growth direction of the conduction band minima and valence band splitting due to the effects of bi-axial strain. Notice that the mass (inversely proportional to the curvature) of the HH band becomes less than the LH band (mass inversion).

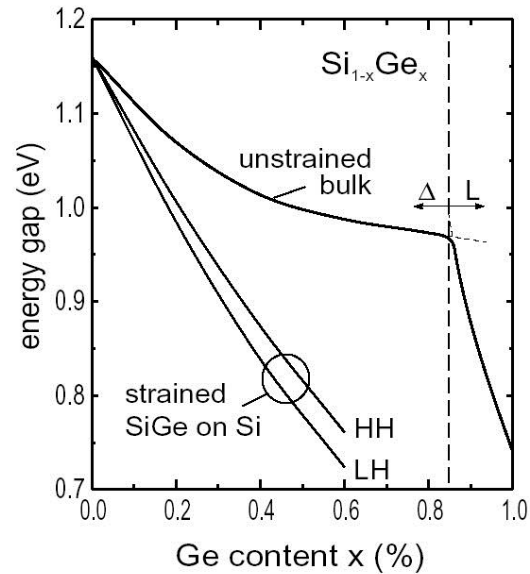


Figure 1.3 Graph showing how the bandgap energy decreases in a $Si_{1-x}Ge_x$ alloy as x increases.[5]

function of x. The effect of this difference in bandgaps leads to a sizeable valence band offset of strained SiGe grown pseudomorphically on silicon which leads to the formation of a quantum well for the confinement of holes. Van de Walle and Martin[6] has shown this to be of the approximate form $\Delta E_v = 0.74x$. This makes possible the growth of buried channel p-MOS devices where the holes are confined to the high mobility SiGe channel, away from the poor Si/SiO_2 interface. It is not the intention of this work to describe the full bandstructure modification of strained SiGe layers, the reader is directed to the excellent review of Schäffler[7] and references therein for further details.

The use of pseudomorphically strained SiGe channels has lead to improvements in hole mobilities, with a record mobility of 17650 $cm^2V^{-1}s^{-1}$ measured at 7 K for channels with $x = 0.065$.[8]

1.3 The Need for a Virtual Substrate

1.3.1 High Germanium Channels

In order that the higher mobilities of high germanium structures can be utilized, it is necessary to grow the channel on a strain adjusted buffer layer. This buffer layer, usually termed a *virtual substrate*, needs to have its lattice parameter tuned so that the strain in the channel is not high enough to cause strain relaxation. Using this technique record hole mobilities in pure Ge channels of 55000 $cm^2V^{-1}s^{-1}$ have been recorded at 4.2 K [9] with room temperature mobilities as high as 1870 $cm^2V^{-1}s^{-1}$.[10] The most common way to produce these virtual substrates is to grow an intermediate layer of $Si_{1-x}Ge_x$ with a value of x lower than the channel in such a way that this layer relaxes to its bulk state. To grow a pure germanium channel it is necessary to use a virtual substrate with $x \geq 0.50$. At such high germanium compositions the virtual substrate is of low quality leading to serious degradation of the performance of these channels. The surface of the virtual substrate tends to have a characteristic undulation called crosshatch and a high density of dislocations which terminate at the surface known as threading dislocations. It is important that the amplitude of the undulations and density of the threading dislocations are kept low in order that devices can be successfully fabricated.

1.3.2 Strained Silicon

A virtual substrate is also needed in order to produce a quantum well for electrons. If silicon is strained under tension, then there is an offset in the conduction band. This offset can be used to produce a 2-dimensional electron gas (2-DEG) with enhanced mobility, record electron mobilities of 8×10^5 $cm^2V^{-1}s^{-1}$ at 15 K [11] and 2830 $cm^2V^{-1}s^{-1}$ at room temperature [12] have been demonstrated in strained silicon

grown on virtual substrates. Recently, there has been major interest in the use of strained silicon for the use of n-MOS *and* p-MOS devices due to their mobility enhancements over bulk silicon (as shown in Figure 1.4). The electron mobility rises rapidly with a maximum of around 80% improvement over silicon. The hole mobility increases more slowly at low strains, but has a larger improvement at very high strains. It must be noted however, that at these large strains the critical thickness of the silicon channel may be too low for practical device applications. As recently as August 2002, Intel Corporation announced that they are to put strained silicon technology into their 90 nm node which is expected to move in to manufacture in 2003.[13] It has not been revealed how Intel grows their virtual substrates, but it is widely believed that a conventional linear grading of SiGe has been used. The next generation of Pentium microprocessors, the Prescott Pentium 4 microprocessor, is also to use this strained silicon technology. Other major figures in the semiconductor market, including IBM, AMD and TSMC, are expected to incorporate strained silicon into their products in the near future.

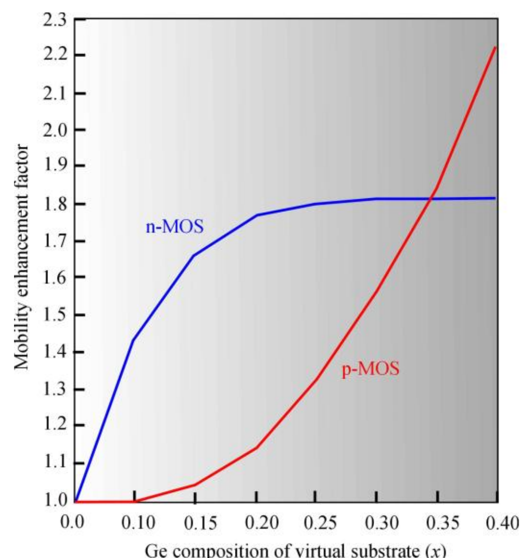


Figure 1.4 The mobility enhancements of p-MOS and n-MOS devices (normalised to bulk silicon).[14,15]

1.3.3 Incorporation of Optoelectronic Materials

Further use of the lattice tuning properties of SiGe virtual substrates is the ability to incorporate III-V semiconductors onto silicon wafers. As can be seen in Figure 1.5 the range of lattice parameters of $Si_{1-x}Ge_x$ includes that of the technologically important III-V semiconductors. It is therefore possible to grow lattice-matched light emitting structures on a properly chosen strain relieved SiGe virtual substrate. This has massive implications in the optoelectronics industry with the possibility of incorporating light emitting components (LED's and laser diodes) with standard Si-CMOS allowing for complete system on chip (SOC) solutions for the communications market.

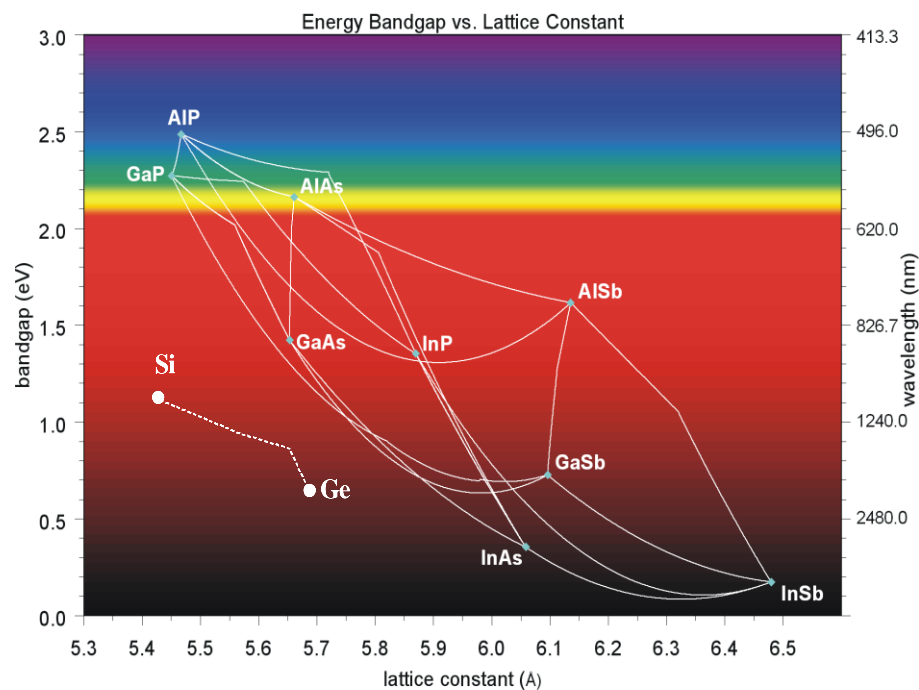


Figure 1.5 Graph of the bandgap energy versus lattice parameter for the main semiconductors. Notice how SiGe spans a range that includes the opto-electronically important GaAs and AlAs.

1.4 Aims of This Work

The aim of this work was to study SiGe virtual substrates grown by solid-source MBE at the University of Warwick. In particular, a new grading technique has been employed in order to control the dislocation networks which are necessary for strain relaxation, leading to a smoother surface and lower threading dislocation densities than conventional grading techniques. The strain relaxation mechanism of this structure has been studied in order that the growth condition could be optimised.

1.4.1 Chapter Summaries

The rest of this work is divided into 6 chapters:

- In **chapter 2** a discussion of the main theoretical concept of virtual substrates will be given. This will include an explanation of crystal structure, the physics of epitaxial growth, the concept of strain relaxation and basic dislocation theory, followed by a review of recent developments in virtual substrates growth.
- In **chapter 3** the main experimental techniques used in this work will be explained. These include the basic operations of the different types of microscopes; optical, TEM, SEM and AFM. A discussion of the phenomena of diffraction is then given with emphasis on its relevance to both TEM and X-ray diffractometry. The use of X-ray diffractometry in determining germanium composition and the state of relaxation is explained followed by details of SIMS and electrical measurements.

- In **chapter 4** an explanation of the new grading technique is given. This new technique is compared to conventional grading techniques using similar growth conditions to produce a $Si_{0.50}Ge_{0.50}$ virtual substrate. The RMS roughness, threading dislocation density and extent of strain relaxation are compared for all three grading techniques. An attempt of reducing the thickness of this virtual substrate is discussed.
- In **chapter 5** the structural and electrical quality of p-channels grown on virtual substrates using this technique will be studied.
- In **chapter 6** a study of the strain relaxation in this new grading profile is undertaken in an attempt to optimise the growth conditions. The quality of an optimised $Si_{0.50}Ge_{0.50}$ virtual substrate is then studied.

2 Theoretical Discussion

In this chapter the main theoretical concepts behind the growth of SiGe virtual substrates will be discussed. A basic review of the crystal structure of semiconductors and the mechanisms of strain relaxation will be given, including an elementary review of dislocation theory and surface roughening. The chapter will conclude with a discussion of the current methods for producing high quality virtual substrates.

2.1 Crystal Structure

The positions of the atoms in a crystal are best described by the *crystal lattice* and *crystal structure*. The crystal lattice is the regular position in real space where every point is equivalent and indistinguishable. For most semiconductors this lattice has cubic symmetry, or in a strained state, tetragonal symmetry. To form the crystal structure a basis (an atom or set of atoms that make up the crystal) must be added to each lattice point in exactly the same orientation. In the case of silicon and silicon-based alloys the crystal lattice is face-centred cubic (FCC). This means that the unit cell (the basic building block of the crystal) is a cube with lattice points at each corner and in the centre of each face making a total of four lattice points per unit cell. The crystal structure of silicon is the diamond structure, this is derived from the FCC lattice by attaching a basis of two silicon atoms at the position (0,0,0) and $\left(\frac{1}{4},\frac{1}{4},\frac{1}{4}\right)$ relative to the unit cell's sides (see Figure 2.1), making a total of 8 atoms

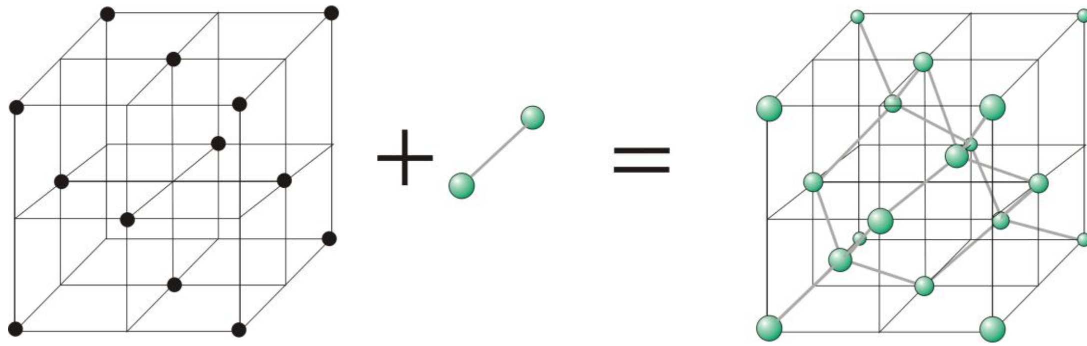


Figure 2.1 The crystal structure of silicon consists of the FCC lattice with a basis of two silicon atoms, forming the diamond structure.

contained within the diamond structure's unit cell. The width of the unit cell, known as the lattice parameter, is usually designated with the symbol $\boldsymbol{a_{Si}}$ for bulk single crystal silicon, and is well known from X-ray measurements to be $\boldsymbol{a_{Si}}$ = 5.431 Å.[16] Germanium's lattice parameter, $\boldsymbol{a_{Ge}}$, has been found to be 5.658 Å,[17] making Ge's lattice parameter 4.2% greater than Si. Each atom in the diamond structure is covalently bonded to 4 nearest-neighbour atoms in a tetrahedral arrangement, the bond length being $\frac{\sqrt{3}}{4}a$.

Germanium is readily soluble in silicon at all compositions and temperatures forming a random alloy, $Si_{1-x}Ge_x$, where x represents the germanium composition and has a value in the range $0<x<1$.[18] In the random alloy the positions of the germanium and silicon atoms are randomly distributed on the lattice sites. In the relaxed, bulk state the lattice parameter of $Si_{1-x}Ge_x$ can be approximated by Vegard's law, a linear interpolation between the lattice parameter of Si and Ge.

$$a_{Si_{1-x}Ge_x} = a_{Si}(1-x) + a_{Ge}x \tag{2.1}$$

Dismukes *et al*[19] and Kasper *et al*[20], have shown that the actual lattice parameter of the $Si_{1-x}Ge_x$ alloy differs slightly from Vegard's law by a small correction factor;

$$a_{Si_{1-x}Ge_x} = a_{Si}(1-x) + a_{Ge}x + 0.02733x^2 - 0.02733x \text{ Å} \tag{2.2}$$

although most authors neglect this small correction for simplicity.

Despite the complete miscibility of Ge and Si, the difference in melting temperatures (938 °C and 1412 °C respectively[18]) and liquidus-solidus lines means that during bulk crystal growth techniques the Ge segregates from the Si on solidification, making the growth of a homogenous $Si_{1-x}Ge_x$ wafer very difficult. As explained earlier, the ability to grow on a $Si_{1-x}Ge_x$ substrate is very attractive for producing electronic devices and has lead to the use of buffer layers grown on top of standard Si substrates in order to terminate the wafer with $Si_{1-x}Ge_x$, with its tuneable lattice parameter. These buffer layers are usually referred to as “virtual substrates” and are the subject of most of this work.

2.1.1 Directions and Miller Indices

Directions and planes in crystal systems are usually designated with reference to the conventional unit cell. This is particularly convenient for cubic systems where the edges of the cubes are used as the axis and the unit of length is usually assumed to be the lattice parameter of the crystal. Specific directions within the crystal are denoted by [*xyz*] for example the furthest corner of the unit cell from the origin is labelled [111], whereas sets of equivalent directions are designated <*xyz*>, so <111> would refer to directions [111], $[\bar{1}11]$, $[1\bar{1}1]$ and $[11\bar{1}]$. The labelling of planes is slightly more complicated.

The crystal contains many sets of planes, these are most often labelled with Miller indices. If a unit cell is drawn in the crystal so that the origin of the cube lies on one plane, then the point at which the next plane intersects the cubic cell is noted.

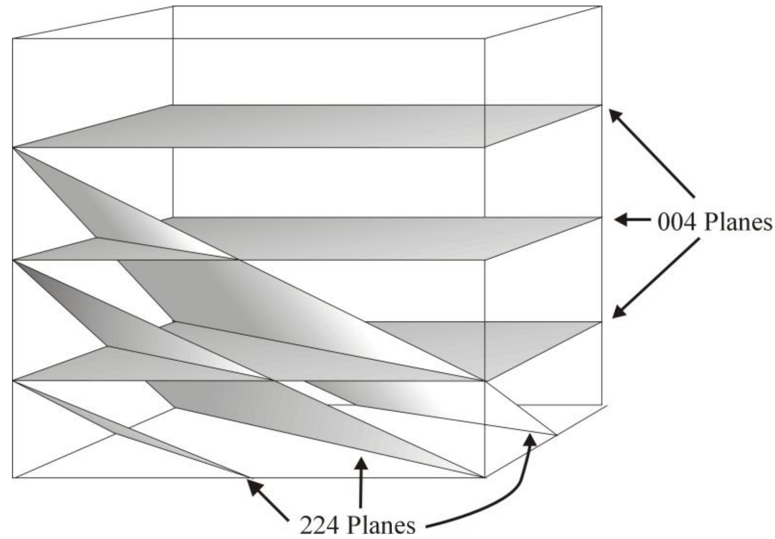


Figure 2.2 Schematic diagram of the cubic unit cell showing the (004) and (224) planes.

The miller indices *h*, *k* and *l* are given by the reciprocals of these intersection co-ordinates. For example, if the next plane intersects the unit cell at positions $(0,0,\frac{1}{2})$, $(0,\frac{1}{2},0)$ and $(\frac{1}{4},0,0)$ the corresponding Miller index would be (224). Specific planes are designated (*hkl*) whereas sets of equivalent planes are designated using {*hkl*}, so {224} would represent the (224), $(\bar{2}24)$, $(2\bar{2}4)$ and $(22\bar{4})$ planes. Figure 2.2 shows a unit cell with the (224) and (004) planes.

In the *orthogonal* crystal systems the planes with Miller index (*hkl*) are perpendicular to the vector $\begin{pmatrix} h \\ k \\ l \end{pmatrix}$ and have inter-planar spacing given by

$$d_{hkl} = \frac{1}{\sqrt{\left(\frac{h}{a_x}\right)^2 + \left(\frac{k}{a_y}\right)^2 + \left(\frac{l}{a_z}\right)^2}} \tag{2.3}$$

where, a_x, a_y and a_z are that lattice parameters (unit cell edge lengths) in the *x*, *y* and *z* directions.

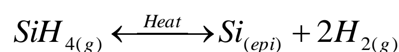
2.2 Epitaxial Growth Techniques.

In order to grow SiGe alloys on silicon substrates it is necessary to use *epitaxial* growth techniques. The term epitaxy describes the layer by layer deposition on a single crystal so that each layer extends the crystal structure of the underlying substrate. In this way the composition, growth rate and growth temperature of the crystal can be controlled so as to optimise the properties of the wafer. The deposition of the same material as the substrate is termed homo-epitaxy, otherwise the term hetero-epitaxy is used.

There are two main epitaxial growth techniques, molecular beam epitaxy (MBE) and chemical vapour deposition (CVD). Theses two techniques differ in the mechanisms by which the material is deposited and the growth conditions obtainable. In the following descriptions the emphasis will be on the growth of Si and SiGe but most of the discussion is equally valid for the growth of other semiconductors.

2.2.1 CVD

Most industrial growth of SiGe employs some kind of chemical vapour deposition (CVD) system. In these techniques, the source of the silicon and germanium adatoms is of a gaseous form, normally silane (SiH_4) and germane (GeH_4) (known as precursors), which react on the surface of the wafer to form the mobile Si and Ge adatoms which incorporate into the crystal.



$$SiH_{4(g)} \xleftrightarrow{Heat} Si_{(epi)} + 2H_{2(g)}$$

The reactants impinge on the wafer, which is heated in order that they decompose into mobile Si and Ge atoms. The sticking coefficients of the reactants (and

contaminants) depend critically on the temperature of the wafer. Early CVD systems where limited by the need to have very high (> 1000 °C) growth temperatures in order that epitaxy of sufficient quality could be attained,[21] this is unsuitable for most device applications due to the thermal budget imposed by strain relaxation and the diffusion of Ge in Si. Later developments in vacuum systems, chemical purity of the reactants and wafer cleaning techniques have allowed growth temperatures in CVD systems to be reduced to as low as 450 °C.[22]

The most common type of CVD is ultra high vacuum CVD (UHVCVD).[23] In these growth systems the background pressure is reduced to ~10^{-10} mbar and so reducing the incorporation of contaminants. The reactants typically have a partial pressure of ~10^{-4} mbar which is intermediate between viscous and molecular flow regimes. The growth rates of the respective reactants are controlled by their partial pressures and the growth temperature. The whole growth chamber is heated and hence a high degree of thermal uniformity is achieved, although this leads to deposition on the sidewalls and slow temperature response. The uniformity of deposition is very high due to the rapid flow of reactants and temperature uniformity and a high throughput is possible due to multi-wafer loading.

Doping is achieved in UHVCVD by the introduction of B_2H_6 (p-type) and PH_3 or AsH_3 (n-type) into the growth chamber. p-type doping is straight forward, but due to the segregation and diffusion of PH_3 and AsH_3 n-type doping is more difficult, leading to poor doping profiles.

There are several variations of CVD systems in common use, of these plasma enhanced CVD (PECVD)[24] and gas source MBE (GS-MBE)[25] are two. In PECVD the cracking of the reactants occurs in an inert plasma (usually argon, but sometimes helium) near to the surface of the wafer. Since the precursor Si and Ge atoms are

produced in the plasma, the temperature of the wafer can be reduced and high growth rates are achievable (over 5 nms^{-1} by LEPECVD[26]). Despite its name GS-MBE is more closely related to CVD systems than MBE. The growth chamber is similar to a MBE system, a cold walled UHV chamber with a heated substrate, but the reactants are silane and germane which are introduced at low pressure into the chamber (hence the name gas-source MBE). However, the deposition reaction at the wafer surface is the same as in UHVCVD, the only difference being the lower partial pressure of the reactants so that the mass flow is in the molecular regime.

As well as high throughputs and high deposition rates in the CVD systems, another advantage is the possibility of selective epitaxial growth (SEG);[27] this is the growth of SiGe in selected areas on the wafer which are defined by windows in an oxide layer deposited on the substrate. This can be used for the incorporation of SiGe devices on the same chip as conventional Si-CMOS. The selectivity of the growth is achieved by using chlorinated reactants (e.g. SiH_2Cl_2) which form HCl at the surface of the wafer. This effectively etches any epitaxy from the oxide surface, whilst allowing growth on the exposed substrate in the oxide windows.

2.2.2 MBE

The samples for the work carried out in this study were grown by MBE in a VG Semicon V90S system which is illustrated in Figure 2.3. In MBE the epitaxial material arrives at the substrate in the form of an atomic/molecular beam. In the V90 system this is achieved by melting a charge of pure silicon or germanium using an electron beam, since the source is a solid charge of silicon or germanium this form of MBE is known as solid source MBE (SS-MBE).[28] The silicon is atomised by the

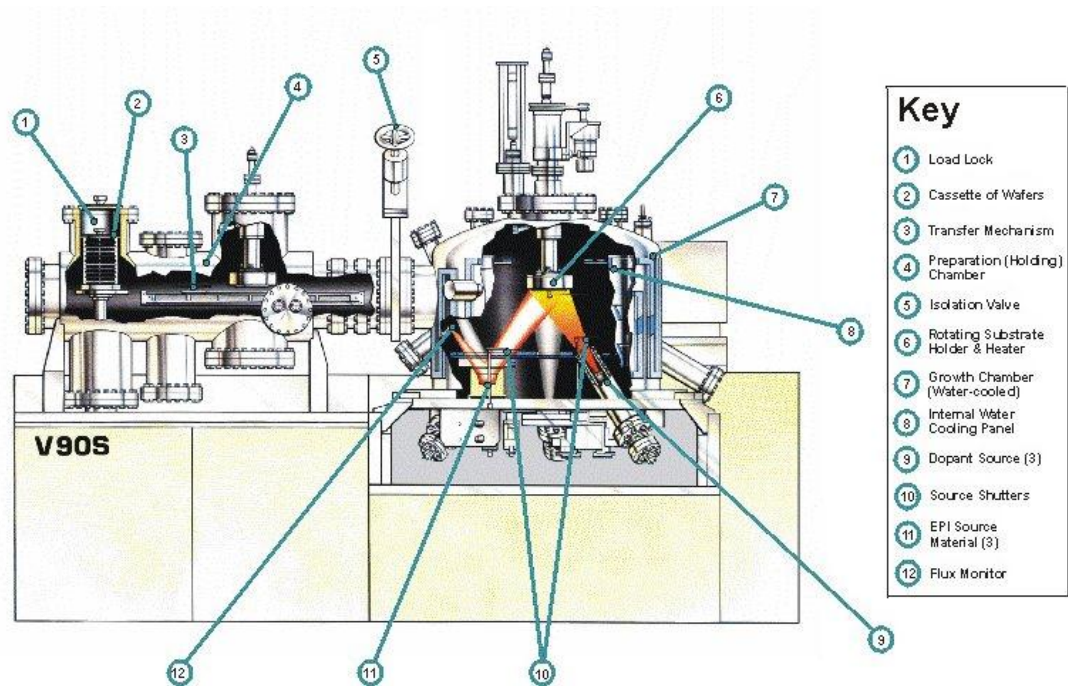


Figure 2.3 Schematic diagram of the V90S MBE growth system.

electron beam and effuses away from the source in all directions. The flux of the silicon and germanium is controlled by the intensity of the electron beam, which has to be carefully calibrated regularly since the flux rates change as the charges deplete. Typical flux rates are about 1 $Ås^{-1}$ but flux rate of 0.1 – 10 $Ås^{-1}$ are possible. This low flux rate enables the heterostructures to be controlled to a single atomic layer. To ensure uniform coverage the wafer is rotated at approximately 5 rpm below a substrate heater. The heater can control the substrate temperature from room temperature to around 1000 °C, this is in contrast to most forms of CVD where the temperature needs to be high for the surface reaction which leads to deposition to occur. It is very important that the entire growth system is be kept under ultra high vacuum (UHV) conditions (>10^{-9} mbar). This is necessary to ensure that the atomic beam of silicon and germanium can reach the substrate without colliding and also to prevent any contaminants from incorporating onto the substrate which is very

reactive once the oxide has been completely stripped. It is simple to show from the kinetic theory of gases that the flux of contaminants, z, is given by[29]

$$z = \frac{p}{\sqrt{2\pi m k_B T}} \qquad (2.4)$$

where p is the partial pressure of the contaminant, m is the mass of a molecule, k_B is the Boltzmann constant and T is the temperature of the gas in Kelvin.

Assuming that all impinging contaminants stick to the substrate, a typical contaminant, for example CO, at room temperature would have a flux rate of ≈ 5 $Ås^{-1}$ for a reasonably low pressure of 10^{-6} mbar. In order that the contamination flux is much less than the Si flux the pressure needs to be reduced to as low as possible, usually at least 10^{-9} mbar.

Before growth the silicon substrate needs to be cleaned in a combination of etches known as a modified RCA clean [30] which removes organic and metal contaminants. After the RCA clean the substrate is dipped into a 2% HF solution and spun dry to remove any oxide and terminate the surface with hydrogen which passivates the surface. The substrate is then quickly loaded into the preparation chamber which is evacuated. The wafer is then moved on a mechanical transfer mechanism to the growth chamber. Immediately prior to growth the substrate is raised to approximately 890 °C in order that the hydrogen and any remaining oxide desorbs.

2.2.3 Growth Kinetics

Which ever growth method is employed the outcome is to leave mobile atoms known as *adatoms* on the surface of the substrate. These adatoms are free to move on the surface until they are chemically bonded to the substrate and incorporated into

the crystal structure. The surface mobility of the adatoms is characterised by the migration length, λ, the average distance that an adatom moves until it is incorporated. The value of λ depends on many factors including the migrating species, the temperature and the crystallographic orientation of the surface. Since the chemical bond strength of the Ge atoms is lower than Si, the migration length is greater for these adatoms. The migration length will depend critically on the size of the energy barriers between adjacent surface sites. The closer packed crystallographic planes have a lower barrier and hence the migration lengths are longer for these planes. The common planes found in the Si system have increasing migration lengths in the order (001) < (011) < (111).[31] Since, at higher temperatures more energy is available to overcome the barriers between neighbouring surface sites, the migration length increases with temperature.

The surface of a Si substrate is not completely atomically flat, but consists of many terraces separated by atomic steps with average separation, *l*. Figure 2.4 shows an AFM image of an "as-received" standard silicon wafer which shows clearly these terraces with atomic steps aligned along the <110> directions. The potential well associated with the surface step is deeper than a surface site away from a step (due to the extra bonds), consequently the adatoms are preferentially incorporated at these steps (Figure 2.5). The growth of the epitaxial layer usually proceeds by the extension of these terraces across the surface if $\lambda > l$. Commercially available Si (001) substrates are usually cut with an intentional off-cut towards the (110) direction of ~ 0.4°. This leaves the surface with a high number of surface steps (~ 10^5 cm^{-1}) which increases the adatom incorporation and encourages 2-D growth by terrace extension.

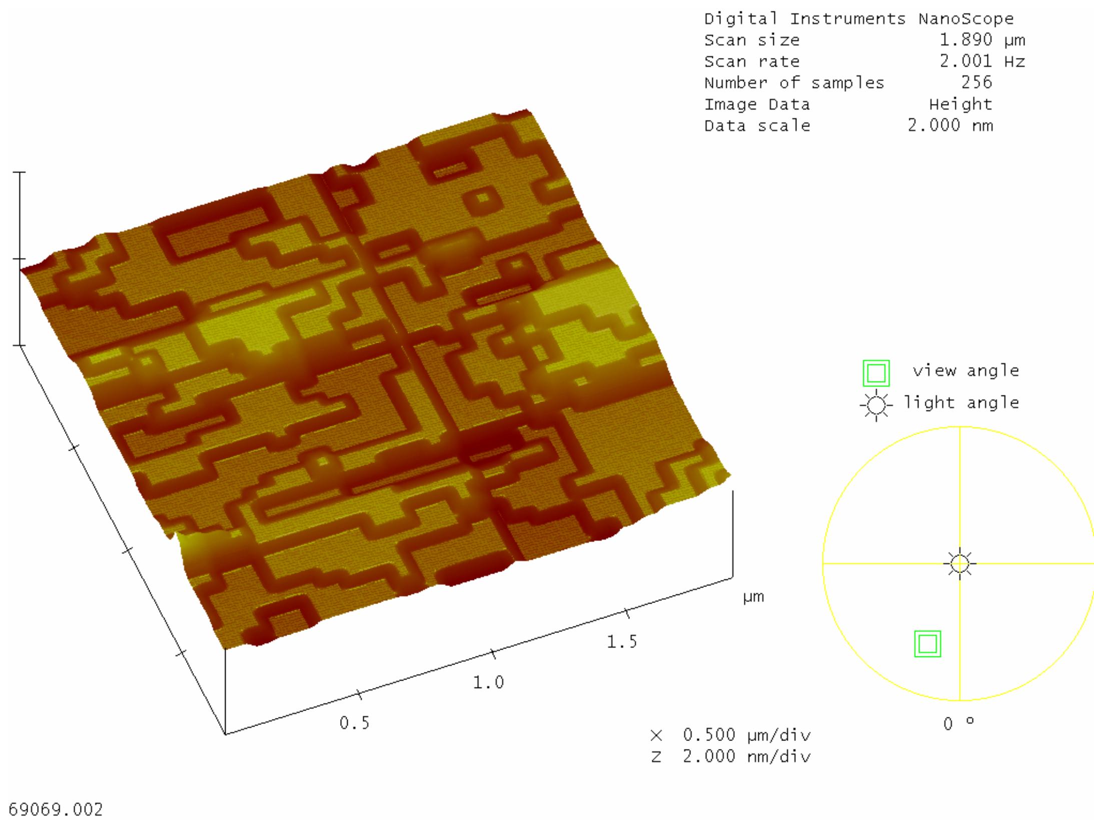


Figure 2.4 AFM image of a (001) Si wafer taken at Warwick. Notice the terraced nature of the surface with many islands and pits. The vertical scale is 2 nm.

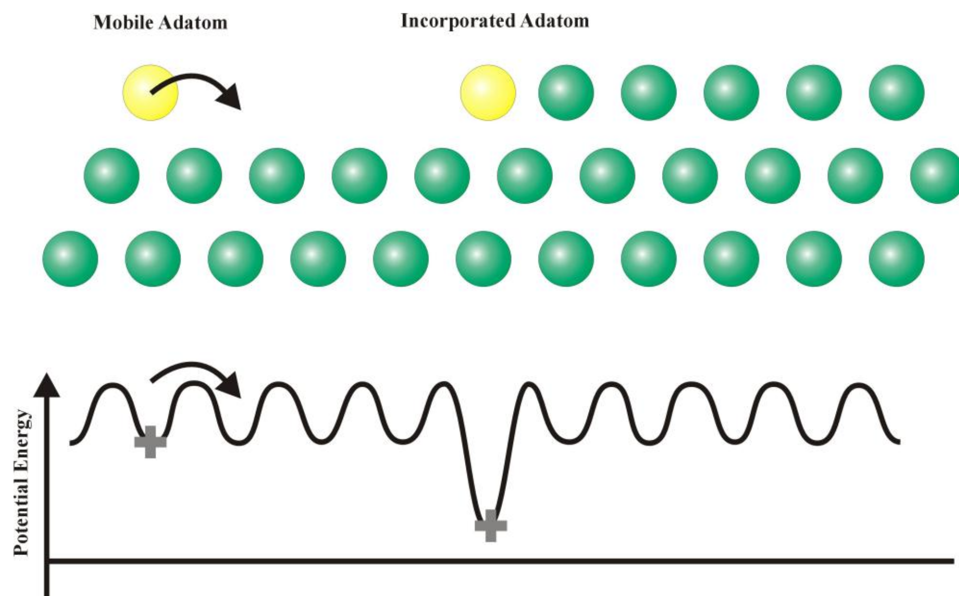


Figure 2.5 Schematic diagram showing the potential well at a surface step where adatoms preferentially incorporate.

Not all of the mobile adatoms will make it to an atomic step before they are incorporated into the crystal. These adatoms form their own steps and local deepening of the potential well, causing more adatoms to be captured and form clusters of atoms on the surface known as islands. Other surface features include pits where terraces have been incompletely covered leaving a small gap in the layer. The growth of these islands and pits leads to a 3-D growth regime which is to be avoided in the growth of thin films. To minimise the 2-D to 3-D transition the migration length, λ, should be made as large as possible to give the adatoms the chance to reach the energetically lower positions at the surface steps. Alternatively, the degree of off-cut of the substrate may be increased (up to 6°) in order to reduce the average step separation. Figure 2.6 shows a schematic representation of the epitaxial growth of a semiconductor on a substrate with terraces, showing how the incoming adatoms move around the flat surface until they are incorporated at a surface step. Growth pits and islands are shown leading to the first stages of 3-D growth.

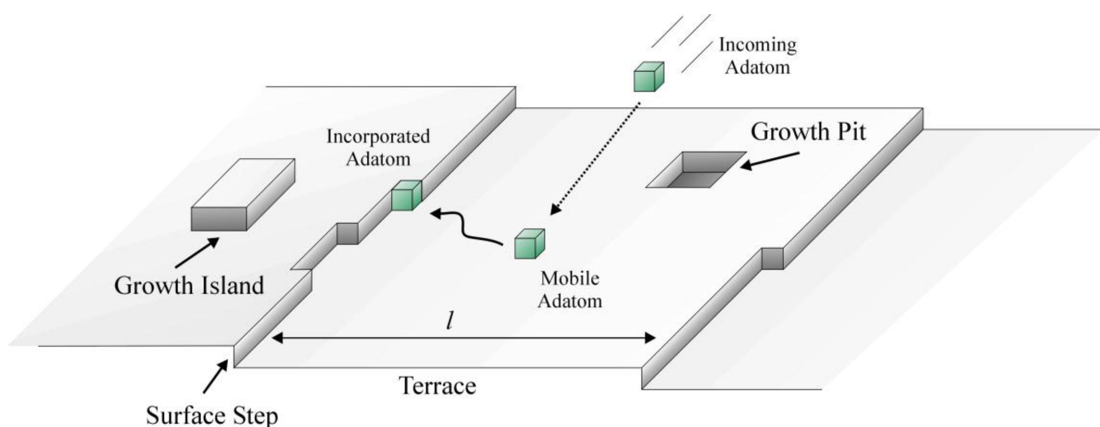


Figure 2.6 Diagram of the growth of a semiconductor on the surface of a substrate. The incoming adatom (from CVD or MBE growth techniques) form mobile adatoms on the surface, which may diffuse until they are incorporated at a surface step, or form a growth island.

There are three modes of epitaxial growth:-

1) Layer by layer growth or **Frank-van der Merwe**[32] (FvdM) mode where the epitaxial layer continues to grow in a 2-D film.
2) 3-D islanding or **Volmer-Weber**[33] (VW) mode where the epitaxial layer grows by clustering into 3-D islands on the surface of the substrate.
3) Layer by layer/3-D islanding or **Stranski-Krastanov**[34] (SK) mode where the epitaxial layer starts to grow in a layer by layer manner, but after several monolayers, growth reverts to the 3-D clustering of the VW mode.

The mode by which the epitaxy continues depends on the balance of free energies of the surfaces and interface involved.[35] The substrate surface free energy (σ_s), the film surface free energy (σ_f) and the interfacial free energy (σ_i) are balanced along with the elastic strain energy of the epitaxial layer. In the absence of strain the FvdM growth mode is favoured when $\sigma_f + \sigma_i < \sigma_s$ and a 2-D layer is formed. During hetero-epitaxy the strain in the layer causes an increase in the value of σ_f and σ_i and consequently the condition for FvdM growth mode may not be fulfilled. If this is the case from the start of the deposition then the 3-D (VW) growth mode is favoured. However, if $\sigma_f + \sigma_i < \sigma_s$ at the start of deposition but the increase of strain causes this condition to become untrue during epitaxy then the growth mode will switch from a 2-D to 3-D mode (SK mode). These thermodynamic processes will govern the growth of the epitaxial layer, however this is only possible if the kinetics of the growth allow thermodynamical equilibrium to be attained. This will be the case for high growth temperature (as usually used in CVD), but may not be the case for the lower temperatures possible in MBE.

2.3 Defects, Dislocations and Strain

Crystals inevitably contain defects, these are regions within the crystal where the atoms do not coincide with the regular positions expected in a crystal. There are many types of defects that may form. These may be classified as zero-dimensional, 1-dimensional, 2-dimensional or 3-dimensional.

- **Zero-dimensional.** These are defects that occur at a point, hence these are usually referred to as point defects. There may be an atom that is missing from the lattice (a vacancy) or an atom that is in between lattice points (an interstitial). These may also occur in pairs (Frenkel defect). As there is a finite energy needed to form these defects, then there will always be a thermodynamical equilibrium concentration at a given temperature.
- **1-dimensional.** These types of defects are known as dislocations and are lines within the crystal that separate regions that have slipped and regions that have not. These are a very important type of defect in the physics of strain relief in epitaxial semiconductors and will be discussed in detail later.
- **2-dimensional.** These are usually termed stacking faults, regions where the ordering of planes within the crystal has been disrupted and consequently a 2-dimensional defect occurs in these planes.
- **3-dimensional.** These are large regions within the crystal where the regular crystallinity has been interrupted, this may be due to precipitates of impurities or large voids (volumes within the crystal with many missing atoms).

2.3.1 Strain

When a semiconductor is grown epitaxially on a substrate it may become strained. This occurs if there is a mismatch in lattice parameter of the epitaxial layer and the underlying substrate. As the epitaxial layer tries to continue the layer by layer growth it has to strain itself onto the lattice spacing of the substrate. A misfit parameter, *f*, may be defined as

$$f = \frac{a_{layer} - a_{substrate}}{a_{substrate}} \tag{2.5}$$

where a_{layer} and $a_{substrate}$ are the (bulk) lattice parameters of the epitaxial layer and substrate respectively. Germanium grown epitaxially onto a silicon substrate has a mismatch of 4.2% (i.e. f=0.042). Hence, a SiGe alloy grown onto a silicon substrate has a mismatch varying (approximately) linearly with Ge composition from 0 to 4.2%.

The mismatch in an epitaxial layer causes the layer to be bi-axially strained in the growth plane. The magnitude of the strain, ε, is equal to (but opposite in sign to) the misfit parameter, *f*, for a fully strained (or pseudomorphic) layer. When $a_{layer} > a_{substrate}$ ε is negative and the layer is compressively strained and when $a_{layer} < a_{substrate}$ ε is positive and the layer is tensionally strained. This strain leads to an elastic energy stored within the epitaxial layer which according to continuum elasticity theory has energy density given by

$$E_{elastic} = \frac{E_Y \varepsilon^2}{1 - \nu} \tag{2.6}$$

where E_Y is the Young's modulus and ν is the Poisson's ratio. As the thickness of the epitaxial layer increases the stored energy can become quite large for even

modest mismatches, for $Si_{0.75}Ge_{0.25}$ $E_{elastic}$ is of the order of 2×10^7 Jm^{-3}. Eventually the stored energy in the epitaxial layer becomes too large for the bonds within the crystal to hold the atoms together and it becomes energetically favourable for dislocations to form to relieve the strain. The thickness of the layer at which dislocations are favoured is known as the equilibrium critical thickness, h_c, and will be discussed later.

2.3.2 Dislocations

When a crystal layer is strained it may relieve the strain by forming dislocations. These can best be thought of as the result of making a cut in the crystal and displacing the two sides of the cut relative to one another (usually by a lattice vector) and then reattaching the two sides of the cut. The result is a line along the base of the cut where the atoms in the crystal do not follow the regular pattern of the rest of the crystal. It must be stressed that dislocations are not actually formed in this way, but this is a convenient way to visualise the structure of a dislocation.

2.3.2.1 Burgers Vector

A dislocation is characterised by its line direction and its *Burgers vector.*[36] The Burgers vector of a dislocation is calculated by performing a Burgers circuit around the dislocation in the following manner. To form a Burgers circuit in a crystal a clockwise circuit is made in a section of a perfect part of the crystal, this is a region in the crystal which contains no defects. The circuit consists of “jumps” between atomic positions within the crystal as in Figure 2.7(a). Then the same circuit, with identical jumps between atomic positions, is performed in a region of the

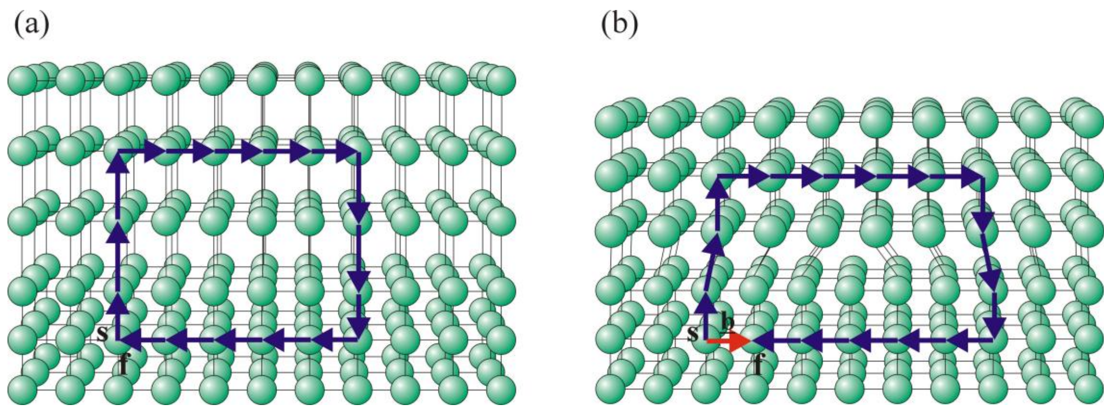


Figure 2.7 The Burgers vector is determined by first making a right-handed circuit in a perfect part of the crystal (a). Then the *same* circuit is made around the dislocation (b). The Burger's vector (red) is defined as the vector needed to close the circuit from start to finish (RH/SF convention).

crystal which contains the dislocation as in Figure 2.7(b), ensuring that the circuit encompasses the dislocation. There will be a resulting *closure failure* of the circuit, i.e. the start of the circuit (marked **s** for start) will not coincide with the end of the circuit (marked **f** for finish) as a result of the dislocation. The vector needed to complete the circuit form start to finish is known as the Burgers vector. The line direction has been chosen arbitrarily to be into the plane of the paper. It is not important which direction is taken to be positive for the line direction, as long as the Burgers circuit is made following a right-hand screw rule *and* the convention of the Burgers vector being from start to finish is used. This convention of using a right-hand screw circuit and a start to finish Burgers vector is known as the RH/SF convention.

In an elastic continuum, the Burgers vector can be defined more formally. If $\boldsymbol{u}(x, y, z)$ represents the elastic displacement vector at a point (x, y, z) then the burgers vector is defined by

$$\boldsymbol{b} = \oint \frac{d\boldsymbol{u}}{dl} dl \tag{2.7}$$

where the integral is a closed line integral performed around the dislocation in a right-hand screw direction.

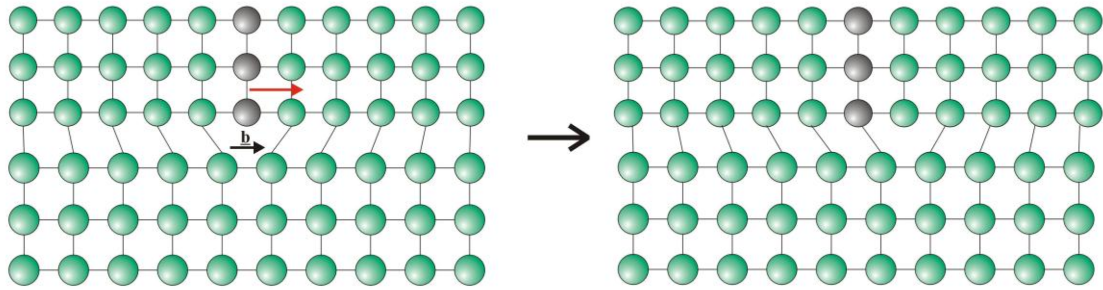
Dislocations are generally classified by the relationship between the line direction and the Burgers vector. When the line vector and Burgers vector are perpendicular the dislocation is said to be an edge dislocation, when they are parallel the dislocation is said to be a screw dislocation. If the two vectors are neither parallel nor perpendicular the dislocation is a mixed dislocation and will have both edge and screw character. The importance of the Burgers vector is that it is a conserved quantity along any dislocation. Although a dislocation may change its direction along its length the Burgers vector will always be the same, of course the type of dislocation will change as the direction changes.

2.3.2.2 Dislocation Motion

Any crystal system will have a preferred plane for slip to occur. This is usually the closest packed plane, as this will have the lowest energy barrier for a row of atoms to move past each other (Peierls barrier[37]). In the diamond crystal structure this is the {111} planes and the Burgers vectors are usually of the form $a/2\langle 110\rangle$.

A dislocation can move by two mechanisms, glide and climb as illustrated in Figure 2.8. Glide is the conservative motion of the dislocations, whereby the dislocation moves by jumping one burgers vector away from its original position. This only needs the local arrangement of bonds to be changed and so no mass transport is needed, consequently this is the easiest form of motion and will be the form chosen if possible. Glide can occur in any plane that contains the line direction and the Burgers vector, hence for an edge dislocation, where the two vectors

(a) glide



(b) climb

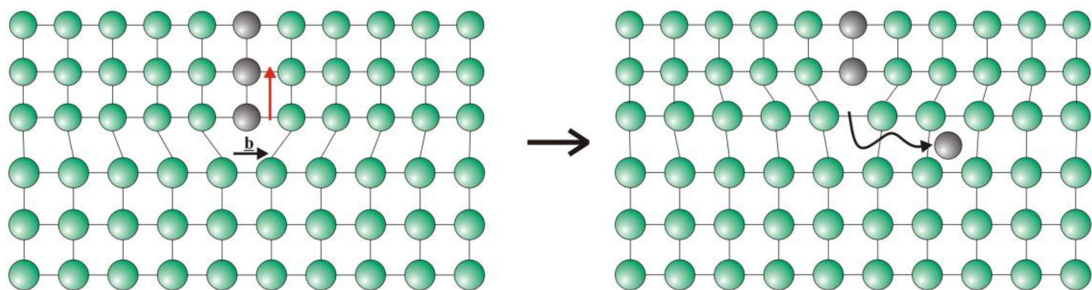


Figure 2.8 Dislocations can move by glide or climb. In glide (a) the dislocation moves in its glide plane by a local rearrangement of bonds. In climb (b) the dislocation can move out of its glide plane by the diffusion of atoms away from the dislocation. In both diagrams above the extra half-plane of the original edge dislocations is shown in black in order to show the movement of the dislocation.

are perpendicular, there is only one possible glide plane for a given dislocation. For a screw dislocation the Burgers vector and line direction are parallel, so any plane containing the dislocation is a possible glide plane.

Climb is the motion of a dislocation out of its glide plane. To do this the dislocation moves in a direction that is not parallel to the Burgers vector, consequently there is a need for atoms to be added or removed from the crystal, requiring mass transport. This mass transport is usually in the form of the production or annihilation of point defects (interstitials or vacancies), which must diffuse away from or towards the dislocation. Hence the motion of a dislocation by climb is very temperature dependent and only occurs at high temperatures and when glide is not possible or has been blocked.

2.3.2.3 The SiGe System

In an epitaxial SiGe layer the slip planes {111} intersect the growth plane (001) along the [110] and $[1\bar{1}0]$ directions as shown in Figure 2.9. Consequently, the Burgers vector and line direction have an angle of 60° (mixed dislocation) or 90° (pure edge dislocation) and are called 60° and 90° dislocations respectively. The 90° dislocations are more efficient at strain relief and so are more predominant at high mismatches, but because the burgers vector does not lie in the preferred glide plane for this crystal system, these dislocations can only move by climb and so only occur at high temperatures.[38] The 60° dislocations have Burgers vectors which do lie in a preferred glide plane and so can easily move by glide. Therefore, in lower mismatched systems and at lower growth temperatures these dislocations are predominant.

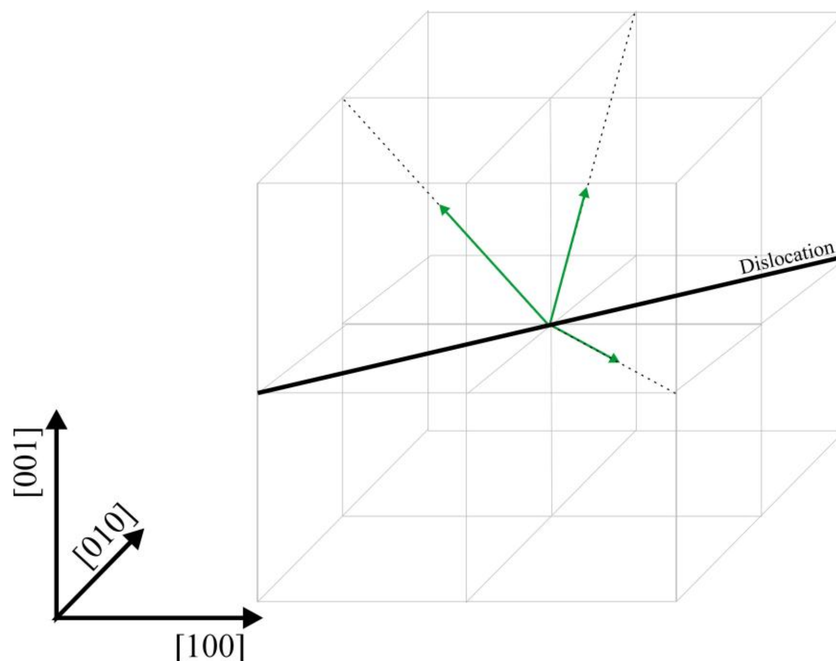


Figure 2.9 The dislocation in silicon and silicon based alloys have line directions along the <110> orientations. The preferred Burgers vectors lie parallel with the <110> directions and consequently have an angle of 60° or 90° to the line direction.

2.4 Relaxation

To produce high quality virtual substrates the buffer layer needs to be in a high state of relaxation. In order that the quality of the virtual substrate can be maximised, a fundamental understanding of the mechanism by which epitaxial layers relax is needed.

2.4.1 Energy of a Dislocation

Because a dislocation introduces a local strain in the crystal structure, due to the deformation of the crystal, a dislocation has an associated energy. The exact form of the energy is complicated because of the discrete nature of the atomic positions. However, continuum elasticity can be used to approximate the form of the energy if the core of the dislocation, where the atomic nature is most apparent, is not included in the calculations but added as an empirical parameter.

The energy per unit length, E_L of a screw dislocation is given by

$$E_L = \frac{Gb^2}{4\pi}\left[\ln(\frac{R}{r_o}) + 1\right] \tag{2.8}$$

and that for an edge dislocation

$$E_L = \frac{Gb^2}{4\pi(1-\nu)}\left[\ln(\frac{R}{r_o}) + 1\right] \tag{2.9}$$

where G is the shear modulus, b is the magnitude of the Burgers vector, ν is the Poisson's ratio of the material, r_o is the (empirical) radius of the core and R is the outer cut-off radius.[39] This cut-off radius represents the distance to the nearest surface or the average dislocation spacing, whichever is smaller. The final term in the square brackets is used as a convenient form for the core energy, here it has the value

1 but more generally it can be given an empirically determined value. For a general dislocation with an angle θ between its Burgers vector and line direction the line energy is given by

$$E_L = \frac{Gb^2(1-\nu\cos^2\theta)}{4\pi(1-\nu)}\left[\ln(\frac{\alpha R}{b})\right] \quad (2.10)$$

Here, the value α accounts for the core energy and is usually given a value between 1 and 4.

The line energy of the dislocation is not very sensitive to the exact value of r_o or R, which is fortunate since these are very hard to determine. From these equations for the line energy several important points should be highlighted. Firstly, since the energy is proportional to the length of the dislocation, the dislocation tends to make itself as straight as possible, i.e. the dislocation has a line tension. It is sometimes useful to use an analogy with an elastic band. Secondly, the energy of a screw dislocation is less than the energy of an edge dislocation with the same Burgers vector. This implies that a dislocation will arrange itself to have as large a screw component as is possible. Lastly, the energy for all dislocations is proportional to the square of the Burgers vector, so dislocations with small Burgers vectors are preferred over dislocations with large Burgers vectors. The smallest lattice vectors in the diamond structure are the <110> vectors, hence perfect dislocations in these crystal systems nearly always have Burgers vectors equal to one of these lattice vectors. A dislocation which has a large Burgers vector will be unstable against dissociation into two or more dislocations with smaller Burgers vectors. In particular, a perfect dislocation (where the Burgers vector is a lattice vector) may decay into two partial dislocations (where the Burgers vector is less than a lattice vector). The region between two partial dislocations must contain a stacking fault as shown in

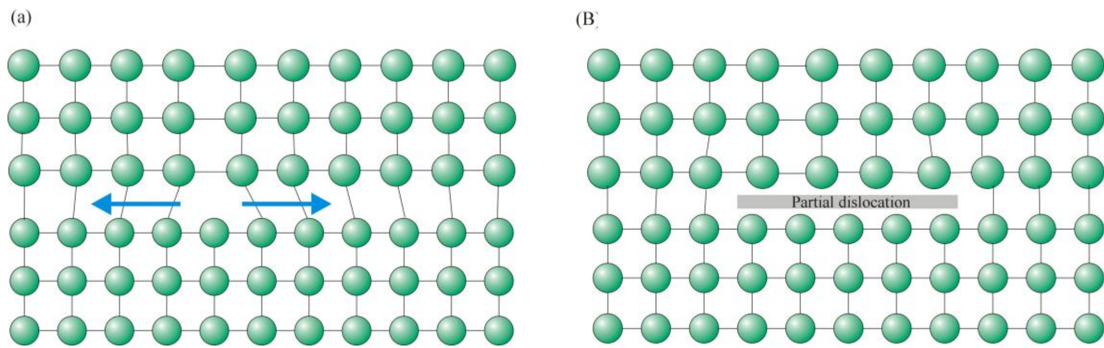


Figure 2.10 A perfect dislocation (a) is unstable against dissociation into an extended partial dislocation (b), by spreading the misfit over a larger area the energy of the dislocation can be reduced.

Figure 2.10. The exact size of the stacking fault is determined by the energy balance between the line energy of the dislocations, the mutual repulsion between the dislocations and the areal energy density of the stacking fault.

In the SiGe system the 60° dislocations with burgers vectors *a*/2<011> are unstable against dissociation into two *Shockley* partials with Burgers vector *a*/6[121] and *a*/6[$\bar{1}$12].[39] However, the distance between these dislocations is generally only a few nm in SiGe[40] so this dissociation is generally ignored and the dislocations are treated as 60° dislocations with Burgers vectors *a*/2<011>, this will also be the case in the present study.

2.4.2 Critical Thickness

As a strained epilayer continues to grow, the strain energy increases in proportion to the thickness of the layer (refer to equation (2.6)). This increase in energy cannot continue indefinitely and eventually the layer will begin to relax. This may be by the introduction of interfacial misfit dislocations or by surface roughening

of the epilayer surface. In the case of plastic relaxation by dislocation formation there exists a *critical thickness*, h_c, above which the layer is unstable against relaxation.

The idea of a critical thickness was first discussed by Frank and Van der Merwe[41] as early as 1949, although only a one dimensional case was considered. Later, Matthews and Blakeslee[42] formulated the most quoted form of the critical thickness models, known as the Matthews and Blakeslee (MB) model. In this approach the effective stress acting on a pre-existing threading dislocation is balanced against the line tension of the dislocation itself. This is equivalent to balancing the energy loss through the relief of strain to the self energy of the dislocation.

Because the strain is in the plane of the hetero-interface, the strain relieving effect of the dislocation is determined by the *effective* Burgers vector b_{eff}. This is the projected part of the Burgers vector onto the interfacial plane which is perpendicular to the line direction. The effective Burgers vector is defined as

$$b_{eff} = b\cos(\lambda) \tag{2.11}$$

where λ is the angle between the Burgers vector and the direction within the interfacial plane which is perpendicular to the line direction.

The strain energy, E_r, relieved by a dislocation is given by

$$E_r = \sigma b_{eff} A \tag{2.12}$$

where σ is the stress in the layer and A is the area of the slipped plane perpendicular to the strained direction. From standard elasticity theory the strain, ε, is related to the stress by

$$\sigma = 2G\varepsilon\frac{(1+\nu)}{(1-\upsilon)} \tag{2.13}$$

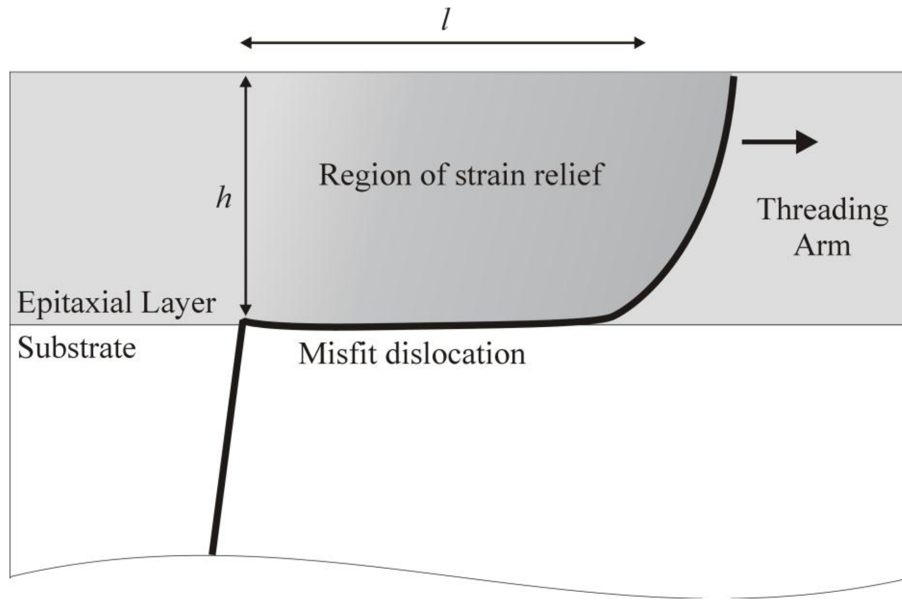


Figure 2.11 As a pre-existing threading dislocation glides, it leaves a misfit dislocation at the interface. This misfit dislocation relieves the strain of the layer above it (darker shaded region). However, the dislocation has a self energy which must be balanced with the reduction of strain energy.

So the strain energy relieved by the presence of a dislocation of length, l, in an epitaxial layer of thickness h is given by

$$E_r = 2G\varepsilon\frac{(1+\nu)}{(1-\nu)}b\cos(\lambda)hl \tag{2.14}$$

The point at which dislocations are energetically favoured in the epilayer is defined as the point at which $E_r > E_L$. So the critical thickness, h_c, is defined as the point at which $E_r = E_L$ giving

$$\frac{Gb^2(1-\nu\cos^2\theta)}{4\pi(1-\nu)}\left[\ln(\frac{\alpha h}{b})\right]l = 2G\varepsilon\frac{(1+\nu)}{(1-\nu)}b\cos(\lambda)hl \tag{2.15}$$

Rearrangement of this equation leads to the Matthews and Blakeslee critical thickness

$$h_c = \frac{b(1-\nu\cos^2(\theta))}{8\pi(1+\nu)\varepsilon\cos(\lambda)}\ln(\frac{\alpha h_c}{b}) \tag{2.16}$$

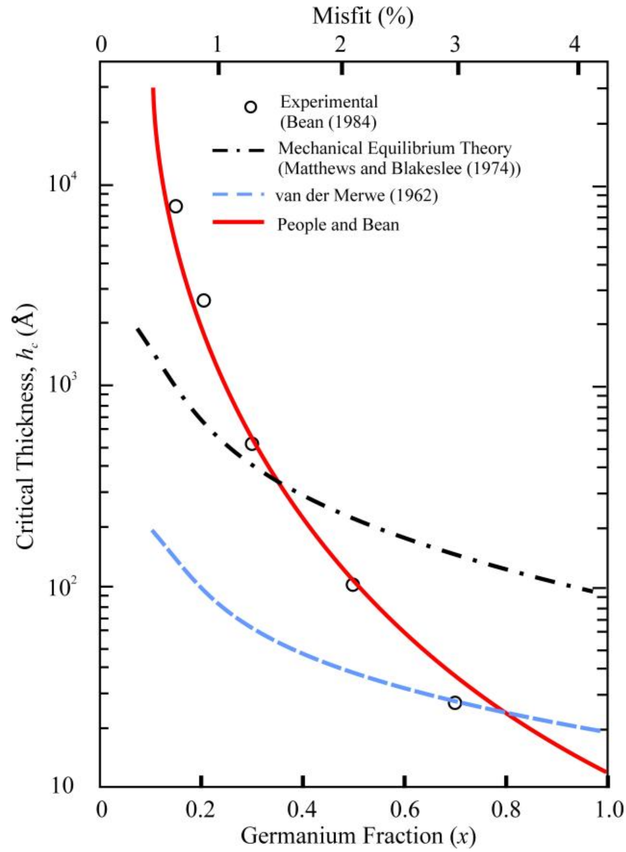


Figure 2.12 Critical thickness as a function of lattice misfit and Ge concentration for a single $Si/Si_{1-x}Ge_x/Si$ heterostructure (adapted from Ref. 43).

which is a transcendental equation, i.e. it has no algebraic solutions, although it is easy to solve numerically and is graphed in Figure 2.12. The following parameters for the SiGe system were used $G = 64$ GPa, $b = 3.9$ nm, $\varepsilon = 0.42x$, $\nu = 0.26$, $\alpha = 2$, $\cos(\theta) = 0.5$ and $\cos(\lambda) = 0.5$ (60° dislocations).

The MB critical thickness is based on thermodynamical equilibrium and does not include the kinetics of dislocation formation. Consequently, the MB critical thickness only provides a lower limit of h_c below which dislocation formation is energetically unfavoured. It is possible to grow a fully strained layer in excess of the MB critical thickness if there is no suitable mechanism for the dislocation formation, this layer is then said to be *metastable*. The value of h_c determined by the MB criterion does compare well with experimental data for growth or anneal

temperatures > 750 °C. In this regime the dislocation density is said to be thermodynamically limited. At lower temperatures h_c can be much greater than that predicted by the MB criterion due to kinetic barriers to dislocation formation. People and Bean[43] attempted to explain the experimental critical thickness observed at lower temperatures. Their approach considered the energy of the dislocation to be concentrated in a narrow core approximately $5b$ from the dislocation and fitted the model to experimental data for MBE growth at 550 °C, their results are plotted in Figure 2.12. There have been several other attempts to improve the accuracy of the MB model for the critical thickness[44,45], although the MB model is still the most widely quoted.

2.4.3 Nucleation

The bottleneck in strain relaxation at low temperatures is due to the kinetic barriers to dislocation formation. In modern high quality silicon substrates the typical threading dislocation density is of the order of 1 m^{-2}, Fitzgerald *et al*[46] have suggested that typical threading dislocation densities needed to relax a SiGe buffer layer are ~10^5 - 10^6 cm^{-2}. Clearly there are insufficient dislocations in the silicon substrate to allow the buffer layer to relax, and so some form of dislocation nucleation is necessary. The nucleation of dislocations can be classified into three main types.

1) Homogenous Nucleation
2) Heterogeneous Nucleation
3) Dislocation Multiplication.

Homogenous nucleation is the spontaneous formation of a dislocation within the semiconductor material. This is usually in the form of a dislocation half loop which

expands from the surface until it reaches the hetero-interface. This mechanism has a high activation energy at low mismatch strains[47] and so is only a significant factor at high mismatch or when other nucleation mechanisms are unavailable

Heterogeneous nucleation is the formation of a dislocation from a source that is extrinsic to the crystal. This may be from a metallic or carbide precipitate within the layer, a surface contaminant due to poor substrate cleaning or other foreign material. Hull[47] has shown that silicon inclusions have acted as nucleation sources in their material. These inclusions were thought to have flaked from silicon deposited on the wall of the growth chamber.

Multiplication is the mechanism by which pre-existing dislocations can act as sources for more dislocations. There are several proposed mechanism for dislocation multiplication, for example the Hagen-Strunk mechanism [48] and the "diamond defect", [49] but the most widely quoted being the modified Frank-Read (MFR) mechanism proposed by LeGoues *et al*[50]. This multiplication mechanisms will be discussed in detail later.

The main source of dislocation nucleation is dependant on the conditions during growth of the epitaxial layer. For small layer thicknesses there are insufficient pre-existing dislocations for the multiplication mechanisms to play a major part and homogenous nucleation barriers will be too high. In this regime, the dislocation formation is limited by the heterogeneous nucleation rate.[47] For very high strains, (~0.02) homogeneous nucleation from surface undulations or surface steps is dominant. At low mismatch strain, the activation energies for dislocation nucleation are high. The lower activation energy for multiplication mechanisms at low mismatch means that if the layer is of sufficient thickness to allow pre-existing

dislocations to glide and interact, this will be the most prevalent mechanism for dislocation formation.

2.4.4 Roughening

Another mechanism by which the strain in an epitaxial layer can be relieved is *surface roughening*. In thermodynamic equilibrium the total free energy of the system must be at a minimum, this includes contributions from the interface, the film surface and the strain energy. In a homo-epitaxial system where there is an absence of strain, the most energetically favoured configuration is a 2-D growth (FvdM) mode. This is because any undulations necessarily increase the area of the surface and consequently increase the surface free energy. However, in a strained system the undulations can lead to an overall decrease in the strain energy, and if this is large enough to offset the increased surface energy these undulations will be favoured. The mechanism by which these undulations lower the strain energy of the system is shown schematically in Figure 2.13. The undulations tend to consist of large rounded mounds with sharp cusps in between. The local spacing of the atoms in the mounds is increased near to the surface and allows local elastic relaxation, but this is partially compensated by the increased strain energy at the cusps. However, the *total* strain energy is lowered as a consequence of the undulations. Because the driving force for roughening is thermodynamic in origin, this form of strain relief is only possible if kinetic barriers are overcome. Consequently, surface roughening only occurs at higher growth temperatures, where the surface migration lengths allow adatoms to arrange themselves in the lowest energy configuration, and can be suppressed by keeping the growth temperature low.

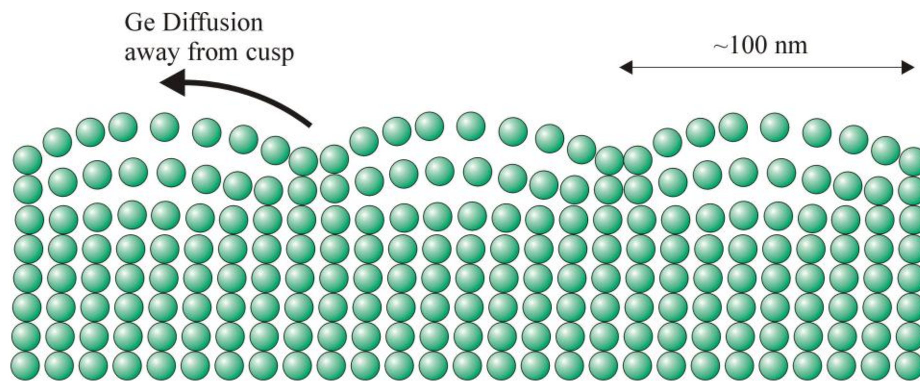


Figure 2.13 Diagram of the mechanism of strain relief by surface roughening.

The thermal activation barrier for roughening has been shown to depend on the strain as ε^{-4},[51] by comparison the barrier to relaxation by dislocation nucleation varies as ε^{-1}. It is apparent that surface roughening will be favoured over dislocation nucleation at high mismatch strain, with dislocations being favoured at lower strain (usually by the MFR multiplication mechanism). Figure 2.14 shows a plot of the relaxation mechanism of SiGe layers grown by UHV/CVD as a function

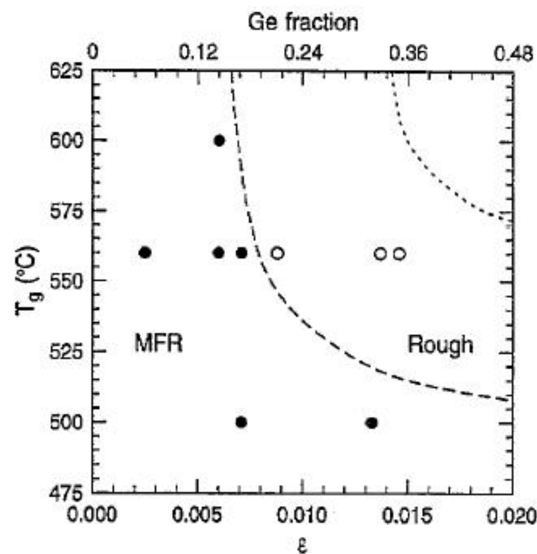


Figure 2.14 Plot of the transition from MFR to surface roughening mechanism for $Si_{1-x}Ge_x$ grown by UHVCVD at different growth temperatures. Open circles indicate relaxation from surface roughening, closed circles indicate MFR mechanism. [Taken from Ref. 52]

of growth temperature and mismatch strain.[52] As can be seen, in order that 3-D growth is suppressed, the growth temperature must be reduced strongly with increasing Ge fraction.

The undulations formed at high temperature tend to have facetted sides orientated at ~11° from the horizontal.[53] This is consistent with a (105) facet and is due the balance between the reduction of elastic energy by roughening, and the increase in energy due to the formation of steps.[54] Subsequent growth over the facetted surface leads to a smoothing effect as there is no longer any strain energy reduction. The varying strain across the undulations can lead to preferential incorporation of migrating Si and Ge adatoms. There is a tendency for the larger Ge adatoms to migrate to the crests of the undulations and the (smaller) Si adatoms to migrate to the cusps, this occurs to reduce the local strain fluctuations at the surface. Another effect of the strain fluctuations is that the increased strain found at the cusps of the surface undulation leads to a local lowering of the dislocation nucleation barriers, these can act as sources for dislocations which form at the surface of the wafer and glide or climb to the hetero-interface.[55]

2.5 Virtual Substrates

In order to fully exploit the strained SiGe system there is a need to be able to control the lattice parameter of the underlying substrate. To grow high germanium layers free of dislocations the buffer layer needs to be strain adjusted so that the strain in the active layer remains lower than that needed to form dislocations. In this way, channels with much higher germanium concentrations can be grown than is possible pseudomorphically on a substrate of just silicon. Also, it is possible to form

a 2-DEG in silicon by tensile straining the channel. This can only be done by growing the silicon on to a substrate whose lattice parameter is larger than that of silicon; hence a relaxed buffer layer is needed. One further use of SiGe virtual substrates is the ability to incorporate III-V semiconductors on to a silicon based platform.[56] As can be seen in Figure 2.15, the lattice parameter of many III-V semiconductors lies within reach of a lattice tuned SiGe virtual substrate. Since these semiconductors have the same crystal structure as SiGe (the diamond structure), if the virtual substrate is lattice matched to the III-V material it would be possible to grow these semiconductors alongside conventional Si-CMOS. This would have huge applications in the optoelectronics industry, with the possibility of incorporating optoelectronic devices with CMOS for complete system on a chip (SOC) solution for signal processing applications.[1]

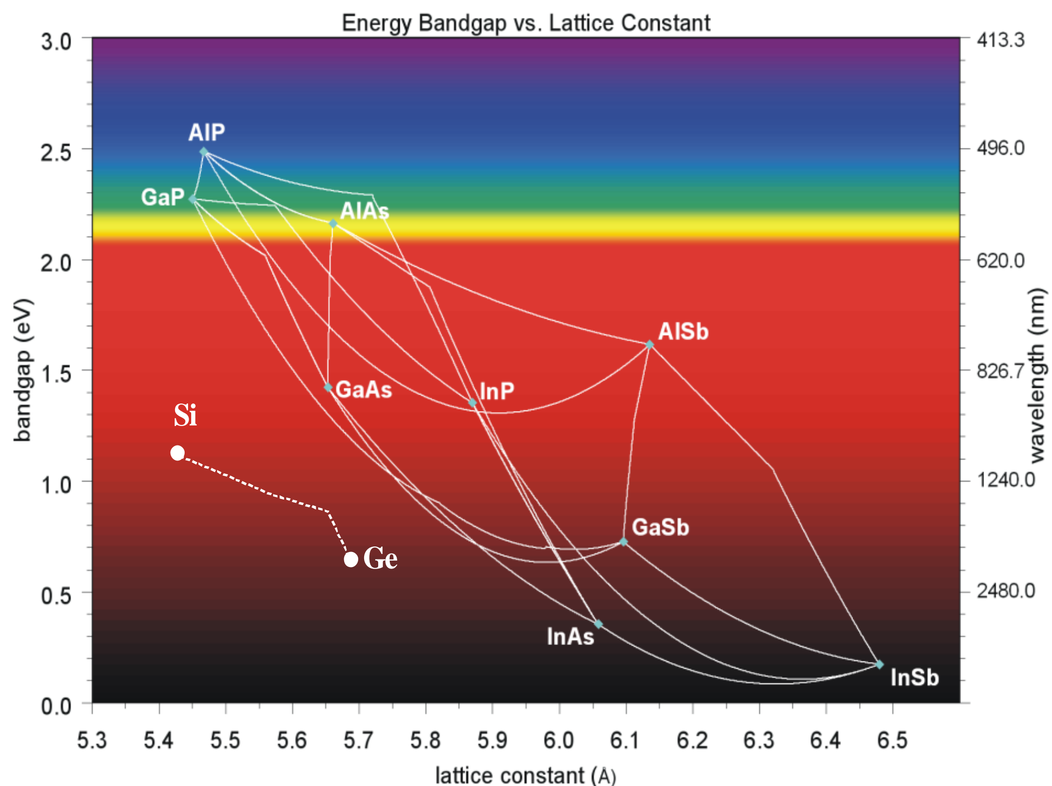


Figure 2.15 Graph showing how the bandgap varies with lattice parameter of the common compound semiconductors.

In order to fulfil this need for a buffer layer whose lattice parameter can be controlled and still be compatible with existing silicon technologies *a virtual substrate* is needed. In its simplest form this is a thick layer of SiGe grown in excess of the critical thickness such that the layer is fully relaxed. The final lattice parameter of the layer can be controlled by carefully choosing the germanium composition. There are, however, many problems associated with the growth of virtual substrates. Since the buffer layer needs to be relaxed, there are inevitably dislocations in the virtual substrates. These dislocations need to be kept away from the channel so that the electrical transport properties are not degraded. The dislocations that provide useful relaxation lie parallel to the growth plane and are known as *misfit dislocations*. A dislocation cannot simply terminate inside the crystal (this is geometrically impossible), instead the dislocation must terminate at a free surface (or at a node with another dislocation). In the case of epitaxial layers, the nearest surface is the growth surface and consequently misfit dislocations tend to terminate with a section which threads to the surface (through the channel) known as a *threading dislocation*. In order to produce useful virtual substrates the density of threading dislocations needs to be kept as low as possible, how low has not been fully determined but for optoelectronics applications dislocation densities lower than 10^4 m^{-2} are needed.[57] Consequently, much research has been carried out in this area.

2.5.1 Strain Relaxation of Virtual Substrates

As stated above, the simplest virtual substrate consists of a SiGe layer of constant composition grown in excess of the critical thickness. As the strain energy increases with the thickness of the layer, it becomes energetically favourable for the strain to be relieved by the formation of misfit dislocations. In a constant

composition layer the dislocations tend to form by nucleating a half loop from the surface, which extends by glide to the Si/SiGe interface.[58] The threading segments then glide away from each other leaving a long misfit dislocation between them. In the ideal case these threading segments would travel to the edge of the wafer, leaving a long misfit dislocation and no threading components. However, in practice since many of these dislocations are needed to relax the layer[59], the dislocations will interact with each other and may become pinned. This has a detrimental effect in two ways. Firstly, the threading dislocations are left penetrating the surface. Secondly, in order to continue relaxation more dislocations need to be nucleated. Both of these consequences cause the threading dislocation density in the substrate to increase. Typical values for threading dislocation densities in constant composition SiGe virtual substrates are 10^9 - 10^{12} cm^{-2}. [60]

If the threading dislocation density is to be reduced then considerable thought must be put into understanding the mechanism by which the layers relax. The relaxation is primarily due to the misfit dislocations that lie on the mismatch interface. Since each misfit dislocation generally has two threading segments (one at each end) then it is clearly advantageous for the misfits to be long. To produce long misfits the threading sections must be allowed to glide as far as possible. This implies that the glide velocity should be high and that pinning events must be reduced. If the misfit dislocations are unable to extend any further, then the continued relaxation requires more dislocations to be introduced, i.e. there needs to be dislocation *nucleation*, which increases the density of threading dislocations.

The glide velocity (v_g) of the dislocations can be approximated by[61]

$$v_g = B\varepsilon_{Eff} \exp\left(-\frac{E_g}{KT}\right) \qquad (2.17)$$

where ε_{Eff} is the effective stress in the layer, E_g is the activation energy for the glide process and has a value $E_g = 2.156 - 0.7x$ eV,[62] and B is a constant. Houghton[63] however, has found that E_g has a value of 2.25 ± 0.05 eV and is independent of the germanium composition, which is close to the value in bulk silicon of 2.20 eV, although he found that the glide velocity depended on ε_{Eff} quadratically, which may account for the discrepancy. The lower value of E_g as x increases causes the dislocation velocity to increase sharply with germanium composition, the glide velocity in pure germanium is ~ 5000 times greater than that in silicon at 550 °C (since the pre-factor is not strongly dependant on x). To ensure long misfit lengths it is advantageous to have a high growth temperature and low grading rate so that the dislocation glide can keep up with the need for relaxation. However, glide velocity is not the only limit to misfit line extension. Since in general, a large number of misfit dislocations are needed to relax the alloy layer, the misfits will inevitably encounter pre-existing misfits in an orthogonal direction. The strain fields of these orthogonal dislocations can cause the gliding dislocation to become pinned and so misfit extension is limited. This is especially evident when alloys with high germanium content ($x > 0.10$) are grown, since higher composition alloys need more dislocations to relax them.

2.5.2 Dislocation Interactions.

Stach *et al*[64] have recently been able to study the interactions of dislocations in thin SiGe films grown on a silicon substrates using *in-situ* TEM. Their UHV-TEM microscope allowed for the observation of dislocations *during* annealing and CVD growth. Their results showed that at low strains *all* interactions result in threading

dislocations being pinned against the orthogonal misfit dislocation. This was attributed to the strain field associated with the misfit dislocation reducing the excess strain which drives the dislocation motion. As the strain was increased, by growing a thicker layer, the effect on the interactions was strongly influenced by the relationship of the burgers vectors of the two dislocations. The strain due to the interactions of two dislocations is very complicated, but is often simplified to the form[65]

$$\sigma_{Int} = k \frac{\boldsymbol{b_1} \cdot \boldsymbol{b_2}}{r} \qquad (2.18)$$

where σ_{Int} is the interaction strain, $\boldsymbol{b_1}$ and $\boldsymbol{b_2}$ are the burgers vectors of the two dislocations and r is the separation of the dislocations cores. It can be seen that when the burgers vectors are perpendicular the strain is not strongly affected, this allows the threading dislocation is to pass freely. Experimental results showed that dislocations with perpendicular burgers vectors did not lead to pinning events in thicker layers. However, dislocations with parallel burgers vectors showed a new form of interaction. The gliding threading segment of the dislocation caused the pre-existing misfit dislocation to split and connect to the gliding dislocation. The result was to form two dislocations, each with a 90° bend. Figure 2.16 shows a schematic diagram of this interaction. This interaction can cause two possible outcomes, and it is proposed that the outcome depends on the direction of the incoming thread with respect to the two glide planes of the dislocations. In Figure 2.16(a) the gliding thread passes through the misfit dislocation and continues unhindered, there is no pinning of the threading dislocation. In Figure 2.16(b) the threading dislocation is unable to continue and becomes pinned. The difference in the two outcomes occurs because of the strains on the dislocations. In (b) the dislocations lie in the

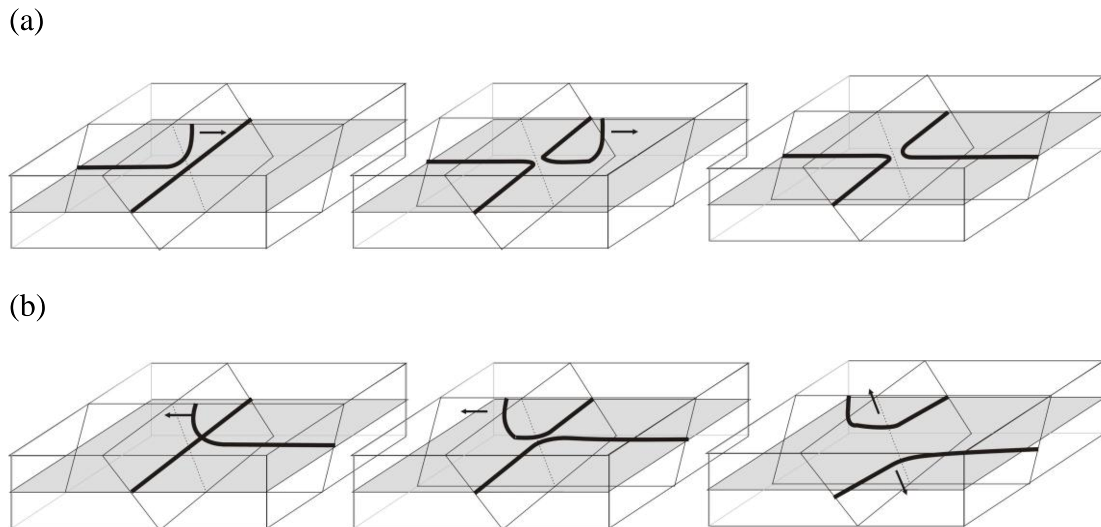


Figure 2.16 Depending on the orientation of the planes in which the corner dislocation lies, the dislocations can either pass unhindered (a), or become pinned by the mutual repulsion of the arms (b).

two glide planes in such a way that the energy of the dislocations can be reduced by gliding away from the point of intersection (this reduces the length of the dislocations and consequently their energy). This pushes the section with the threading arm up towards the surface, effectively blocking the continued glide, the other dislocation is pushed into the substrate. In (a) the dislocations cannot shorten their length by gliding and so no repulsion occurs. The threading dislocation is able to continue gliding and is not pinned.

It was also shown by Stach *et al* that the pinned threading dislocations need much higher stresses to free them, than the stresses needed to prevent pinning in the first place. It can be postulated, therefore, that in order to prevent dislocation pinning a high initial stress is advantageous.

2.5.3 Conventional Approaches to Virtual Substrates

One method to improve the virtual substrate is to grow a relaxed uniform layer of low germanium composition at high temperature, so that the growth is not limited by glide. If the layer has a low enough Ge composition so that the number of orthogonal misfit dislocation interactions is low, then nucleation of new dislocations will be minimised. If now a new layer is grown with a (slightly) higher Ge composition, the existing threading dislocations can continue to glide on this new growth interface. Since this is spatially separated from the original network of misfit dislocations there are minimal interactions. The second layer can be grown in a similar fashion to the first, keeping the pinning events to a minimum. This procedure of growing "steps" can be repeated up to the desired final Ge composition. This technique of *step grading* greatly reduces the density of threading dislocations with typical values of $10^6 – 10^7$ cm^{-2}.[66]

A more common way to reduce the pinning events and hence increase the length of the misfit dislocations is to *linearly grade* the germanium composition with depth. This ensures that the dislocation networks are distributed over the entire width of the graded region. This greatly reduces the number of dislocations on a single atomic plane, and so reduces the chance of interactions. It is usual to reduce the temperature with increasing Ge fraction in order to suppress surface roughening due to the increased adatom mobility of Ge.[67] Threading dislocation densities using linear grading generally have threading dislocation densities of ~ 10^6 cm^{-2},[50] although densities down to 10^3 cm^{-2} have been reported in very thick (10 µm) $Si_{0.68}Ge_{0.32}$ linearly graded layers.[68]

2.5.4 Multiplication Mechanisms

The mechanism proposed by LeGoues *et al*[50] for the relaxation in linearly graded buffer layers involves the multiplication of misfit dislocations from orthogonal interactions, and the subsequent annihilation of their threading segments. Referring to Figure 2.17 as a gliding threading segment approaches a pre-existing misfit dislocation with the same burgers vector, the dislocations can rearrange themselves to form two corner dislocations as discussed above. In this way the blocking of dislocation glide can be avoided. These corner dislocations will have the tendency to repel each other as they have the same burgers vector (a). However, since the two sides of the corner dislocation lie on different glide planes, the dislocation is locked to the line which is common to both. Since the dislocations repel each other, but are constrained to this line, one is pushed down into the substrate. The dislocation can then form a half loop (b,c) which expands (d,e) until it reaches the surface creating two new threading arms and re-forming the original corner dislocation (f). The two new threading dislocations can then glide away from the interaction forming new misfit dislocations on the same glide plane as the original dislocations (g,h). This process can be repeated many times and hence is known as a multiplication mechanism and leaves the tell-tale signature shown in (i). This mechanism is similar to the much studied Frank-Read mechanism found in other systems and is known as the modified Frank-Read (MFR) multiplication mechanism. An advantage of this relaxation mechanism is that the pre-existing misfit dislocation can cause many multiplication sites. Threading dislocations formed from these sites all have the same burgers vector and lie on the same glide planes. Hence, these dislocations will tend to annihilate each other, reducing the total threading dislocation density.

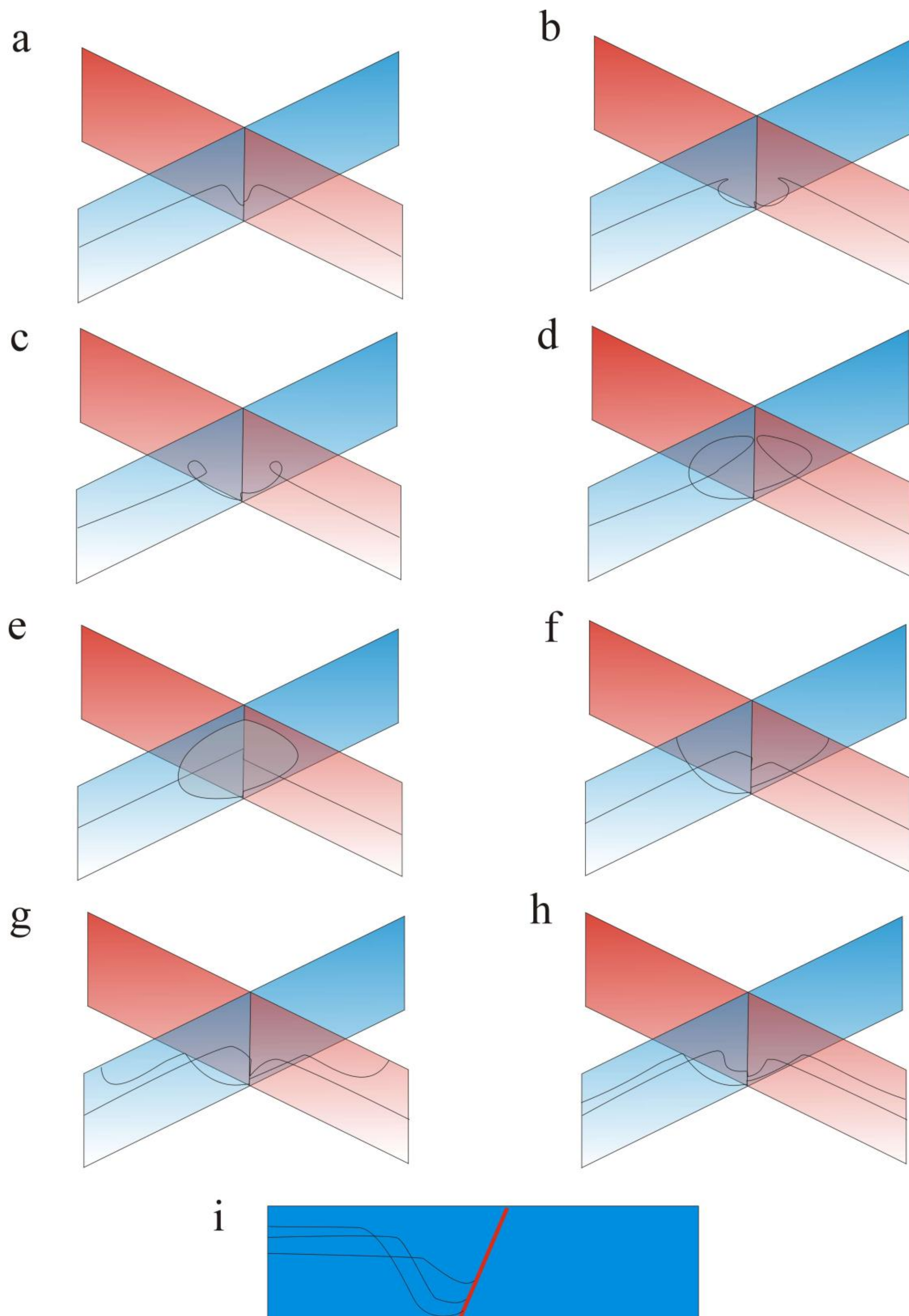


Figure 2.17 Schematic diagram of the modified Frank-Read dislocation multiplication mechanism. A corner dislocation lying on two different glide planes (blue and red in the diagram) is pushed down into the substrate. A half loop forms and expands to the surface creating two more threading dislocation which can glide away from the corner. The distinctive dislocation structure of this mechanism is shown in (i), viewed along the [110] direction.

Layers relaxed by the MFR mechanism leads to large pile-ups of misfit dislocations on the same glide plane. Lutz *et al* [69] have showed that a misfit dislocation causes a surface step of approximately 2.5 Å where the glide plane intersects the surface. Also, if there is a pile-up of dislocations on the same glide plane the total surface step is simply the sum of the steps due to each individual dislocation. Consequently, when layers relax by the MFR mechanism, the large pile-ups occurring on the same glide planes cause a surface step which can become quite large. The network of surface steps caused by the buried misfit dislocation is known as *crosshatch*. This crosshatch is detrimental to the quality of the buffer layer. It can cause roughening in electrically active channels which are grown on the buffer layers and non-uniform oxidation which causes variations in the properties of MOS devices across a wafer.[70] Also, crosshatch causes difficulty in the lithography needed for device processing and as device scaling continues, this problem will become worse.

Fitzgerald *et al* [46] have shown that the crosshatch pattern on the surface of the virtual substrate can lead to pinning of threading dislocations. In their model proposed for strain relaxation, misfit dislocations form by glide of threading dislocations when there is enough residual strain to overcome kinetic and thermodynamic barriers. This can only occur when there is a critical channel width (h^*) above the dislocations which is still strained, the relaxation then only occurs below this channel width, with threading dislocations continually gliding in the strained channel. In order to maintain this dynamic equilibrium of strain relief, it is shown that a density of *mobile* threading dislocations of ~ 10^5 - 10^6 cm^{-2} are needed under most growth conditions. As the virtual substrate grows, crosshatch starts to form on the surface by the mechanism described above or from the inhomogeneous strain fields associated with the underlying dislocation network. As the mobile

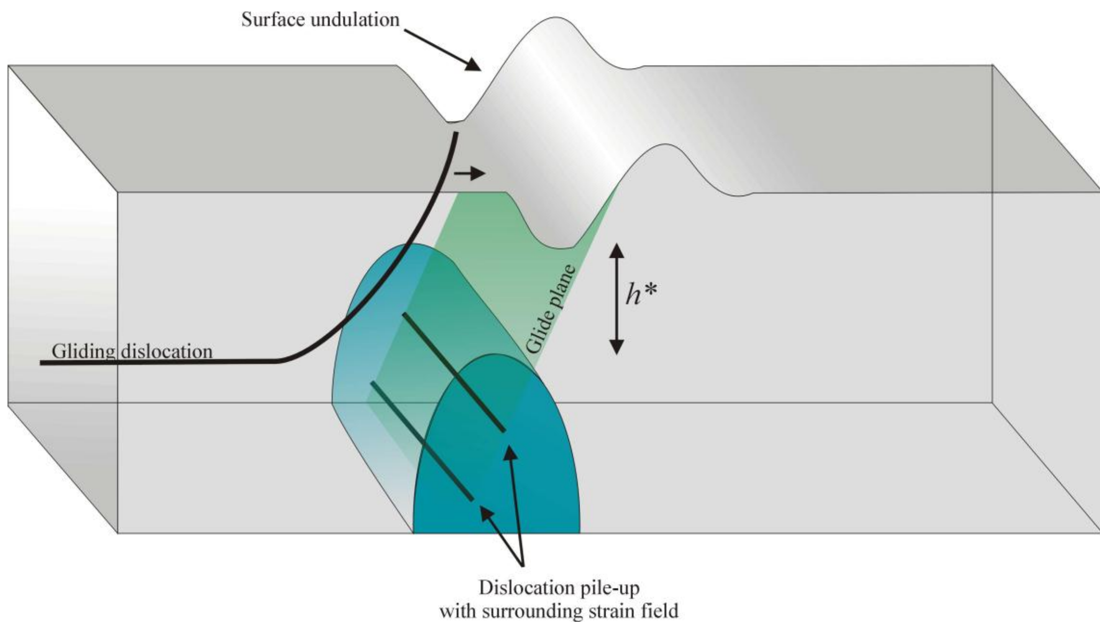


Figure 2.18 The height h^* that the gliding dislocation may pass through is reduced by the presence of pre-existing dislocations and their associated strain fields (shaded blue) and the undulation at the surface.

threading dislocations cross the deep troughs associated with this crosshatch they can become trapped, this is due to a narrowing of the channel width below these troughs as shown schematically in Figure 2.18. Once trapped, these threading dislocations can no longer play a part in the relaxation process. Since the density of *mobile* threading dislocations is reduced by these pinning events, the strain will increase until sufficient energy is available to nucleate more dislocations and so restore their equilibrium concentration. The strain field at the surface of the wafer associated with the threading dislocations causes the adatom incorporation to be reduced locally, further contributing to the amplitude of the undulations. This vicious circle quickly increases both the RMS roughness of the surface of the virtual substrate and the threading dislocation density. In order to reduce this effect, the grading rate should be kept as low as possible. In doing this, the critical channel width is kept reasonably thick and hence the surface is far from the underlying misfit dislocation networks.

This will reduce the effect of the strain fields associated with these networks to a minimum at the surface of the wafer and so reduce the depths of the troughs. Also, the thicker layers that are inevitable at low grading rates will give a smoothing effect on the surface, again reducing the trough depth. There are, however, problems associated with growth of thick epitaxial layers needed for low grading rates. Firstly, the deposition time needed for such thick layers can be unsuitable for integration into industrial processes, especially for MBE which would have a crippling long growth times for thick buffer layers. Secondly, the step height associated with the thick layers will make integration with Si-CMOS difficult due to lithographical and metallization issues. Also, the lower thermal conductivity of SiGe alloys over silicon would cause problems for thick layers, since at large integration levels heating effects due to the devices fabricated on top of the buffer layers can cause severe problems. A compromise between high grading rates needed for reduced threading dislocation densities and the need to keep the buffer layers reasonably thin needs to be found. Fitzgerald *et al* [46] have shown that the optimal grading rate, taking into account the above arguments, is 10 % Ge/μm.

2.6 Recent Developments in the Growth of Virtual Substrates

Much research is currently being carried out on the growth of SiGe virtual substrates. The quality of these substrates is currently not sufficiently high for incorporation into Si technologies, there is a need to reduce both the threading dislocation density and the RMS roughness of the surface in order that the full potential of SiGe technology can be achieved. In the following section a quick review of some of the most recent developments will be given.

2.6.1 CMP

Currie *et al*[71] has overcome the problem of crosshatch pinning threading dislocations by the introduction of chemical mechanical polishing stages (CMP) during production. The virtual substrates are linearly graded with a grading rate of 10 %/µm to keep the threading dislocation density low as described above. During the growth of the virtual substrate the growth is stopped and the wafer is removed from the growth chamber in order that a CMP step can be employed to polish the wafers surface removing the crosshatch. The wafer is then reloaded into the chamber and the growth of the virtual substrate continues. Since the crosshatch which pinned the threading dislocations has now been removed, the dislocations are free to glide again. In this way the pre-existing dislocations can be useful for strain relaxation again and there is no need for further dislocations to be nucleated. Further, now that the dislocations are free to move, there is an enhanced probability of dislocation annihilation which reduces the threading dislocation density. The threading dislocation density of a virtual substrate grown to pure germanium ($x = 1$) using this CMP stage was found to be 2.1×10^{6} cm^{-2} compared to > 10^{7} cm^{-2} for a similar virtual substrate without the CMP stage. The RMS roughness was similarly reduced from 47 nm without the CMP stage to 24.2 nm with CMP.

2.6.2 In-situ Annealing

Kissinger *et al*[72] have shown that the threading dislocation density can be drastically reduced by the introduction of a high temperature *in-situ* anneal during growth of the virtual substrate. In their experiments virtual substrates where grown using APCVD up to $x = 0.20$ with a step-graded profile with a total thickness of

500 nm. After the growth of each step the wafers were subjected to a high temperature anneal of 1050 °C for 1 hour. Without the annealing stages the threading dislocation density was found to be ~10^6 cm^{-2}, but the introduction of the annealing steps reduced this to as low as 3×10^2 cm^{-2}. This was attributed to the high glide velocity of the dislocations in this temperature regime which is further enhanced by interdiffusion of the atoms within the crystal. The aim of this sequence of growth and anneals was to allow misfit dislocation formation from the edge of the wafer instead of the more usual multiplication mechanisms. In this way, if the dislocations are able to traverse the entire width of the wafer, the density of threading dislocation can be kept to a minimum. The anneals had to be long since as the strain was relieved the glide velocity decreased, taking nearly an hour for full relaxation of the layers. The surface still had the typical crosshatch pattern but the amplitude of the undulations is not discussed. In order that contamination of the wafer during the anneal steps is kept to a minimum, the anneals were carried out under an atmosphere of hydrogen.

2.6.3 Point Defect Injection

Kasper *et al*[73] suggested that the intentional injection of point defects into a SiGe virtual substrate might lead to a reduction in threading dislocations. Since point defects allow the process of dislocation climb (as described in section 2.3.2.2) then by allowing threading dislocations to cross glide planes the probability of dislocation annihilation is enhanced Figure 2.19. Also, point defects can condense to form large prismatic loops within the layer. This would allow dislocation nucleation without the need for threading dislocations. Kasper calculated that point defect densities of ~ 10^{12} cm^{-2} would be needed. Two methods were proposed for the creation of these point defects, Si^+ bombardment and low temperature epitaxy (LTE). The

bombardment of Si^{+} ions can knock Si atoms out of the lattice sites from within the substrate creating vacancies and interstitials. Whilst LTE can cause vacancies due to the low adatom mobilities preventing complete coverage of the wafer surface during growth, consequently small voids form in the crystal.

A number of authors have used low temperature epitaxy (LTE) during the initial stages of growth of virtual substrates. Bauer *et al*[74] grew a fully relaxed $x = 0.28$ virtual substrate of only 70 nm thickness. An LTE SiGe layer was grown at a growth temperature of 200 °C and then the temperature was raised to 550 °C before a second SiGe layer was grown. It was shown from XTEM that the threading dislocation density was reduced compared to a similar thickness of SiGe grown without the LTE. Gaiduk *et al*[75] extended this approach to thin virtual substrates by growing a series of step graded layers by MBE using a sequence of LTE (to produce point defects) followed by a higher growth temperature layer (to allow dislocations to form) and finally a high temperature anneal to ensure relaxation. Using this technique they were able to reduce the threading dislocation density of $x = 0.30$ SiGe virtual substrates to 10^4 cm^{-2}, compared to 10^9 cm^{-2} for similar layers grown without the LTE and annealing steps The RMS roughness of the surface was also found to be significantly reduced. A second virtual substrate with $x = 0.50$ was grown using this technique with a threading dislocation density of 4×10^4 cm^{-2}.

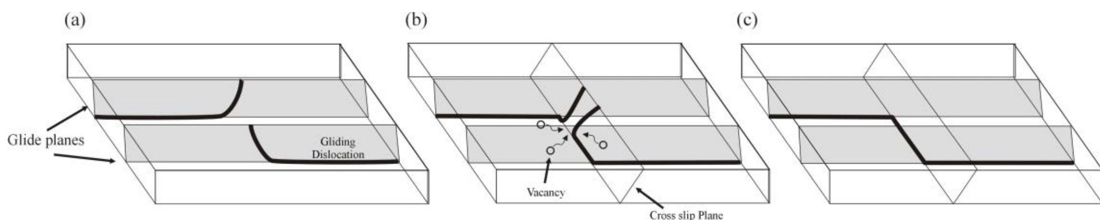


Figure 2.19 Schematic diagram of how dislocations on different glide planes (a), can cross slip by the diffusion of vacancies (b), leading to threading dislocation annihilation (c).

2.6.4 Surfactant Mediated Growth

Another technique for improving the quality of SiGe virtual substrates is to use surfactants. Antimony (Sb) has been used during growth of virtual substrates in order to change the surface properties of the wafer. Liu *et al*[76] have grown an $x = 0.50$ SiGe virtual substrates using MBE with a thickness of 2 μm. A monolayer of Sb was first deposited onto the surface of the wafer before a linearly graded SiGe layer was grown at 510 °C. For comparison a similar wafer was grown without the Sb surfactant for comparison. It was found that the Sb reduced the threading dislocation density from 6×10^{7} cm^{-2} for the substrate grown without the Sb to 1.5×10^{4} cm^{-2} for the substrate with the Sb. Also, AFM showed that the Sb improved the RMS from 16 nm to 2 nm. These improvements in both threading dislocation densities and RMS roughness could be due to the Sb reducing the surface free energy of the wafer as it was growing. This would increase the glide velocity of the dislocation, thus allowing dislocations annihilation, and should reduce the effect of pinning of the dislocations by the surface roughness (as described in section 2.5.4).

2.6.5 Limited Area Growth

As early as 1989 Fitgerald *et al*[77] showed that by the reduction of the growth area in epitaxial systems that the dislocation density could be reduced. By defining mesa pillars on GaAs wafers followed by MBE epitaxy of $In_{0.05}Ga_{0.95}As$ (350 nm) a linear relationship between misfit dislocation density with lateral size of the growth zone was found. This linear relationship was shown to be due to the elimination of multiplication mechanisms leaving only heterogeneous nucleation as a source for dislocation formation. These nucleation sites are proportion to the area of the growth

zone, so a reduction of this area leads to a reduction in dislocation density. For a thickness of 700 nm a superlinear relationship was shown indicating the onset of multiplication mechanisms. Noble *et al*[78] showed similar results for the SiGe system. In their work the growth zones were defined by selectively growing $Si_{0.8}Ge_{0.2}$ using limited reaction processing (LRP)-CVD in oxide windows etched on a silicon substrate. If the only dislocation sources are fixed (proportional to area) then for a growth zone of lateral dimensions L, the total number of active sources, N_s, is given by

$$N_S = \rho L^2 \tag{2.19}$$

If the dislocations undergo no interactions, they will traverse the entire growth zone giving a total number of dislocations N_D of

$$N_D = \alpha N_s L = \alpha \rho L^3 \tag{2.20}$$

where α is a geometric factor ~1-2. Since the area of the growth zone is given by L^2 the dislocation density ρ_D is given by

$$\rho_D = \frac{N_D}{A} = \alpha \rho L \propto L \tag{2.21}$$

Luryi *et al*[79] gave a theoretical model for the elastic relaxation of thin epitaxial layers grown with limited lateral dimensions. They showed that the strain could be relieved by an elastic distortion of the lattice which extended the critical thickness of the system. This critical thickness is a strong function of the lateral dimensions of the growth zone and for some critical length, L_c, the strain of an arbitrary layer thickness can be completely relieved without the formation of dislocations.

Hammond *et al*[80] used this technique of limited-area growth to produce a linearly graded $Si_{0.77}Ge_{0.23}$ virtual substrate of only 500 nm thickness. These layers were grown using SS-MBE on mesa pillars etched into a silicon substrate. It was found that by the suppression of multiplication mechanisms at lateral dimensions ≤ 10 µm, the surface crosshatch could be completely eliminated, although a maximum relaxation of only 80% was achieved. Wohl *et al*[81] has continued this work with a view to fabricating hetero-CMOS transistors. Their approach involves growing 750 nm layers up to $x = 0.30$ in oxide/nitride windows using MBE. The poor relaxation in growth zone ≤ 10 µm was overcome by the introduction of an LTE-Si stage prior to the growth of the virtual substrate. A 50 nm layer of silicon was grown at 330 °C in order that point defects are introduced to aid dislocation formation at the Si/SiGe interface. A much reduced threading dislocation density was reported from TEM studies. However, the suitability of limited-area growth to CMOS applications is questionable due to the inherent non planar surfaces produced. If this approach is to be successfully incorporated into existing technologies, a planarisation stage is likely to be necessary.

3 Experimental Techniques

The following chapter describes the main experimental techniques used in this study and the theories behind them. These techniques have been performed by the author or by specialists in those particular techniques at the author's request. A large part of this chapter concerns the use of microscopes including differential interference contrast (DIC) microscopy, transmission electron microscopy (TEM), scanning electron microscopy (SEM) and atomic force microscopy (AFM). Other techniques include X-ray diffractometry (XRD) secondary ion mass spectroscopy (SIMS) and electrical measurements.

3.1 Microscopy

3.1.1 Introduction

In order to structurally characterise semiconductor materials it is necessary to actually *look* at the material. To do this, some kind of microscope is needed. There are several types of microscope available to the microscopist, of these the most important are; the optical microscope, the Transmission Electron Microscope (TEM), the Scanning Electron Microscope (SEM) and the Atomic Force Microscope (AFM). An optical microscope is limited in its resolution by the diffraction limit imposed by the wavelength of light. Even the best optical microscopes available cannot overcome this fundamental limit. Under ordinary lighting a minimum feature size of 400 nm is the best one can expect to resolve, this leads to a maximum useful magnification of no more than 1000 times. As many of the features of interest in semiconductor materials are on a much smaller scale than this, channel widths of

order 5 – 20 nm and surface undulations as low as 1 - 2 nm for example, optical microscopy is of little use. A second draw back to the optical microscope is the very small depth of field, this is the depth of an object that is in focus at any one time. Because of the large convergence angle of an optical microscope the depth of field is limited to values around 1 μm. Any increase in depth is at a cost of reduced resolution.

The electron microscope has many advantages over the optical microscope including a very large increase in both resolution and depth of field. The source of illumination is not light with its comparatively long wavelength, but a beam of fast moving electrons whose wavelength, λ, can be controlled through the de Broglie relation,[82]

$$\lambda = \frac{h}{p} \tag{3.1}$$

by its momentum p. Hence, by increasing the electrons momentum (by accelerating them through a large electric potential difference) we can reduce the associated wavelength to many orders of magnitude less than that of visible light. For this reason electron microscopes typically run at 100 – 400 KV with very high magnification microscopes operating in excess of 1 MV allowing atomic resolution (~ 1 Å).

Other advantages of the electron microscope include the ability in a Transmission Electron Microscope (TEM) to choose only electrons that have met a certain diffraction condition to form an image. This enables the operator to look at specific crystallographic planes, giving the opportunity to characterise the type of features that are seen (for example, the burgers vectors of dislocations can be determined) and also to maximise the contrast from strained regions within the layers.

One drawback of the TEM is the necessity to painstakingly prepare the samples to around 100 nm in thickness so that enough electrons can pass through the sample to give an intensity suitable to form an image. This is because electrons interact very strongly with matter and hence will only pass through very thin specimens. In order to produce such thin samples the preparation time is typically 2 days and can have a high failure rate, particularly for the inexperienced.

3.1.2 Image Formation in the Microscope

The optical microscope and the TEM form an image in a very similar manner. Consequently, the main components in the TEM have an analogous component in the optical microscope. In order to understand how an image is formed in the TEM it is useful to revise how an image is formed in the optical microscope.

Figure 3.1(a) shows the main components needed to form an image in an optical microscope. A source of light (usually a lamp) is placed at the focus of a condenser lens in order to collimate the beam of light. The condenser aperture placed near to the condenser lens limits the width of the beam to ensure that light which forms the image will have a well defined path near to the optical axis. The object is illuminated (in reflection or transmission) by the beam and an intermediate image of the object is formed by the objective lens. The objective aperture limits the angle of light that is allowed to form the image, this limits the final resolution. Finally, the eyepiece magnifies the intermediate image and produces a virtual image at infinity. With the aid of the lens in the eye, this is projected onto the back of the retina as a real, magnified image.

3.1.3 Differential Interference Contrast (DIC)

In order to study the surface topogoraphy of a sample in an optical microscope it is very useful to use differential interference contrast (DIC). Figure 3.1(b) shows the principle features of this technique (also called Nomarski

microscopy). The collimated light from the source is plane polarised by a polariser and is reflected down towards the sample with a half-silvered mirror. Before the light reaches the objective lens it passes through a birefringent wedge (Wollaston Prism). This has the effect of splitting the beam into two rays with orthogonal polarisations. The objective lens focuses the two beams onto the sample which are then reflected back into the objective lens. As the beams pass through the wedge the two beams are brought together, although interference cannot occur due to their polarisations being perpendicular. A second mirror reflects the light towards the eyepiece (if a camera is to be used this would be movable) where a second polariser, crossed with the first, acts as an analyser. This analyser recombines the two beams with the same polarisation, thus allowing interference to occur. Any phase difference between the two beams of light through the system will cause the image to have interference fringes. This may be due to the wedge, but also due to the slope of the sample. Since the intensity of light reflecting from a surface depends on the polarisation vector relative to the surface normal, any slope of the sample will introduce a phase shift of the light. This will show up in the final image as a change of colour. Consequently,

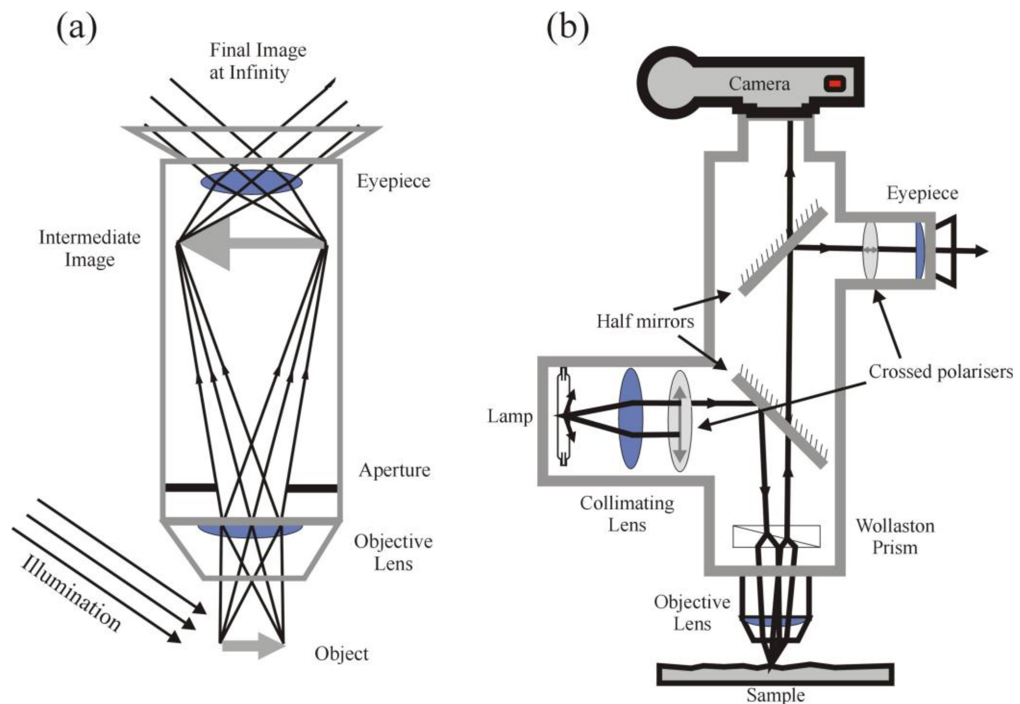


Figure 3.1 (a) Ray diagram of the optical microscope showing main lenses and aperture. (b) Diagram of a Nomarski (DIC) optical microscope showing the beam splitting due to the Wollaston prism.

in the Nomarski microscope a slope on the surface leads to a change in contrast in the image.

The Nomarski microscope is very sensitive to changes in height of the sample and so can be used to observe quite small surface features of only a few nm's. However, because of the complicated nature of the image formation, this technique is not quantitative. It is useful as an initial examination of the quality of a wafer since it requires no preparation and is non-destructive, and also for studies where height calibration is not needed (i.e. counting defect etch pits). For further discussion on optical techniques the reader is directed to the numerous texts on the subject.[83]

3.1.4 Image Formation in the TEM

The following description concerns the JEOL JEM-2000FX TEM used in this study but the main features are common to most TEM systems currently available. A diagram of the main features of the TEM is shown in Figure 3.2. The source of electrons is an electron gun which consists of a fine tungsten wire filament connected to the cathode of a high voltage supply. The electrons are accelerated towards the grounded anode by high voltage supply which can be varied between 80 – 200 kV, although the normal mode of operation is at the highest setting of 200 kV. The filament is heated to produce thermionic emission by inducing a filament current which is set so that the electron beam is *just* saturated to produce a uniform beam of electrons. A Wehnelt cap shields the electron gun and produces a focussing effect of the electron beam, termed *crossover*. Since all electrons pass through this crossover it can be considered as the source of electrons.

The electron beam passes through the condenser aperture to limit its width and two condenser lenses control the size of the beam at the specimen. The first lens produces a narrow spot, hence is called the *spot size* control. The second lens diverges the beam and so controls the intensity of electrons at the sample, hence this

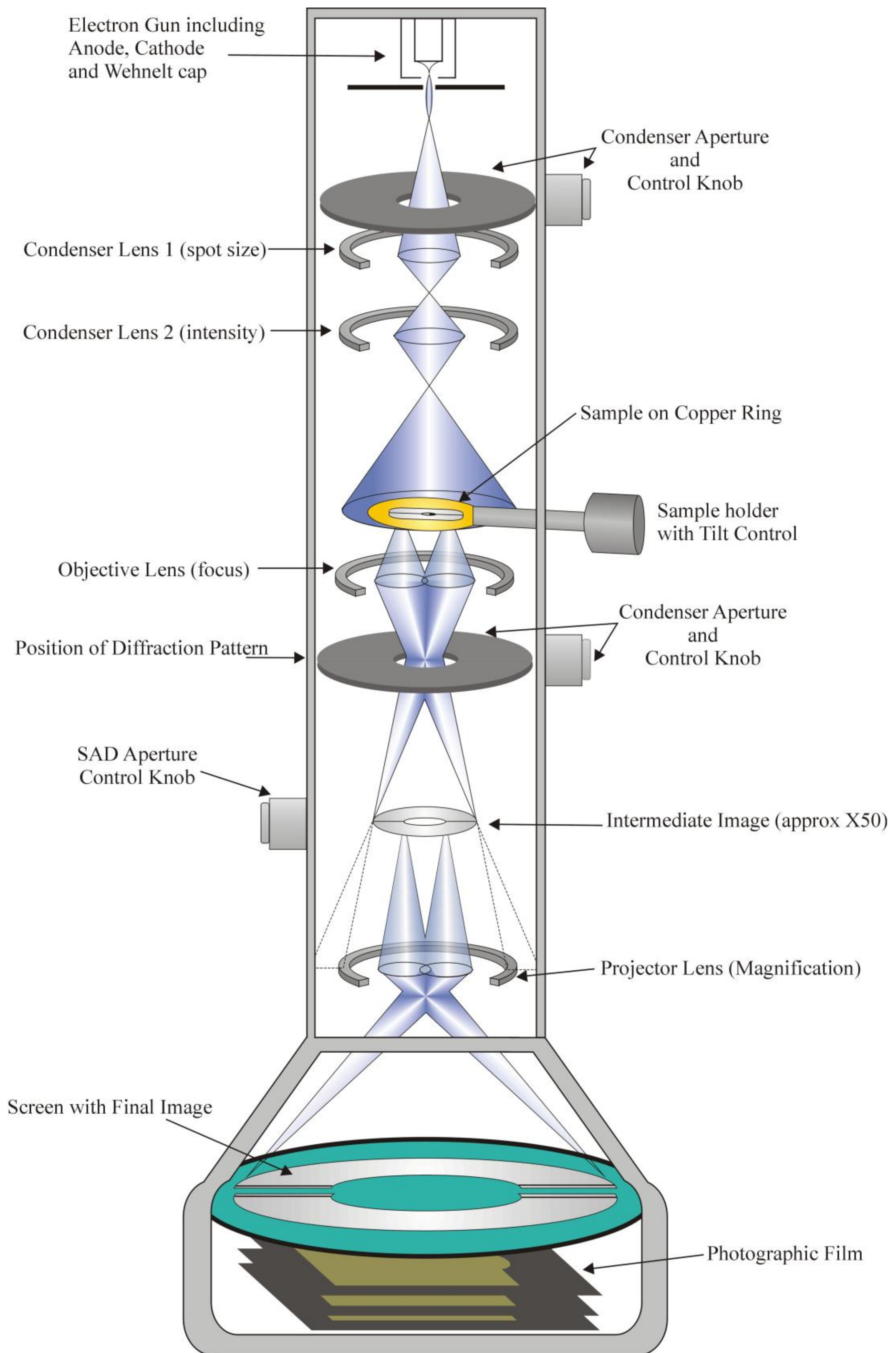


Figure 3.2 Schematic diagram of the TEM showing the main components and the electron beam path.

is called the *intensity* control. This second lens is usually set to be over-focussed so as to produce the best image.

The sample is held in place by the specimen holder which has the ability to control the tilt of the sample in the two orthogonal directions, this is important to control the crystallographic orientation used to produce the image. Immediately after the sample, is the objective lens which produces a real image of the sample at approximately 50 times magnification. The objective lens can be controlled in order to focus the final image. Between the objective lens and the intermediate image is the position of the diffraction pattern image, this occurs at the back focal plane of the objective lens. It is at this position that the objective aperture is used to select specific diffraction conditions, this is very important to produce contrast in the TEM. At the position of the intermediate image is the Selective Area Diffraction (SAD) aperture. When the TEM is in imaging mode this aperture is moved out of the beam, but in diffraction mode the aperture is used to select the area of the sample allowed to contribute to the diffraction pattern.

The magnification occurs mainly in the final section of the TEM column using the projector lens to produce a final real image on a phosphor screen for viewing. There are in fact several lenses in the projector system, the intermediate lenses are active or inactive depending on the actual magnification required. In order to produce a lasting image a camera system beneath the phosphor screen is used. When a photograph is taken a shutter prevents the electron beam from forming the final image while the screen is lifted up, the shutter is then moved out of the way for the duration of the exposure time and the screen then returns to its original position.

Electrons interact strongly with matter and consequently the entire TEM column is kept under high vacuum, typically $< 10^{-7}$ mbar. A combination of turbo-molecular, diffusion and ion pumps keep the column under vacuum, with isolation valves between the different chambers within the system, such that samples can be loaded without venting the entire column.

3.1.5 The SEM

Another type of electron microscope is the Scanning Electron Microscope (SEM). In contrast to the TEM samples to be studied in the SEM do not need to be thinned. There is no sample preparation needed for electrically conducting materials at all (non-conducting materials need to be coated in a thin film of gold). The main use of an SEM is to study the surface morphology of samples.

An SEM has a similar electron gun system to a TEM. The beam is focused to a small spot on the sample and can interact in many ways with the material. In the most common mode of operation the *secondary electrons* that are emitted after interaction with the electron beam are used to form an image. As the beam hits the sample the electrons penetrate the sample and collide with the atoms within the material. Secondary electrons are emitted from the material within a small interaction region near the surface of the sample. A detector placed near to the sample collects these secondary electrons and produces an electrical signal proportional to the number of electrons detected.

The electron beam is scanned in a manner identical to a television raster. The signal is fed to a CRT which has its raster in time with the beams scan. The result is an image based on the number of electrons emitted at different points on the sample, by reducing the scanned area of the electron beam the image is magnified with a typical magnification range of 10× to 100000×. In the SEM a significant increase of resolution, magnification and depth of field can be achieved over that of an optical microscope.

Other modes of imaging are possible in the SEM using other effects of the electron beam on the sample. The main modes are Electron Beam Induced Current (EBIC) which uses the current induced within samples consisting of a pn junction,

back scattered electrons which are sensitive to the atomic number of the sample, cathodoluminescence which uses light which is emitted and X-ray analysis which can be used for elemental determination.

3.1.6 AFM

Atomic Force Microscopy (AFM) can form a very high magnification height profile of the surface of a sample, using this technique a minimum feature size as low as 0.1 Å can be imaged. In an AFM a small probe of approximately 20 nm is dragged over the surface of the sample on the end of a cantilever. Sensitive piezo-electric ceramics control the rastering of the probe over the sample surface, the maximum area that can be rastered is 100 μm × 100 μm with a typical raster frequency of ~ 2 Hz. There are two main imaging modes of the AFM, contact mode and tapping mode. In this study a Digital Instruments Nanoscope III was used in contact mode, a diagram of which is shown in Figure 3.3.

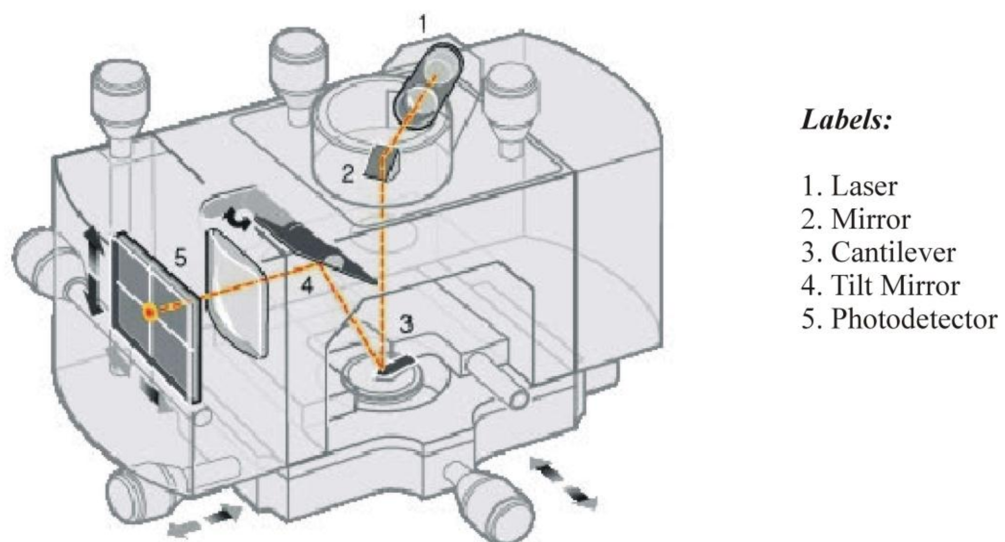


Figure 3.3 Schematic diagram of the working head of the Digital Instruments Nanoscope III AFM instrument. [Taken from the Digital Instruments user manual]

3.1.6.1 Contact Mode

In contact mode the tip of the probe is allowed to rest on the surface of the sample. There is a mutual repulsion between the probe tip and the atoms of the surface due to the overlap of their atomic orbitals. This repulsive force is kept constant throughout the data acquisition by a sensitive measurement and feedback circuit. When the probe encounters a raised feature on the surface of the sample, the increased repulsion on the probe causes a small deflection of the cantilever. This is detected by the deflection of a laser spot focused on the tip of the cantilever and reflected onto a sensor. The sensitive four part light sensor detects the tiny motion of the laser spot and sends a signal to the computer controlled feedback circuit. This feedback circuit compensates for the increased repulsion on the probe by changing the electric potential across the piezo-electric height controller, until the repulsive force is maintained to its preset level. Figure 3.4 shows a schematic diagram of the AFM apparatus used in this study.

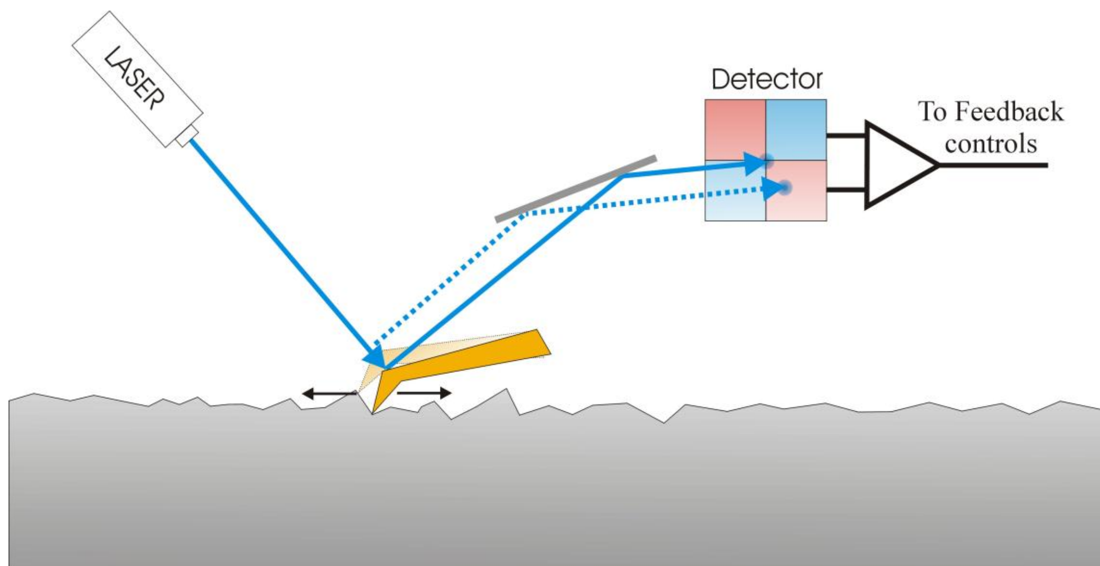


Figure 3.4 Schematic diagram of the operation of an AFM. The sample is scanned underneath the cantilever and tip. The rise and fall of the cantilever is detected by the shift of a laser spot reflected off the AFM tip.

The imaging data is derived from the voltage used by the feedback circuit to maintain a constant force. Sophisticated software is able to convert this data into a 3-D topographical map and allow data analysis, for example measurement of the RMS roughness. This mode of operation of the AFM is the easiest to set up having the least free parameters to control. For some applications the force of the cantilever on the surface in this mode can damage the sample making this mode unsuitable for these purposes. Fortunately, the Si/SiGe materials used in this study are sufficiently hard that they can be safely imaged in contact mode, and hence only this mode was used.

3.1.6.2 Tapping Mode

In this mode the probe is not in contact with the surface of the sample and so causes no damage to soft samples. Instead the cantilever is oscillated at its resonant frequency so that the probe is close enough to the sample to feel the attractive/repulsive forces but without actually touching it. Any change in the force that the probe feels due to changes in the height of the sample will result in a change of the oscillations of the cantilever. The image data can be acquired by using the change in amplitude of the oscillation (Amplitude mode) of the phase shift of the oscillation (Phase mode).

Great care must be taken in analysing data taken from an AFM, there are many reasons why the data may not be a fair representation of the actual topography of the surface of the sample. The shape of the probe has a significant impact on the resolution of the final image. A blunt rounded tip will be unable to follow a sharp feature on the surface and will give a much larger width of this feature. In this case it is the shape of the probe which is actually being imaged. This is most evident in contact mode where the constant abrasion of the probe by the hard surface can blunt

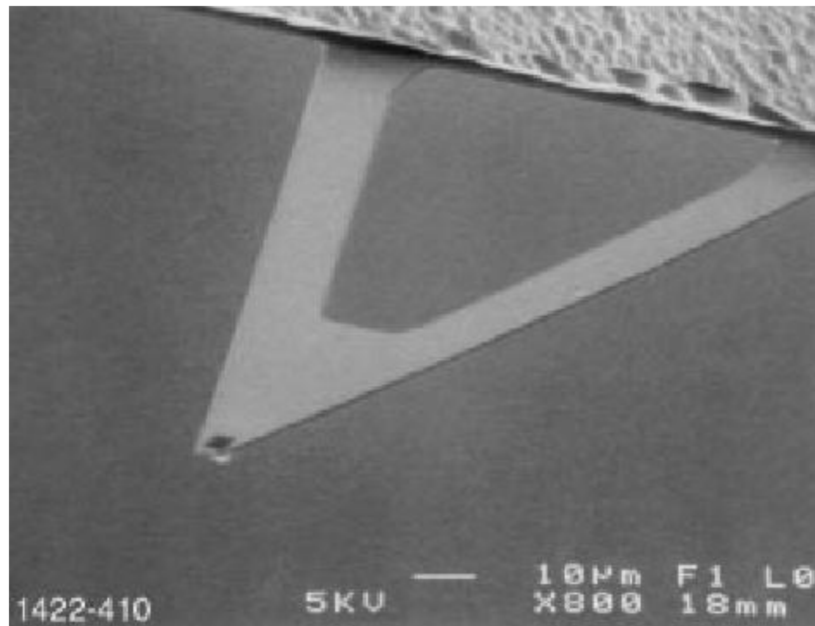


Figure 3.5 SEM micrograph of a silicon nitride tip used in the AFM.

the sharp tip. The fineness of the AFM tip is shown in Figure 3.5 which is an SEM micrograph of a typical AFM tip. The cantilever can clearly be seen with the very fine tip on the underside at the end of the cantilever. Tapping mode is less prone to this problem of blunting due to its non-contact nature. Another problem that may occur is that any contamination on the surface can lead to erroneous readings since it can cause "sticking" of the tip to the sample, it is very important therefore to thoroughly clean and dry the specimen before it is examined in the AFM. The operator of the AFM must carefully set up the instrument so that the probe follows the surface morphology as closely as possible. There are three main feedback controls, differential gain, proportional gain and look-ahead gain as well as controls for the tip velocity over the surface. Incorrect set up of the AFM can lead to poor images and low resolution.

3.1.7 TEM: Sample Preparation

As discussed earlier, in order to examine samples in a TEM the thickness of the sample must be reduced to around 100 nm. This is to ensure that enough electrons can pass through the sample to form an image. There are many ways in which a TEM specimen may be prepared, the exact technique depends on the individual. Consistent preparation of TEM samples is more of an art than a science and is a laborious task. The method used in this study is outlined below and in Figure 3.6.

Two samples of approximately 2.5 mm × 10 mm are cleaved using a diamond scribe along the <110> directions. These are then glued together, uppermost sides in, using Gatan M-Bond-610 adhesive which is applied using an artist paint brush. Tweezers are used to rub the two samples together to ensure even distribution of the glue and to expel any air bubbles. Two support blocks are then glued to each side of the sandwich in the same manner. The whole block is then clamped and placed in an oven which is ramped in temperature to approximately 180 °C over 1 hour. The sample is held at this temperature for 2 – 3 hours and allowed to cool slowly, this ensures that the glue is cured without undue stress due to rapid temperature changes.

The sample is then cut in half along its length using a diamond saw, this ensures that there is a flat side to mount the sample and also doubles the area available for thinning whilst reducing the thickness to be ground down. The sample is mounted (flat side down) on a glass slide using beeswax melted on a hotplate. The slide is then taped to a brass block and ground on a Metaserv 2000 polisher/grinder wheel using progressively finer grit of 240, 400 and then 1200 grit and then two polishing stages using 6 μm and 1 μm diamond paste. The samples are then flipped over and ground in the same manner to approximately 20 μm thickness before polishing to a mirror finish. In later work a South Bay Technology Lapping fixture 150 was used instead of a brass block to keep the sample level and enable the

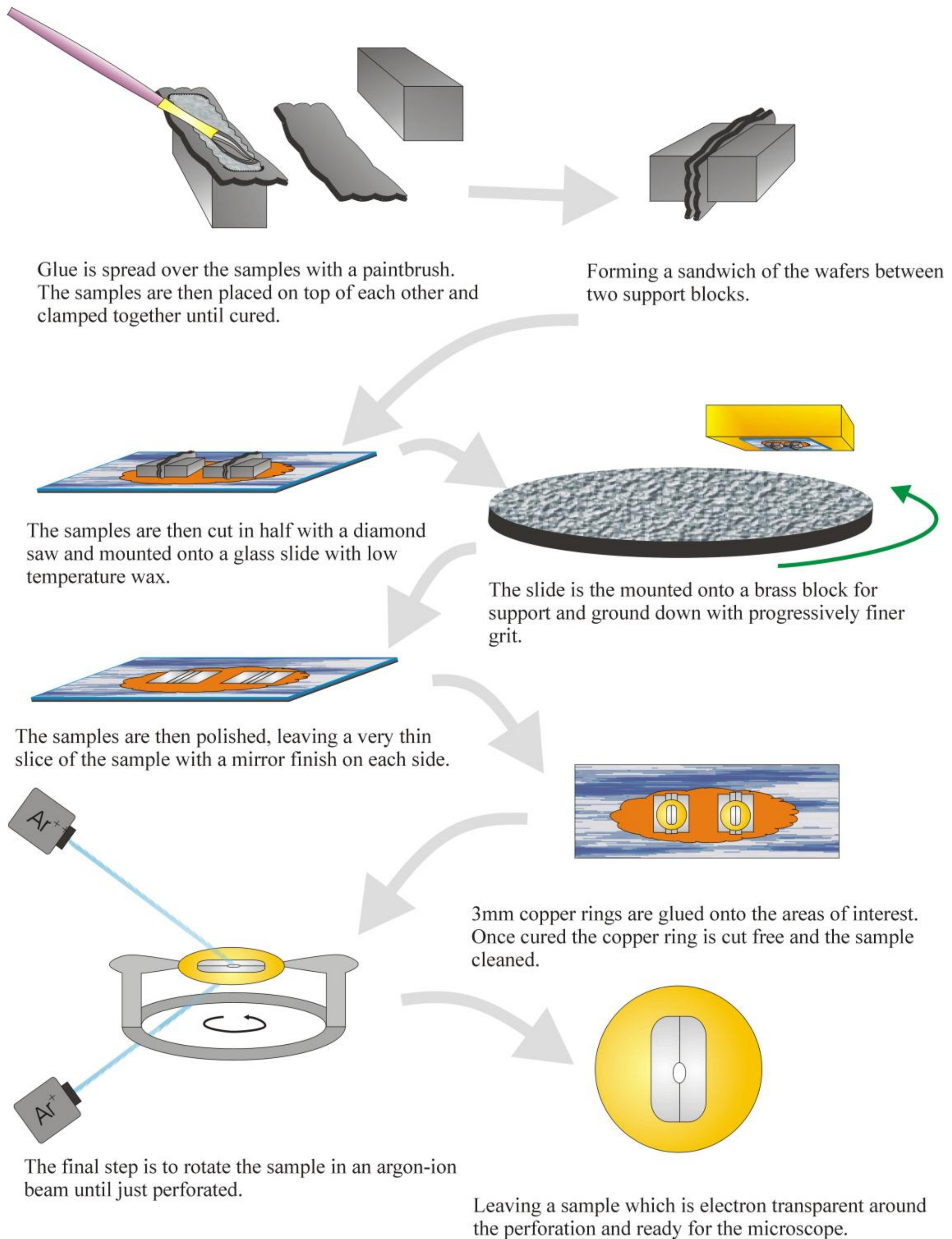


Figure 3.6 Schedule for preparing cross-sectional TEM samples as described in the text.

thickness of the sample to be monitored. At this stage the sample should be around 20 μm thick and polished on both sides. A copper ring of 3 mm diameter with a slot of dimensions 1 mm × 2 mm is then glued with the slot along the glue line of the samples using a very thin layer of Araldite Precision adhesive. This is allowed to cure overnight at room temperature. Once cured the copper rings are cut free of the surrounding silicon and removed from the slide using a hot plate to melt the wax. They are then cleaned in hot propan-2-ol and dried on filter paper.

The final stage of preparation is to erode the sample until it is just punctured in a Gatan PIPS ion beamer. This accelerates a beam of argon ions at the sample from both sides using a potential of 4.5 kV. The sample is ready for the TEM when a hole just appears in the centre of the sample. The material surrounding the hole is then electron transparent and suitable for examination in a TEM.

3.1.8 Diffraction

In order to understand the phenomena of diffraction a wave-like view of matter must be taken. Because of wave-particle duality, wave-like properties can be attributed to all manifestations of matter and energy. In particular X-rays and electrons can both experience diffraction effects. If a wave is considered to impinge on a crystalline material then an elastically scattered wave may emanate from each and every atom. These waves will usually be in phase with the incident wave (or at least have a constant phase relationship), hence they are *coherent*. In Figure 3.7 the incident wave (red) has wavevector $\boldsymbol{k}$ Consider the scattered wave (blue) in one direction only, this has wavevector $\boldsymbol{k'}$. Here the scattering from only two atoms with displacement vector $\boldsymbol{r}$ has been considered.

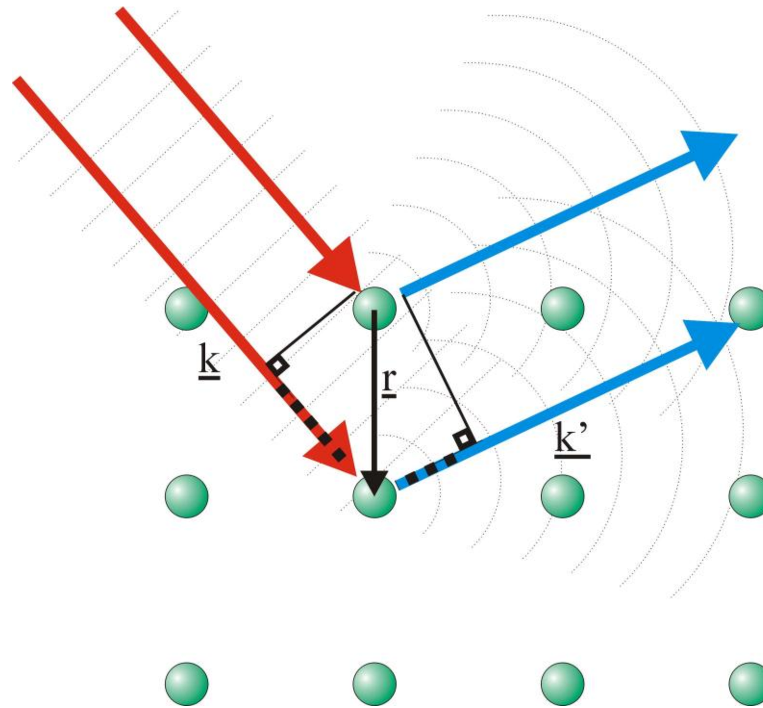


Figure 3.7 Schematic diagram of the phenomenon of diffraction from a crystal lattice. The dotted line represents the path difference between waves scattered from the two different atoms.

In general, the rays reflecting from the two atoms have a path difference (dotted) and hence phase shift (ϕ) which can be shown to be given by $\phi = \boldsymbol{r} \cdot \Delta \boldsymbol{k}$. In order to calculate the total intensity (*I*) scattered in a general direction, ***k'***, a sum over all possible scattering centres with displacement vector ***r*** must be performed.

$$I = \sum_{\boldsymbol{r}} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \tag{3.2}$$

For a perfect lattice, this has a none-zero result only if the difference in wavevectors ($\Delta\boldsymbol{k}$) is a *reciprocal lattice vector* ($\boldsymbol{G}$), giving the condition for diffraction

$$\Delta\boldsymbol{k} = \boldsymbol{G} \tag{3.3}$$

a form of the well known Braggs Law. This is often quoted in the simplified form

$$2d_{hkl} \sin\theta = \lambda \tag{3.4}$$

where d_{hkl} is the separation of the planes with miller indices *h, k* and *l*, λ is the wavelength of the incident wave and θ is the angle of incidence measured from the diffracting planes.

3.1.9 The Reciprocal Lattice

As described in section 2.1 the position in real space of the atoms of a crystal are described by the crystal lattice. In diffraction theory it is more useful to consider the *reciprocal lattice*. The vectors describing the reciprocal lattice are derived from the real lattice vectors by the following transformations

$$\boldsymbol{a}^* = \frac{\boldsymbol{b} \times \boldsymbol{c}}{abc} \qquad \boldsymbol{b}^* = \frac{\boldsymbol{c} \times \boldsymbol{a}}{abc} \qquad \boldsymbol{c}^* = \frac{\boldsymbol{a} \times \boldsymbol{b}}{abc} \tag{3.5}$$

where $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ are the real lattice vectors and the superscript (*) represents their respective reciprocal lattice vectors. Just as the real lattice describes all the positions of the crystal in real space, the reciprocal lattice describes a set of points in *reciprocal space*. Each reciprocal lattice point represents a set of equivalent planes in the real lattice.

If the incident wavevector is plotted in reciprocal space the locus of all possible elastically scattered wavevectors will describe a sphere (the Ewald sphere) and is depicted in Figure 3.8. Whenever the Ewald sphere crosses the reciprocal lattice, the Bragg condition (equation (3.4)) will be satisfied and diffraction can occur. For a perfect lattice (hence infinite in all directions) the reciprocal lattice has infinitesimally small points. This means that the Bragg condition must be obeyed *exactly* if there is to be any diffracted intensity. Clearly, this will only happen very

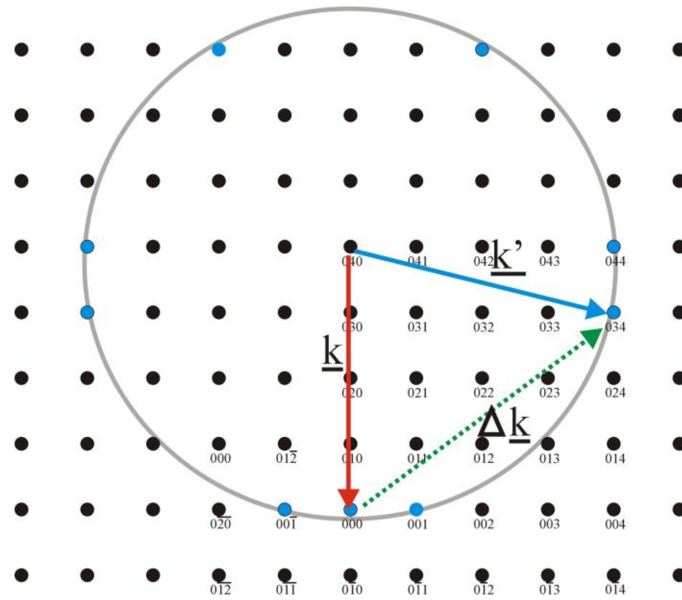


Figure 3.8 Representation of the Ewald sphere construction. When the incident wavevector (red) and scattered wavevector (blue) are separated by a reciprocal lattice vector, diffraction can occur. Possible diffraction vectors are shown as blue dots.

occasionally and hence, for a given incident wave, there will be very few diffracted rays. In a TEM, however, the diffraction pattern consists of many spots even though the Bragg condition cannot be satisfied exactly for all of them. This occurs because in order to produce a TEM image the sample must be very thin parallel to the electron beam. Since the real lattice has restricted dimensions parallel to the electron beam, the reciprocal lattice is extended in this direction. Instead of infinitesimal points in space, the reciprocal lattice consists of a set of rod like shapes often called *relrods*. Since the Ewald sphere must pass through the reciprocal lattice for

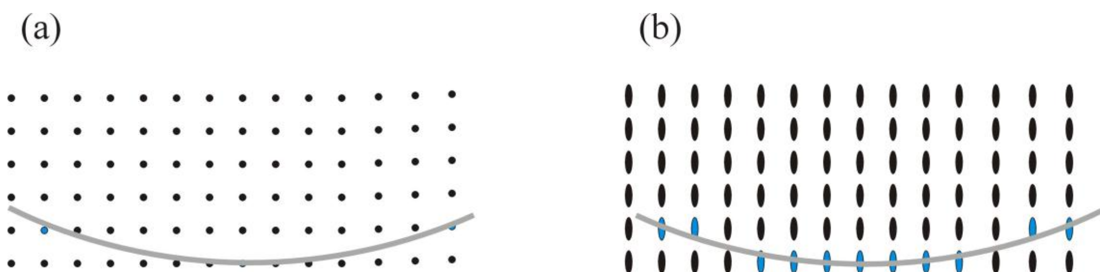


Figure 3.9 In the perfect lattice (a) the Ewald sphere only crosses a few points. In a thin foil the reciprocal lattice points are stretched into rods, so the Ewalds sphere crosses much more points (b).

diffraction to occur, the long shape of the relrods for a thin sample allows many more spots to occur in the diffraction pattern as shown in Figure 3.9.

3.1.10 Structure Factors

It was shown earlier that the diffraction from a set of scattering centres can be calculated from equation (3.2). For a real lattice, $\boldsymbol{r}$ can be split into several vectors as follows (see Figure 3.10)

$$\boldsymbol{r} = \boldsymbol{R}_{UC} + \boldsymbol{R}_{Lattice} + \boldsymbol{R}_{Structure} + \boldsymbol{r}_{Electron} \quad (3.6)$$

where

$\boldsymbol{R}_{uc}$ represents the position of the Unit Cell,

$\boldsymbol{R}_{Lattice}$ represents the vector from the origin of the unit cell to each atom in the lattice,

$\boldsymbol{R}_{Structure}$ represents the vector from each lattice point to the basis atoms,

$\boldsymbol{r}_{Electron}$ represents the position of the electrons within each atom.

From this Equation (3.2) can be recast into

$$I = \sum e^{i(\boldsymbol{R}_{UC} + \boldsymbol{R}_{Lattice} + \boldsymbol{R}_{Structure} + \boldsymbol{r}_{Electron})\cdot\Delta\boldsymbol{k}} \quad (3.7)$$

and because of the exponential dependence this can be split into four factors:

$$I = \sum_{\boldsymbol{R}_{UC}} e^{i\boldsymbol{R}_{UC}\cdot\Delta\boldsymbol{k}} \sum_{\boldsymbol{R}_{Lattice}} e^{i\boldsymbol{R}_{Lattice}\cdot\Delta\boldsymbol{k}} \sum_{\boldsymbol{R}_{Structurue}} e^{i\boldsymbol{R}_{Structurue}\cdot\Delta\boldsymbol{k}} \sum_{\boldsymbol{r}_{Electron}} e^{i\boldsymbol{r}_{Electron}\cdot\Delta\boldsymbol{k}} \quad (3.8)$$

The first term is known as the *geometric factor* and for a cubic system only has none zero result for $\Delta\boldsymbol{k}$ being a reciprocal lattice vector with (any) integer Miller indices, *hkl.* The second and third terms are known as the *structure factor*. This can

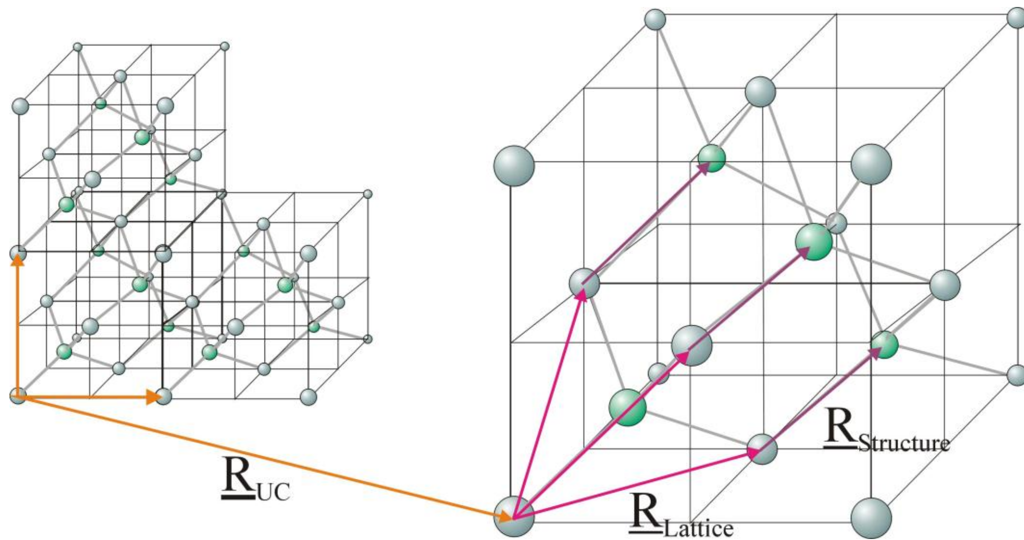


Figure 3.10 The crystal structure is made up of cubic unit cells. Any atom in the crystal can be described by the addition of a unit cell vector, a lattice vector and a basis vector.

give a zero result for some reciprocal lattice vectors and consequently lead to *missing orders*. The final term is known as the *form factor* and represents the scattering from one atom. This is generally considered to be constant for a given atom and proportional to the atomic number, *Z*.

The diamond crystal structure is based on a cubic unit cell, so any Miller indices (*hkl*) will have a non-zero geometric factor. The FCC lattice has 4 atoms relative to the cubic unit cell at $(0,0,0)$, $\left(\frac{1}{2},\frac{1}{2},0\right)$, $\left(\frac{1}{2},0,\frac{1}{2}\right)$ and $\left(0,\frac{1}{2},\frac{1}{2}\right)$. Giving the first structure factor as

$$\sum_{\boldsymbol{R_{Lattice}}} e^{i\boldsymbol{R_{Lattice}}\cdot\Delta\boldsymbol{G}} = 1 + e^{\pi i(h+h)} + e^{\pi i(h+l)} + e^{\pi i(k+l)} \tag{3.9}$$

This has a non-zero result for *h*, *k* and *l* all even or all odd, giving the selection rule for the FCC lattice as *h*, *k* and *l* all even or all odd. The second structure factor is due to the basis attached to each lattice point of an atom at $(0,0,0)$ and $\left(\frac{1}{4},\frac{1}{4},\frac{1}{4}\right)$, which gives:

$$\sum_{R_{Structurue}} e^{iR_{Structurue}\cdot\Delta k} = 1 + e^{\frac{\pi}{2}i(h+k+l)} \quad (3.10)$$

This has a non-zero result only if the value of $h+k+l$ is *not* an odd multiple of 2. Hence, the selection rule for diffraction in the diamond structure is h, k, and l must have the same parity (i.e. all even or all odd) and $h+k+l$ is not an odd multiple of 2. So (001), (002), (110) and (003) are all forbidden or *missing orders*. Whereas (004), (111) and (220) are all allowed and expected in a diffraction pattern of silicon.

3.1.11 Double Diffraction

As shown above the (002) diffraction is forbidden and should not be seen in a diffraction pattern of silicon. However, in a TEM the (002) diffraction spot often appears to be present. This is not a breakdown of the selection rules but is due to *double diffraction*. This occurs if a diffracted wave undergoes a second diffraction. For example, if an electron beam is incident along the [110] direction it may be diffracted by the $(1\bar{1}1)$ planes. This diffracted wave may then be diffracted by the $(\bar{1}11)$ planes. The total diffracting vector would then appear to be due to the forbidden (002) planes whereas in reality it is due to the two allowed $(1\bar{1}1)$ and $(\bar{1}11)$ planes.

3.1.12 Kikuchi Lines in the TEM

The diffraction pattern in a TEM is due to elastically scattered electrons that have met the Bragg condition for diffraction. However, for fairly thick samples another feature of the diffraction pattern can be seen. Kikuchi lines are pairs of long

lines of higher or lower intensity than the background which criss-cross the diffraction pattern. This phenomenon is due to inelastically scattered electrons.

When electrons are inelastically scattered they lose coherence with the incident wave. Referring to Figure 3.11 the point at which they are scattered can be considered a new (instantaneous) source of electrons. This new source is within the sample and produces electrons that may have many different wavevectors (but approximately the same magnitude) than the incident beam. In general, these electrons simply give a background intensity in the diffraction pattern. However, some of these electrons will have the right angle of incidence to be diffracted by the planes which the source lies between. The possible diffracted rays describe a double cone in real space and it is the intersection of these cones with the viewing screen that gives rise to the pair of Kikuchi lines. If the diffracting crystal is rotated the

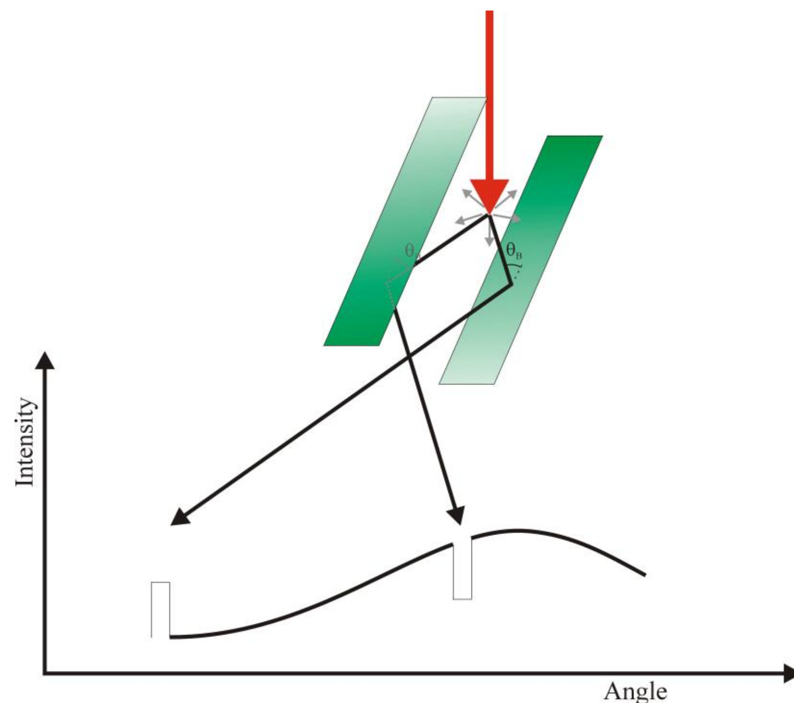


Figure 3.11 Diagram showing the formation of Kikuchi lines. See text for explanation.

Kikuchi lines move as if they are rigidly attached to the lattice. This is because the electrons are scattered in all directions, so there will always be electrons scattered so as to have the Bragg angle. Hence, there will always be Kikuchi lines at the Bragg angle from the diffracting planes. If these planes are rotated, so will the Kikuchi lines. These Kikuchi lines can be very useful for the microscopist to orientate the crystal in the TEM and to correctly identify specific diffraction conditions.

3.1.13 Contrast in the TEM

Contrast in the TEM comes mainly from the use of apertures. The electron beam hitting the sample is partially diffracted. Any change in the lattice parameter of the crystal causes a change in the relative intensities of the diffracted and un-diffracted rays. However, without an objective aperture both rays are allowed to form the image and any contrast information is lost. In order to improve contrast an aperture is placed to allow only one beam, usually the un-diffracted beam (bright-field), to form the image.

In order to get the best contrast in the TEM it is necessary to set the correct diffraction condition. The highest contrast occurs when there is strong diffraction from a set of planes whose spacing changes across the feature. This is because as the spacing of the planes changes, the intensity of the diffracted ray changes and hence there is strong contrast in the image. In a strained pseudomorphic layer the largest change in spacing occurs in the growth direction, the [001] direction. Hence to image the strain, diffraction from these planes should be maximised. To do this the microscope should be set to have the (000) and (004) spots strongly excited (the (001), (002) and (003) diffraction spots are forbidden as explained earlier). To excite

a specific diffraction spot, the sample is orientated so that the Kikuchi lines from the spot that is to be excited is coincident with this spot and the (000) spot. In this case only these two spots are strongly excited and the sample is said to be in the *two-beam condition*.

When imaging dislocations in a TEM the relationship between the Burgers vector, ***b***, and the diffracting vector, ***g***, plays an important role. As can be seen in the Figure 3.12(a) when ***g*** is perpendicular to ***b*** (hence ***g·b*** =0) there is little change in spacing of the diffracting planes around the dislocation, hence there is little or no contrast. When ***g*** is parallel to ***b*** as in Figure 3.12(b) the change in spacing of the diffracting planes is the greatest so there is the largest contrast in the final TEM image. For this reason if a dislocation is to be imaged the diffraction condition is

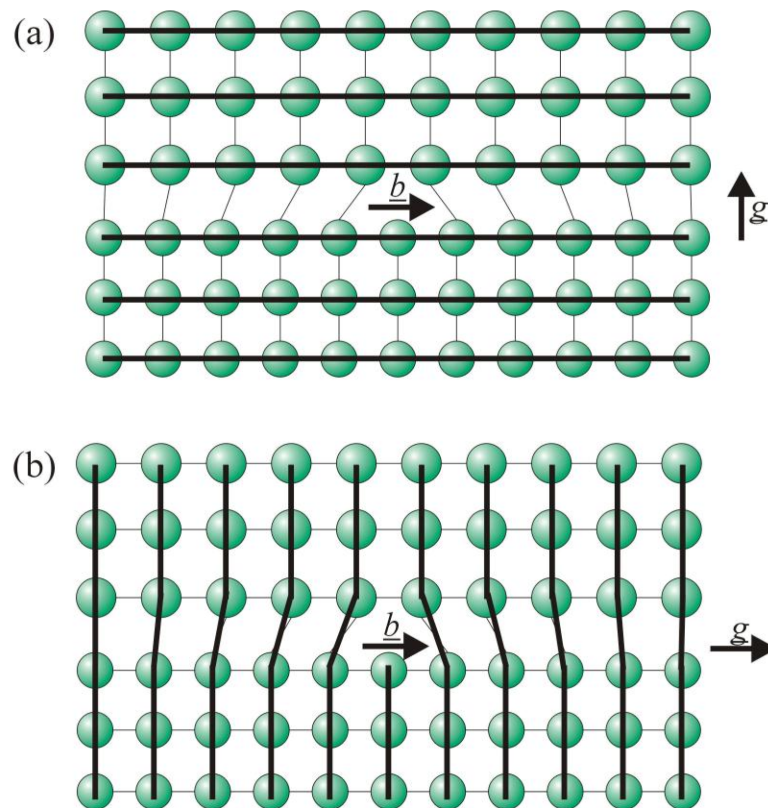


Figure 3.12 Diagram showing the effect of a dislocation on two different sets of planes, in (a) ***g*** is perpendicular to ***b*** and the planes are largely undistorted. In (b) ***g*** is parallel to ***b*** and there is a large distortion to the atomic planes.

chosen to be parallel to the expected Burgers vector. In the silicon system this is usually the [110] directions and the (220) diffraction spot is excited (the (110) spot being forbidden). It is possible to determine the Burgers vector of an unknown dislocation from the TEM by finding 2 non-parallel diffraction conditions in which the contrast of the dislocation is low. Since this only happens when $\boldsymbol{g}\cdot\boldsymbol{b} = 0$, it should be possible to determine $\boldsymbol{b}$, this is know as $\boldsymbol{g}\cdot\boldsymbol{b}$ (*g* dot *b*) analysis.

3.2 X-ray Determination of Composition and Strain

As shown in section 3.1.8 the condition for diffraction of a wave is given by the Bragg equation

$$2d_{hkl}\sin\theta = \lambda \tag{3.4}$$

where d_{hkl} represents the spacing of the planes with Miller indices *h*, *k* and *l* and θ is the angle of diffraction measured from the *hkl* plane. If we consider a tetragonally distorted layer grown on a silicon substrate then the spacing (d_{hkl}) of the layer will not in general be the same as that of the substrate. There may also exist a tilt of the crystallographic planes (ϕ) of the layer with respect to that of the substrate. There may also be a tilt of the sample with respect to the sample holding stage, hence all measurements are taken with reference to the bragg angle of the silicon substrate which is well defined.

With reference to Figure 3.13 the measured angle $\theta^{measured}$ of the diffracted ray from the layer is given by

$$\theta_{layer}^{measured} = \theta_{layer} - \phi \tag{3.11}$$

If the sample is now rotated about the vertical axis by 180° the new value of $\theta^{measured}$

is given by

$$\theta_{layer}^{measured,180} = \theta_{layer} + \phi \tag{3.12}$$

Since the angle of diffraction is to be referenced to the substrate we can define the difference between substrate and layer peaks by

$$\Delta\theta = \frac{(\theta_{layer}^{measured,0} - \theta_{substrate}^{measured,0}) + (\theta_{layer}^{measured,180} - \theta_{substrate}^{measured,180})}{2} \tag{3.13}$$

and the angle of tilt (ϕ) by

$$\Delta\theta = \frac{(\theta_{layer}^{measured,0} - \theta_{substrate}^{measured,0}) - (\theta_{layer}^{measured,180} - \theta_{substrate}^{measured,180})}{2} \tag{3.14}$$

The angle of the diffraction peak from the silicon substrate can be shown from equation (3.4) to be

$$\theta_{substrate} = \sin^{-1}(\frac{\lambda}{2d_{hkl}^{substrate}}) \tag{3.15}$$

where λ is the wavelength of the X-ray source and d_{hkl} is given by

$$d_{hkl} = \frac{1}{\sqrt{\left(\frac{h}{a_x}\right)^2 + \left(\frac{k}{a_x}\right)^2 + \left(\frac{l}{a_z}\right)^2}} \tag{3.16}$$

for a tetragonal unit cell.

The angle of diffraction for the layer is then given by

$$\theta_{layer} = \theta_{substrate} + \Delta\theta \tag{3.17}$$

and from this the spacing of the planes within the layer, d_{hkl}^{layer}, is given by

$$d_{hkl}^{layer} = \frac{\lambda}{2\sin(\theta_{layer})} \tag{3.18}$$

Since the layer will in general be tetragonally distorted we can characterise the unit cell with the in plane and out of plane lattice constant a_x and a_z. If we assume the

material is elastic with a Poisson's ratio ν, the in and out of plane lattice constants are related to the relaxed lattice constant, *a(x)*, by

$$a(x) = a_x + \frac{1-\upsilon}{1+\upsilon}(a_z - a_x) \tag{3.19}$$

In order to determine a_x and a_z we need to find d_{hkl} for two non-parallel sets of planes. Usually the (004) reflection is used and an asymmetric reflection, usually (224) or (115). From these measurements a_x and a_z can be calculated from

$$a_z = a_{Si}\left(\frac{\sin(\theta^{004}_{\text{substrate}})}{\sin(\theta^{004}_{\text{layer}})}\right) \tag{3.20}$$

and

$$a_x = \left(\sqrt{h^2 + k^2}\right)\frac{\lambda}{2\sin(\theta^{hkl}_{layer})\sin(\phi)} \tag{3.21}$$

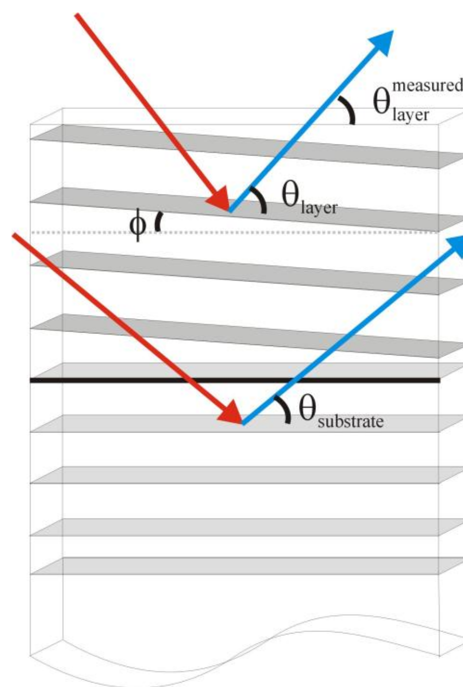


Figure 3.13 Diagram of the diffracted rays from the (004) planes of the substrate and an epitaxial layer. The epitaxial layers may have a tilt φ with respect to the substrate which must be taken into account

Using these values of a_x and a_z in equation (3.19), x can be calculated by numerically solving the correction to Vergard's law

$$a(x) = a_{Si}(1-x) + a_{Ge}x + 0.02733x^2 - 0.0278x \text{ Å} \tag{3.22}$$

and the state of relaxation is defined as

$$\mathrm{Re}laxation = 100\left(\frac{a_x - a_{Si}}{a(x) - a_{Si}}\right) \tag{3.23}$$

It should be noted that this method is only accurate for high states of relaxation (~ 90 % or higher),[84] otherwise a full reciprocal space map needs to be taken.

3.3 SIMS

Secondary Ion Mass Spectrometry (SIMS) utilises a beam of ionised *primary* particles to sputter ions from the surface of the sample. As the surface erodes from the impact of the primary ions, many types of particles are released from the surface. In SIMS the particles of interest are the *secondary ions*, these are particles from the surface of the sample that have been ejected in an ionised state. The masses of these secondary ions are recorded using a mass spectrometer so that the type and quantity of different atomic species can be determined as a function of the erosion time. It is assumed that the erosion rates of the sample remain constant, hence by measuring the crater depth after the sample has been analysed in the SIMS apparatus, the erosion rate can be determined. Using this erosion rate it is straightforward to convert the mass vs. time data into a mass vs. depth profile.

The instrument used in this study was an Atomika – 4500 SIMS Profilometer, this utilises a floating low-energy ion gun so that a high beam intensity is achievable at a low beam energy (typical beam energies are 500 - 1000 eV). The primary ions used in this study were O_2^+, although the equipment has the capability of using Cs^+

which is particularly useful if an oxygen depth profile is required (O_2^+ being clearly unsuitable for this purpose). The primary ion beam is rastered over a square of width 180 - 250 μm and the mass spectra recorded as a mesh of 256×256 pixels.

When the primary beam interacts with the surface of the sample a variety of outcomes are possible as shown in Figure 3.14. For SIMS analysis the only useful product is the ejection of ionised particles from the surface of the sample which accounts for only ~ 1% of the sputtered particles. These particles are ejected from the sample by momentum transfer from the incoming primary ions. The primary ions typically have a penetration depth of a few nm's, which limits the maximum resolution of the instrument. Along with these secondary ions, other particles are removed including neutral atoms, small molecular species and particles which have been knocked from deeper inside the sample. These particles reduce the sensitivity

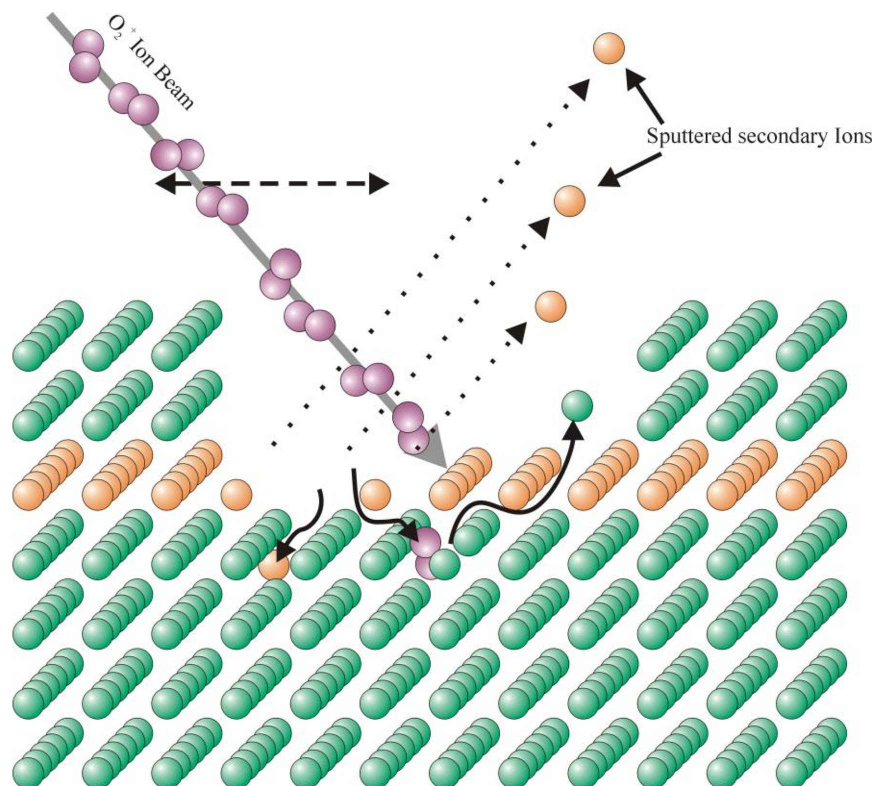


Figure 3.14 Diagram of the process of sputtering during SIMS. As the primary ion beam is scanned across the sample, it sputters secondary ions from the surface. Other effects can lead to inter-mixing.

and resolution of the SIMS analysis and so the beam condition is set to minimise their effect. Other effects which reduce the resolution of the SIMS include:

- Intermixing - This is caused by the impinging primary ions knocking surface atoms deeper into the sample where they are later sputtered as secondary ions. This effectively mixes the surface atoms with the atoms deeper into the sample (< penetration depth) and causes a broadening of the SIMS depth profile.
- Crater sidewall effects – As the sample is eroded the impinging primary ions can erode the sides of the crater, this smears the profile.
- Uneven surfaces – If the surface of the sample is uneven the averaging over the analysed area will cause a broadening of the SIMS profile.
- Dislocations – A highly dislocated sample can effect the erosion rates of the sample and lead to a distortion of the profile.

Steps can be taken to minimise these effects. To reduce intermixing the energy of the primary ion beam is kept low (the floating low-energy ion gun is invaluable for this reason) also increasing the angle of incidence reduces the penetration depth. The resolution of the depth profile in the Atomica SIMS profilometer can be reduced to about ~ 1 nm, although in order to profile deeper samples the resolution must be sacrificed for an increased erosion rate. Crater sidewall effect are minimised by the use of gates. The array of mass data is sub-divided into concentric groups of pixels known as *gates*, each of which contains all the data from a given radius from the centre of the crater. The use of these gates allows the effect of the sides of the crater to be eliminated, this is done by selecting smaller and smaller gate widths, and so collecting data from a further away from the crater wall, until the profile remains

unchanged with further reduction in gate width. It can then be safely assumed that the profile is free from sidewall effects. The resolution of the profile is also limited by the ability to measure the secondary ions that are sputtered from the sample. The mass spectrometer used in the Atomika - 4500 is able to detect dopant concentrations down to ~ 5×10^{16} cm^{-3} (dependant on species).

The measured flux of a particular species of secondary ions (I_i) of a particular species is dependant on many factors

$$I_i = \dot{z}\beta_i\eta_i ANc_i \tag{3.24}$$

where

I_i = count rate of species i (s^{-1}),

$\dot{z}$ = sample erosion rate (cm s^{-1}),

β_i = fraction of species i sputtered as an ion,

η_i = detection efficiency factor for the emitted ions,

A = analysed area (cm^2),

N = atomic density of sample (cm^{-3}),

c_i = atomic concentration of species i.

The sample erosion rate depends on

$$\dot{z} = \frac{j_p Y_m}{N} \tag{3.25}$$

where

j_p = flux density of primary ions (cm^2 s^{-1}),

Y_m = sputter yield (atoms/primary ion).

To calculate the composition from these equations needs an accurate determination of all the parameters involved. Unfortunately, this proves to be very difficult since parameters like Y_m and β_i are strong functions of the composition and so are very sensitive to changes in the matrix (the matrix effect). Instead, the quantification of the Ge composition in this study was achieved in a manner described by Dowsett *et al.*[85] Since the Si yield shows a linear relationship with Si concentration, this was used for the quantification. A reference level for pure silicon was used for calibration, the silicon substrate being convenient for this (otherwise a separate silicon sample could be analysed at the same time). The silicon composition in the sample was then defined as

$$C_{Si} = \frac{\frac{1}{N_{sample}} \sum I_{Sample}}{\frac{1}{N_{Reference}} \sum I_{Reference}} \tag{3.26}$$

where

C_{Si} = atomic concentration of silicon,

N = number of data points of the sample / Si reference used in the summation,

I = count rate of the sample / Si reference (s^{-1}).

The Ge concentration (x) is then calculated from

$$x = 1 - C_{Si} \tag{3.27}$$

The accuracy of this method has been determined to be better than 1 % up to a Ge fraction of at least $x = 0.3$ by independent measurements using XRD and Rutherford Back Scattering (RBS).[85]

3.4 Electrical Measurements.

To characterise the electrical properties of channels grown on top of SiGe virtual substrates the method first outlined by van der Pauw[86] was used. The van der Pauw (vdP) method requires that several conditions of the electrically active channel are met:-

- The charge carriers are confined to a thin lamella so that the conduction can be considered 2-dimensional.
- All contacts to the charge carriers are made at the periphery of the sample and have negligible size compared to the sample.
- The channel is homogenous and isotropic throughout the measured area
- The channel is singularly connected, i.e. there are no holes within the conducting layer.

Other than these constraints the sample can be of any size or shape, however irregular. Under these conditions the sheet resitivity, the carrier sheet concentration and the carrier drift mobility can be calculated.

3.4.1 Resistivity

In the vdP method four contacts are made to the edge of the sample to be measured, these are labelled A, B, C and D in *cyclic* order around the edge of the sample (see Figure 3.15(a)). A current is passed through adjacent contacts (A and B, say) and the resulting potential difference is measured across C and D, this defines a "resistance" $R_{AB,CD}$

$$R_{AB,CD} = \frac{V_D - V_C}{i_{AB}} \quad (3.28)$$

where

V_D and V_C are the potentials measured at contacts D and C,

i_{AB} is the current through terminals A and B.

Similar resistances can be measured by cyclically permuting the contacts A, B, C and D producing a total of four resistances $R_{AB,CD}$, $R_{BC,DA}$, $R_{CD,AB}$ and $R_{DA,BC}$. According to van der Pauw, these resistances define the sheet resistivity (ρ_s) of the conducting layer as

$$\rho_s = \left(\frac{\pi}{\ln 2}\right)\left(\frac{R_{AB,CD} + R_{BC,DA} + R_{CD,AB} + R_{DA,BC}}{16}\right) \times f\left[\left(\frac{R_{AB,CD}}{R_{BC,DA}}\right) + \left(\frac{R_{BC,DA}}{R_{CD,AB}}\right) + \left(\frac{R_{CD,AB}}{R_{DA,BC}}\right) + \left(\frac{R_{DA,BC}}{R_{AB,CD}}\right)\right] \tag{3.29}$$

where the f- factor is defined by the root of the transcendental equation

$$\cosh\left[\left(\frac{R_{AB,CD} - R_{BC,DA}}{R_{AB,CD} + R_{BC,DA}}\right)\frac{\ln 2}{f}\right] = \frac{1}{2}\exp\left(\frac{\ln 2}{f}\right) \tag{3.30}$$

For highly symmetrical samples the value of f is close to unity and so the approximation

$$f \approx 1 - \left(\frac{R_{AB,CD} - R_{BC,DA}}{R_{AB,CD} + R_{BC,DA}}\right)^2 \frac{\ln 2}{2} \tag{3.31}$$

is usually good enough.

3.4.2 Hall Coefficient

The hall coefficient R_H is measured in a similar manner by passing a current through diagonally opposite contacts (A and C, say) whilst the sample is in a strong perpendicular magnetic field (B) and measuring the Hall potential V_H across the other two contacts (see Figure 3.15(b)) giving

$$R_H = \frac{10^4 V_H}{BI} = \frac{1}{n_s q} \quad cm^2 C^{-1} \qquad (3.32)$$

where

B is the magnetic field strength in Tesla (T),

V_H is the Hall voltage in volts (V),

I is the current in Amperes (A),

n_s is the sheet density (cm^{-2}),

q is the electronic charge $=1.602\times10^{-19}$ (C).

The mobility (μ) of the carriers can be calculated from these values of ρ_s and R_H by the equation

$$\mu = \frac{R_H}{\rho_s} \quad cm^2 V^{-1} s^{-1} \qquad (3.33)$$

3.4.3 Apparatus

The mobility and sheet concentration of the sample is usually measured in the temperature range 10 - 300 K using a Leybold closed cycle cryostat. Helium gas is continually pumped around the head of the cryostat in order reduce the temperature to as low as 10 K by an Air Products Helium pump. The temperature is raised by a 25 W resistive element heater attached to a Lakeshore DRC 91C temperature controller, the temperature is monitored by means of a Si diode thermometer bonded to the head of the cryostat. Oxford Varnish was used to bond the sample onto the brass mounting head of the cryostat to ensure good thermal contact but electrical isolation. A program written specifically for these measurements controls the

switching of the current and voltage contacts by means of a Keithley 705 scanner. Measurement of the current and potentials of the contacts is made by an E G & G Princeton Applied Research 5110 lock-in amplifier, giving a sensitivity down to nanovolts. During the measurement of the Hall voltage (V_H) the sample is subject to a magnetic field strength of 0.2 T from a Varian magnet powered by a 3-phase power supply. The cryostat head is placed between the poles of the magnet in such a way that the field is perpendicular to the sample.

The sample is prepared by evaporating a thin layer of Al onto the sample through a mask, leaving ~ 1 mm diameter holes arranged in a square array with a spacing of ~ 5 mm. The sample is then cleaved into squares in such a way the there is an aluminium contact at each corner of the sample. In order to diffuse the Al contacts into the conducting layer a 15 minute anneal at 450 °C under a N_2 ambient atmosphere is used. The electrical contact is then made by thin copper probes that are clamped onto the aluminium pads, the wires of the measurement system are then soldered onto these copper strips. A diagram of the sample and electrical contact arrangements for vdP rsistivity and Hall mobility is below shown in Figure 3.15

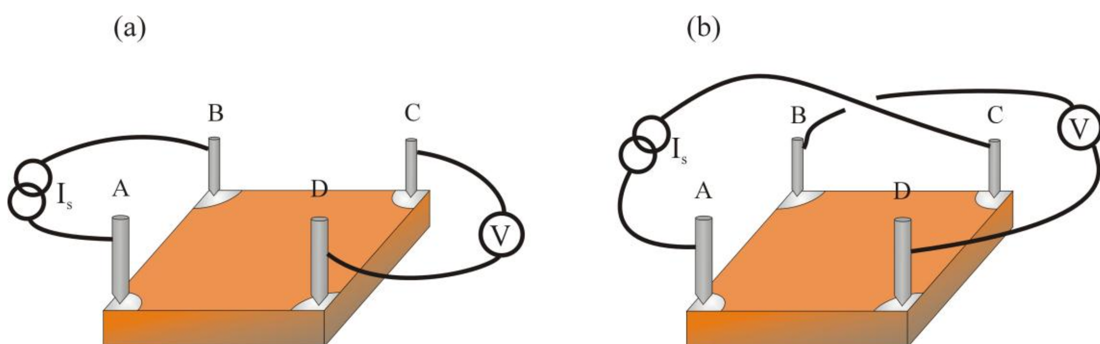


Figure 3.15 Diagram of the electrical connection made to the sample during (a) vdP resistivity measurements, and (b) Hall coefficient measurements.

4 Novel Terrace Graded Germanium Profile

4.1 Introduction

In this section the results of a series of experiments to try to produce high quality virtual substrates are presented. A high quality virtual substrate must be fully relaxed (greater than 95% relaxation as defined by equation (3.23)) and the threading dislocation density must be as low as possible.[57] Fitzgerald *et al*[46] suggest that threading dislocation densities of $10^5 – 10^6$ cm^{-2} are needed during the relaxation process of a virtual substrate. However, several authors have shown dislocation densities lower than this in virtual substrates grown by several different methods.[72,74,75,76,77] The surface topography should preferably be as flat as possible so that an electrical channel can be grown with a smooth interface (to reduce scattering) and that lithography is not affected.

4.1.1 Terrace Grading

This new technique for producing high quality virtual substrates consists of grading the germanium composition in a series of steps up to the required final composition. Each step consists of a bi-layer (or terrace) containing a *linearly* graded layer followed by a *uniform* layer of constant composition. The uniform layer has the same composition as the terminating composition of the preceding linearly graded layer. This sequence of steps is continued up to the desired final composition of the virtual substrate, with the starting composition of each bi-layer being the same as the final composition of the preceding step. A schematic of the germanium profile is shown in Figure 4.2. Because of the apparent similarity of the germanium profile to a terraced hill, this grading technique will be referred to hereafter as *Terrace Grading*.

In the first Terrace Graded wafer (**TG1**) the germanium composition was linearly graded from $x = 0$ to 0.10 over 200 nm. A uniform layer (terrace) of $x = 0.10$ was then grown on top to a thickness of 200 nm. The Ge composition was then increased using a similar linear grade/uniform layer sequence in steps of $x = 0.10$ up to a final composition of $x = 0.50$. The growth temperature started at 700 °C and was reduced in synchronisation with the Ge composition in order to suppress any strain induced surface roughening.[67] A schematic of the composition and growth temperature with depth is shown in Figure 4.2.

The reasoning behind this growth schedule is that as the initial linear grade is grown the misfit dislocations needed for strain relaxation begin to nucleate. The uniform layer grown on top adds to the excess strain driving dislocation propagation. The linear grade ensures that the misfit segments of the dislocations lie on different atomic planes. This should reduce the number of dislocation intersections, and because the Ge composition was kept low the probability of dislocation interactions is kept to a minimum. It is hoped that the thickness of the $x = 0.10$ step is great enough that this step is completely relaxed. In this way the growth of the next step occurs on top of a relaxed layer, no further dislocations should be needed in this lower layer. The threading dislocations nucleated in the lower layer will still be present with the exceptions of any that may have reached the edge of the wafer or have been annihilated. Any dislocations that have become pinned in the lower layer should now be able to glide free as the interface to be relaxed is now spatially separated from the dislocation interaction causing the pinning. The result of this is that the second graded layer should relax without the need for many further dislocations to be created. Also, any dislocations that are needed are more likely to be formed by the multiplication mechanisms explained in section 2.5.4. The

subsequent layers will relax in a similar manner, and because the pinning events are kept low and any pinned dislocations should be freed at each step, the final density of threading dislocations should be lower than a conventional linear graded layer.

The surface morphology of the virtual substrate should be smoother than conventional linearly graded virtual substrates as a result of the relaxation mechanism proposed. Since the misfit dislocation network needed to relax the virtual substrate is split into five spatially separated areas (the linearly graded regions) the strain fields associated with these networks should be completely independent. The strength of the strain fields should average to a much lower value than that of a buffer layer where the misfit networks are not separated. If relaxation has occurred through multiplication mechanisms like the MFR mechanism, then the dislocation pile-ups usually associated with these relaxation processes should also be split between the five graded layers. The consequence of this is that the crosshatch, caused by the accumulated surface steps of the dislocations associated with the MFR mechanism, should be significantly reduced. Figure 4.1 show schematically the principle of the separation of the dislocation pile-ups and the effect on the surface morphology.

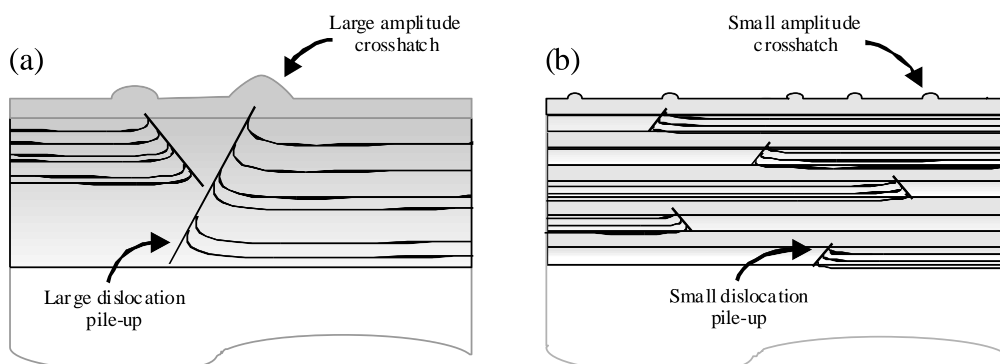
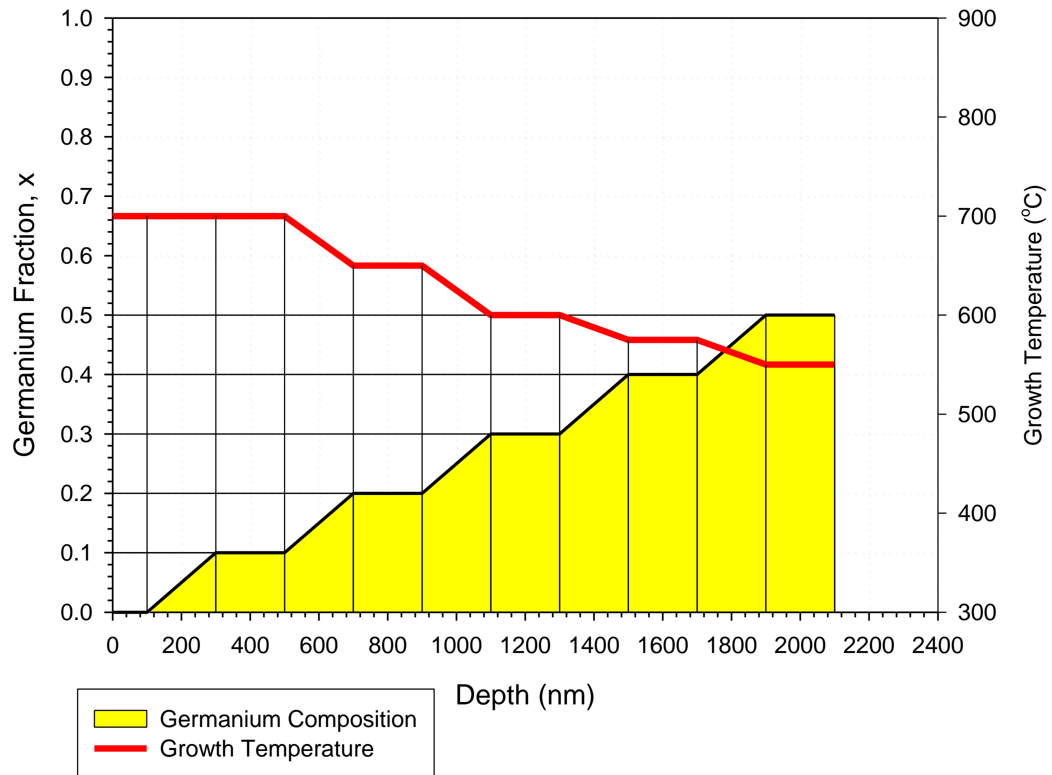


Figure 4.1 In a linearly graded layer (a) the dislocations form large pile-ups that penetrate the whole layer. Each pile-up causes a large surface undulation. In the terrace graded layers (b) pile-ups exist in the graded layers, but since these graded layers are reduced in size the pile-ups are smaller. Consequently, the surface undulations are smaller and more closely spaced.

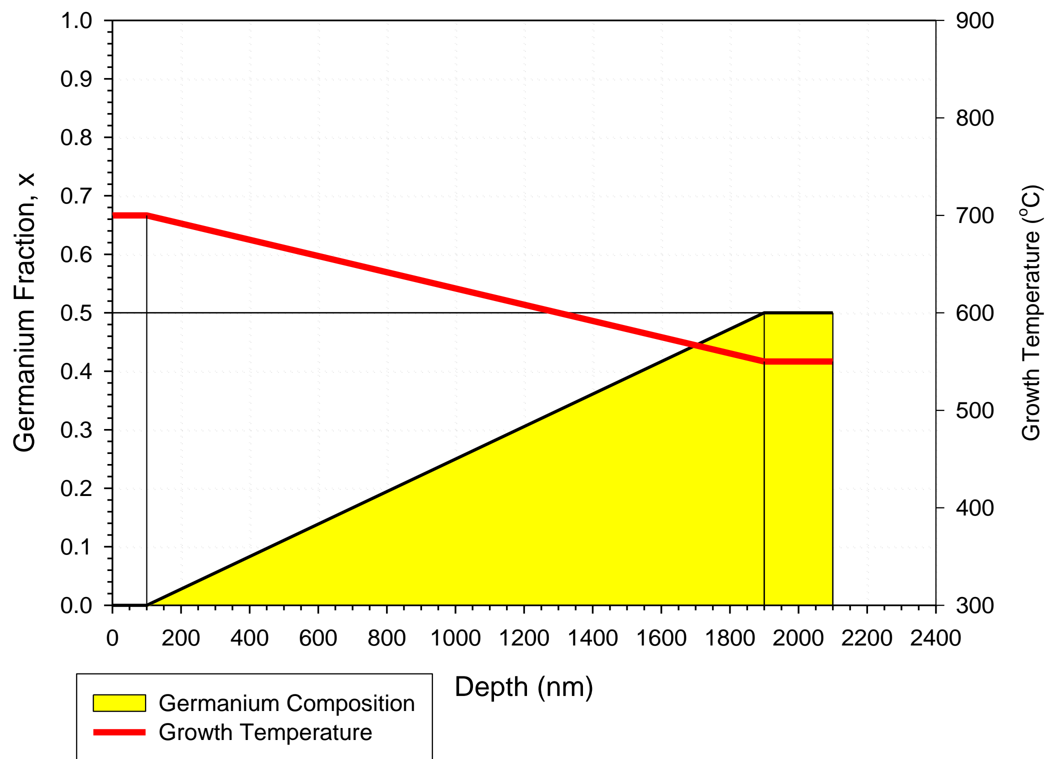
To assess the effectiveness of the new layer structure two wafers were grown using conventional grading techniques. The first wafer (**LG1**) was grown using a conventional Linearly Graded technique from $x = 0$ to 0.50. The thickness of this layer was kept the same (2 μm) as the terrace graded wafer (TG1) and the growth temperature was reduced synchronously with Ge composition in a similar manner. The linearly graded wafer should relax by the MFR multiplication mechanism described in section 2.5.4. It is expected that there will be large dislocation pile-ups that penetrate the entire structure leading to large surface undulations (crosshatch). A schematic of the growth specifications is summarised in Figure 4.3. The second wafer (**SG1**) was grown using a Step Grade. The germanium composition was increased with abrupt steps of $x = 0.10$ every 400 nm. The steps of $x = 0.10$ were chosen so that the closest comparison between the terrace graded profile and the step graded profile could be made. Again, the growth temperature was reduced in synchronisation with the germanium composition. The temperature was ramped down during the first 200 nm of each step and then held constant for the final 200 nm this is exactly the same as the sample TG1 so that all growth conditions are identical in the two samples. A summary of the growth specification is shown in Figure 4.4.

These three test wafers to determine the effectiveness of the proposed terrace graded germanium profile have been chosen to be as similar to each other as is possible. In this way any differences in the quality of the virtual substrates can be attributed to the different germanium profiles. The thickness of 2 μm is very thin for a $Si_{0.50}Ge_{0.50}$ virtual substrate, and so the quality of a conventional virtual substrates is not expected to be that high.



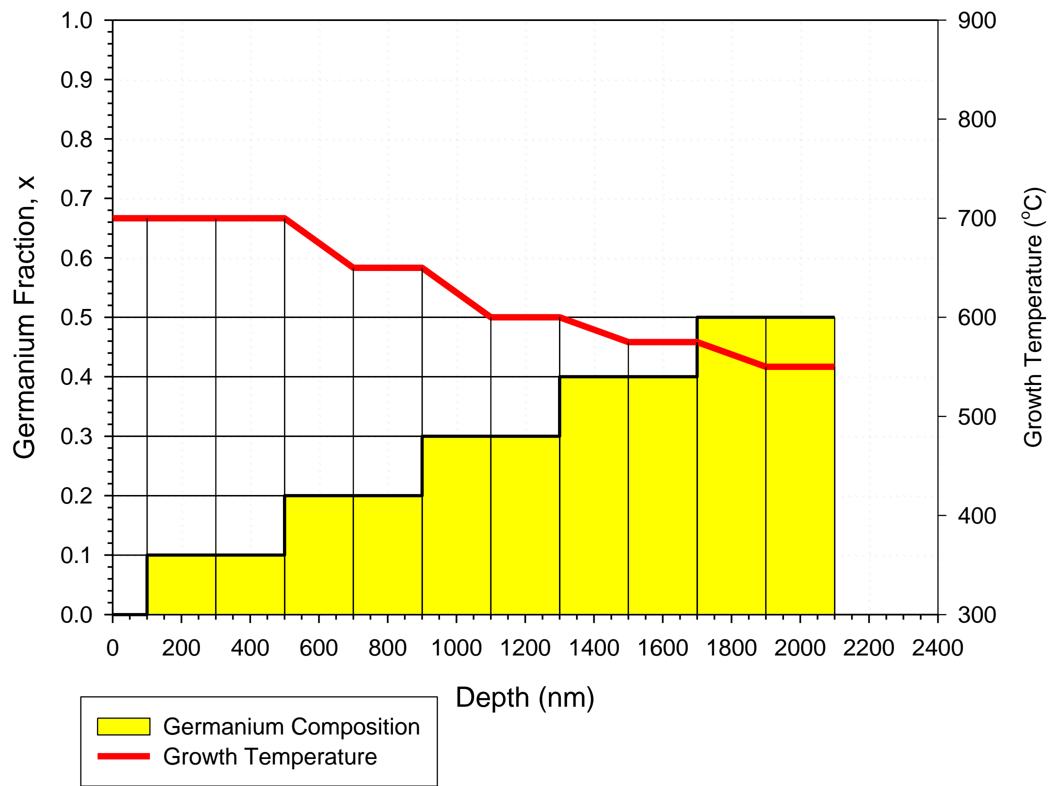
Step	Growth Temperature (°C)	Germanium Composition (%)	Thickness (nm)
0	700	0	100
1	700	0 – 10	200
2	700	10	200
3	700 – 650	10 – 20	200
4	650	20	200
5	650 – 600	20 – 30	200
6	600	30	200
7	600 – 575	30 – 40	200
8	575	40	200
9	575 – 550	40 – 50	200
10	550	50	200

Figure 4.2 Growth specification of wafer TG1.



Step	Growth Temperature (°C)	Germanium Composition (%)	Thickness (nm)
0	700	0	100
1	700 – 550	0 – 50	1800
2	550	50	200

Figure 4.3 Growth specifications for sample LG1.



Step	Growth Temperature (^{o}C)	Germanium Composition (%)	Thickness (nm)
0	700	0	100
1	700	10	200
2	700	10	200
3	700 – 650	20	200
4	650	20	200
5	650 – 600	30	200
6	600	30	200
7	600 – 575	40	200
8	575	40	200
9	575 – 550	50	200
10	550	50	200

Figure 4.4 Growth specification for sample SG1.

4.2 Structural Characterisation

4.2.1 XTEM

Figure 4.5(c) and Figure 4.6(c) show XTEM micrographs of the terrace graded wafer with the diffraction condition set to (220) for high dislocation contrast. They clearly show that the dislocations have formed in the linearly graded regions as expected. The uniform layers have threading dislocations passing through them, but a much reduced number of misfit sections. It is important to note that there is no evidence of any threading dislocations in the final uniform layer. Many XTEM images were taken and no threading dislocations were seen. An estimated 100 µm in length have been photographed, giving an upper limit of the threading dislocation density of approximately 10^7 cm^{-2}. The dislocation structure shows many dislocations that have piled-up on the same glide plane. This is a clear indication that relaxation has occurred by the MFR or similar multiplication mechanism. However, the dislocation pile-ups do not extend through the entire thickness of the virtual substrate as is usually seen in layers relaxed by the MFR mechanism. Instead, the pile-ups only span one or two of the linearly graded regions. Note also that no dislocations have been pushed deep into the substrate as can happen in layers relaxed by the MFR mechanism. This suggests that the separation of the dislocation pile-ups has been successful and that the misfit dislocation networks on each of the layers are indeed independent. If many dislocations *were* to form on the same glide plane, then the lower dislocations are usually pushed into the substrate by their mutual repulsion. Since the pile-ups have been separated by the uniform layers, the number of dislocations from each multiplication site is kept low enough that this has not occurred. However, there is a higher density of dislocations in the lower layers

Figure 4.5 Bright field XTEM micrographs taken with ***g*** = (220). The step-graded virtual substrate (a) shows most dislocations are confined to the interfaces of the steps, large dislocation pile-ups are seen to penetrate the entire thickness of the sample and the surface is seen to be rough. The linearly graded virtual substrate (b) also shows dislocation pile-ups which penetrate through the entire structure. Threading dislocations can be seen to penetrate the uniform capping layer and some dislocations have been pushed deep into the substrate. The terrace graded virtual substrate (c) shows clearly delineated dislocations in the graded layers. No threading dislocations are evident in the final uniform capping layer and dislocation pile-ups only extend through the graded layers.

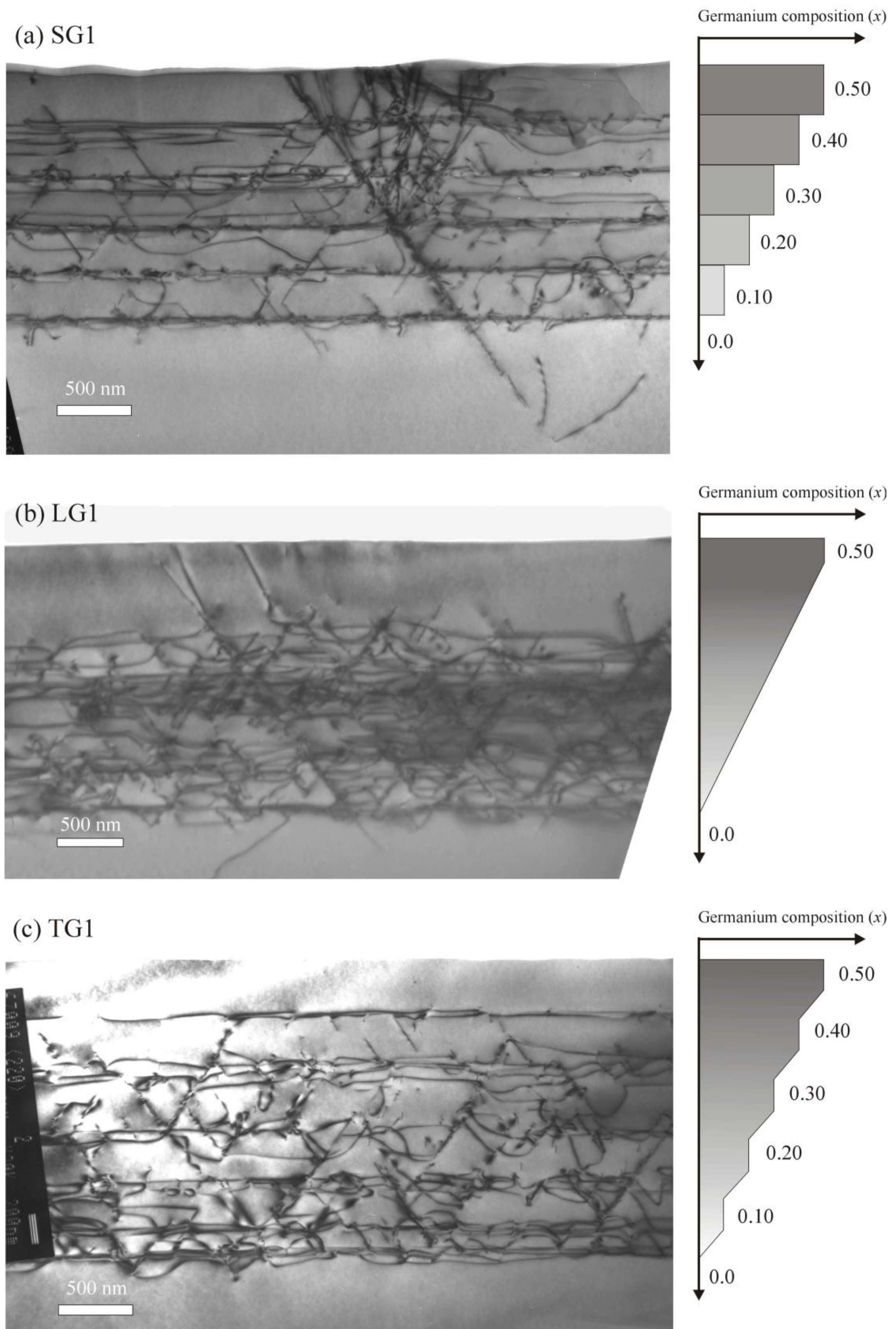


Figure 4.6 Bright field XTEM micrographs taken with ***g*** = (220) with higher magnification of (a) SG1, (b) LG1 and (c) TG1. The terrace graded virtual substrate (c) shows clearly delineated dislocation networks in the graded layers.

of the structure, indicating that dislocations have been pushed down from interactions in the upper layers. In contrast, the XTEM micrographs of the linearly graded virtual substrate, shown in Figure 4.5(b) and Figure 4.6(b), show that the dislocation pile-ups extend the entire width of the graded region. Since many dislocations are nucleated by the same multiplication sites, as clearly suggested by the large dislocation pile-ups, then the dislocations will be pushed into the substrate. Figure 4.5(b) clearly shows this is indeed the case, further strengthening the case that the many dislocations are sourced from only a few multiplication sites. Also, threading dislocations have been pinned at the pile-up and penetrate through the uniform top layer to the surface. These pinned threading dislocations are potentially detrimental to device performance on active layers and need to be reduced and ideally eliminated. This is a clear indication that the terrace graded layer sequence is an improvement on the conventional linearly graded virtual substrates. Figure 4.5(a) shows the XTEM of the step graded wafer. Most of the misfit dislocations are confined to the abrupt interfaces where the germanium composition has a sudden increase, hence there is a large misfit stress. The dislocations that have formed in between the interfaces will most likely have been pushed down by dislocations forming above them. In this sample there is a large number of threading dislocations seen penetrating the upper surface. In particular, very large clusters of threading dislocations have formed with a separation of a few tens of microns. Presumably this is a consequence of the increased number of dislocation interactions due to the dislocations forming on the same atomic planes (the interfaces between the steps). Since so many dislocations are needed at these interfaces to fully relax the layers, the local stress fields associated with these dislocations will tend to pin gliding dislocations as they pass each other orthogonally.

4.2.2 X-ray Diffractometry

In order to check the relaxation and Ge composition of the terrace graded wafer XRD measurements were performed as described in section 3.2 at Imperial College, London. The results of these measurements are shown in Figure 4.7. The intensity of an X-ray curve is proportional to the number of planes with a given atomic spacing. For a uniform layer, where there are a large number of planes with identical plane separation, the X-ray curve will have a pronounced peak. A graded layer has a continually increasing plane separation and hence only a small number of planes with exactly the same separation, this leads to a much lower X-ray intensity with a flat profile. In Figure 4.7 the six peaks are due to the layers of constant composition as these have the largest number of equally spaced atomic planes, and hence the highest intensity in the diffraction pattern. The highest and thinnest peak is due to the silicon substrate. This is because it has the largest number of atomic planes and because of its near perfect crystal quality. The four middle peaks are due to reflections from the $x = 0.10$, 0.20, 0.30 and 0.40 uniform layers counting from the substrate peak. These peaks are considerably lower than the substrate peak due to their limited number of atomic planes. The fact that these peaks are relatively broad suggests that there is a variation in lattice parameter or plane orientation (mosaic broadening) from the reflecting planes.[87,88] Both of these are expected consequences of the dislocations that can be seen to be penetrating through the uniform layers in the XTEM of Figure 4.5(a). Importantly, the peak due to the final uniform layer is considerably narrower than the other peaks. This is consistent with the lower number of dislocations seen in the XTEM and shows that the final layer is of a high quality. The key results obtained using the analysis of section 3.2 are summarised in Table 4.1.

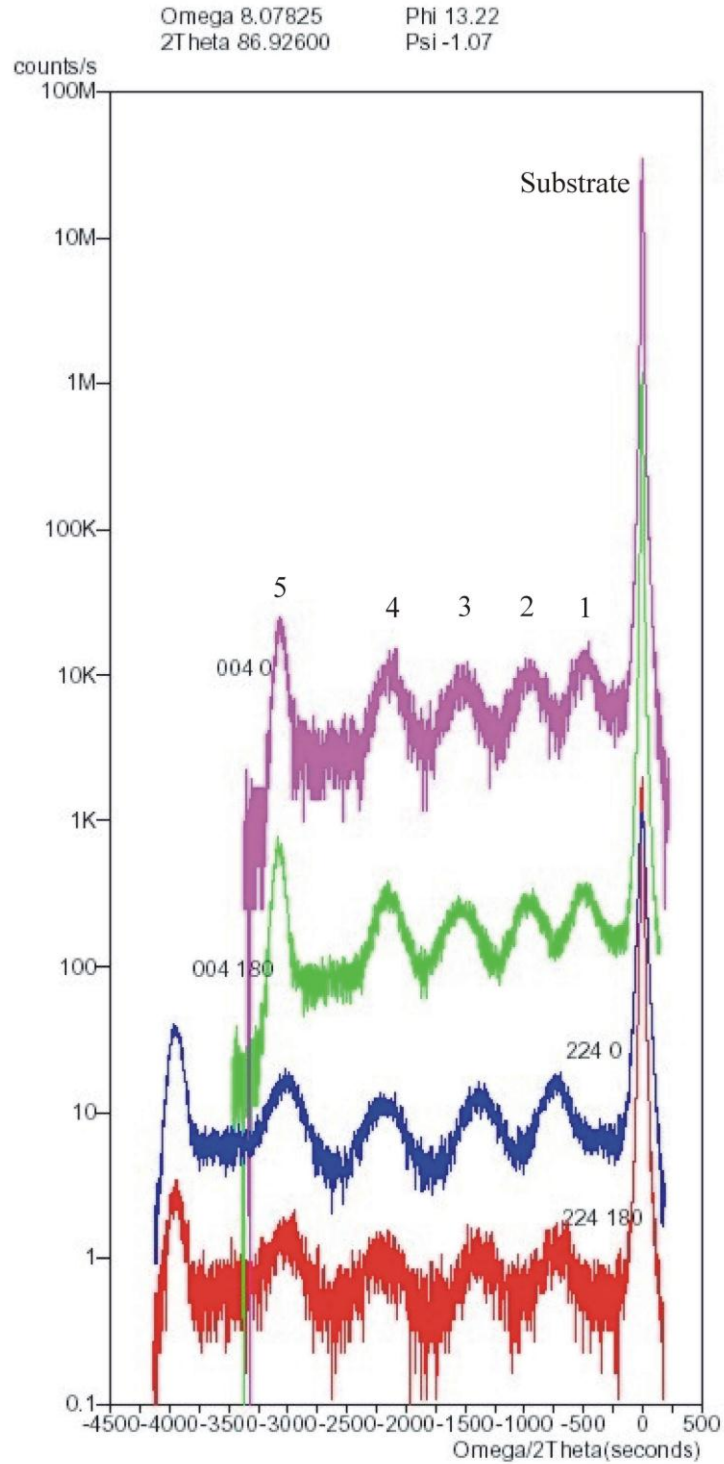


Figure 4.7 XRD rocking curve of the (004) and (224) diffraction spots of sample TG1.

	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5
Composition, x	0.094	0.184	0.289	0.397	0.542
Relaxation (%)	101.0	102.1	101.5	99.9	95.2

Table 4.1 Summary of result derived from the XRD rocking curves of Figure 4.7.

From the XRD data it can be seen that the actual compositions of the uniform layers are very close to the intended values of $x = 0.10$, 0.20, 0.30, 0.40 and 0.50, with the exception of the final layer which has a composition of $x = 0.54$, higher than was intended. The layers are clearly all in a very high state of relaxation. The fact that some of the lower layers are more than 100% relaxed is not due to the inaccuracy of the technique (which is better than 1%), but is thought to be real. The relaxation is a measure of the layers in-plane lattice parameter compared to that of a relaxed layer of the same composition. Hence, relaxation greater than 100% indicates that the layer has a lattice parameter larger than that of a relaxed layer, i.e. it is *tensile* strained. Since the higher layers are all fully relaxed it is possible that they have put a small tensile strain on the lower layers. Also, if dislocations have been forced into these lower layers from above, from the MFR mechanism for example, then there will be an over-saturation of dislocations in these layers. Since misfit dislocations relax the layers they are contained in, an over-saturation of these dislocations will result in over-relaxation.[67] Sample LG1 was also subject to XRD analysis and the final composition was found to be $x = 0.50$ with a relaxation of 96.1%. This is as was intended, but a lower value of x than sample TG1. The XRD of sample SG1 is shown in Table 4.2, again the uniform layers have been grown close to the desired specification with each layer in a high state of relaxation. The final layer has a higher germanium composition than was intended, but this is almost identical to sample TG1 making this a fair comparison.

	Layer 1	Layer 2	Layer 3	Layer 4	Layer 5
Composition, x	0.12	0.211	0.318	0.407	0.539
Relaxation (%)	102.1	100.7	100.0	100.3	96.1

Table 4.2 Summary of the XRD results of sample SG1.

4.2.3 Grading Profile Determination

With the composition of the uniform layers ascertained from XRD it is necessary to check that the actual composition *profiles* are the same as the original growth specifications. The ideal technique for this type of depth/composition analysis is Secondary Ion Mass Spectrometry (SIMS). SIMS was carried out at the SIMS group of the University of Warwick. The incident ion beam was directed normal to the surface of the sample in order to keep the erosion of the sample as uniform as possible. The bombarding ions were O_2^+ and were accelerated through 1 KeV. The results of the analysis are summarised in Figure 4.8, the results for the uniform wafers and the terraced wafers are plotted together for comparison.

Referring to Figure 4.8 we can see the expected terraced profile for the terrace graded wafer (blue). The first two terraces with the highest Ge composition

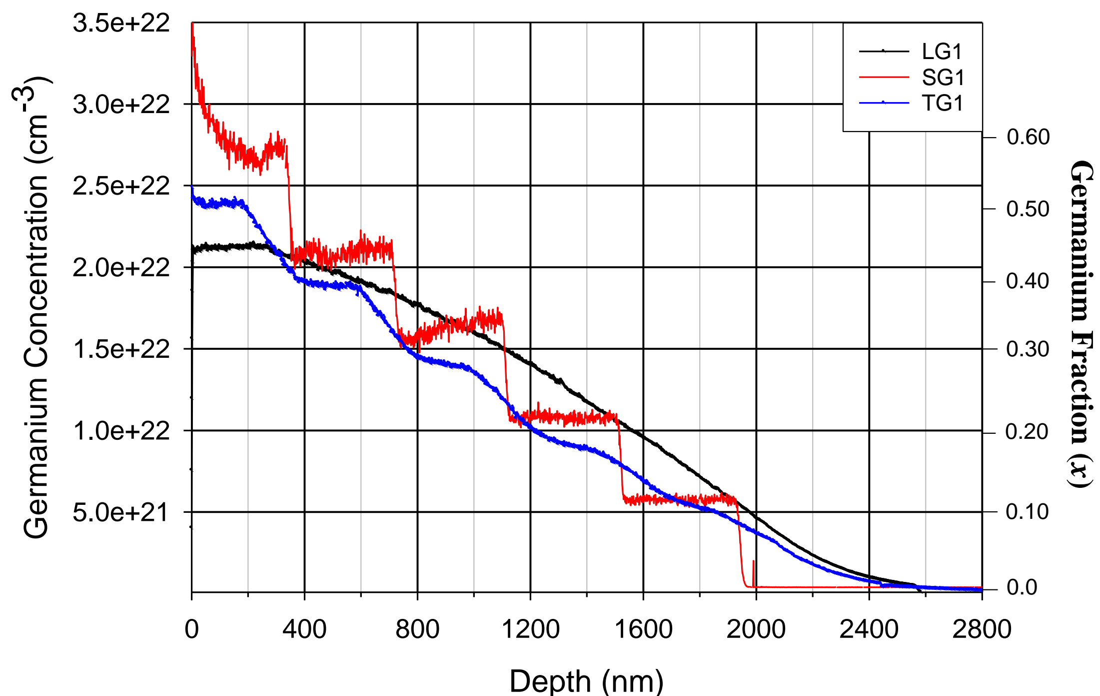
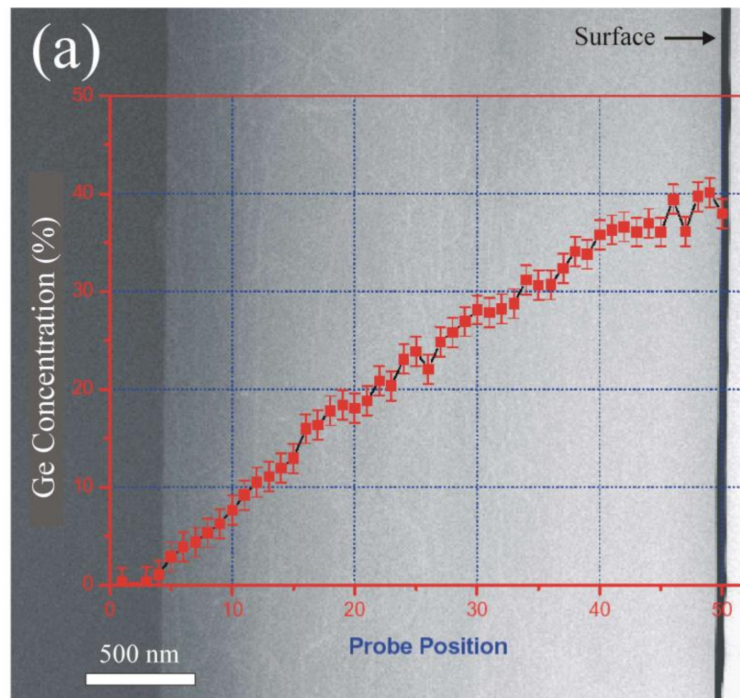


Figure 4.8 SIMS profiles of sample LG1 SG1 and TG1.

are clearly seen to be of constant composition with linear grading in between. As the profile extends deeper into the sample the uniform layers appear to be smoothed out, making the profile approach that of a linearly graded layer. It is believed that this smearing of the profile is an artefact of the SIMS technique rather than an actual feature of the sample. It is proposed that the erosion rates of the sample are influenced by the high density of dislocations which are clearly seen in the XTEM in Figure 4.5(c). As can be seen, the density of the dislocations in the lower layers is much higher than the upper regions. Another factor is that the uniformity of the eroded surface is reduced with increasing depth into the sample. The consequence of these factors is that the SIMS data cannot be accurately quantified for the deeper layers. Instead, another method of quantifying the Ge composition is needed to confirm that the requested profile has been achieved. Energy dispersive spectroscopy (EDS) was carried out in the HRTEM facility at the University of Sheffield. This uses the X-ray spectra from the sample in the TEM to determine the composition of the layers. The result of the EDS analysis is shown in Figure 4.9, as can be seen this Ge profile is not exactly the same as the SIMS profile. In the lower layers the expected terraces at $x = 0.10$, 0.20 and 0.30 can be clearly seen. However, in the higher regions the profile fluctuates, although small terraces at $x = 0.40$ and 0.45 can just be made out.

Considering the data from the SIMS and the EDS together it can be seen that the desired Ge profile has been achieved. The SIMS is accurate for the upper regions of the layers and show that the actual profile is close to the desired profile. In the lower regions where the SIMS is known to be inaccurate the EDS show the profile still has the desired uniform layers. Although neither technique shows the full profile as was intended, they both show the uniform terraces and linear grades in between.



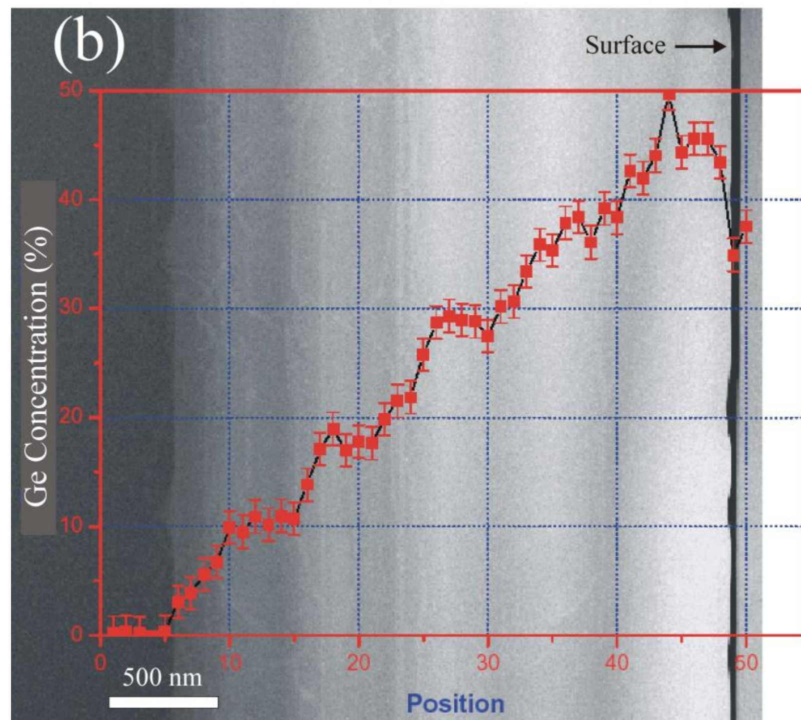


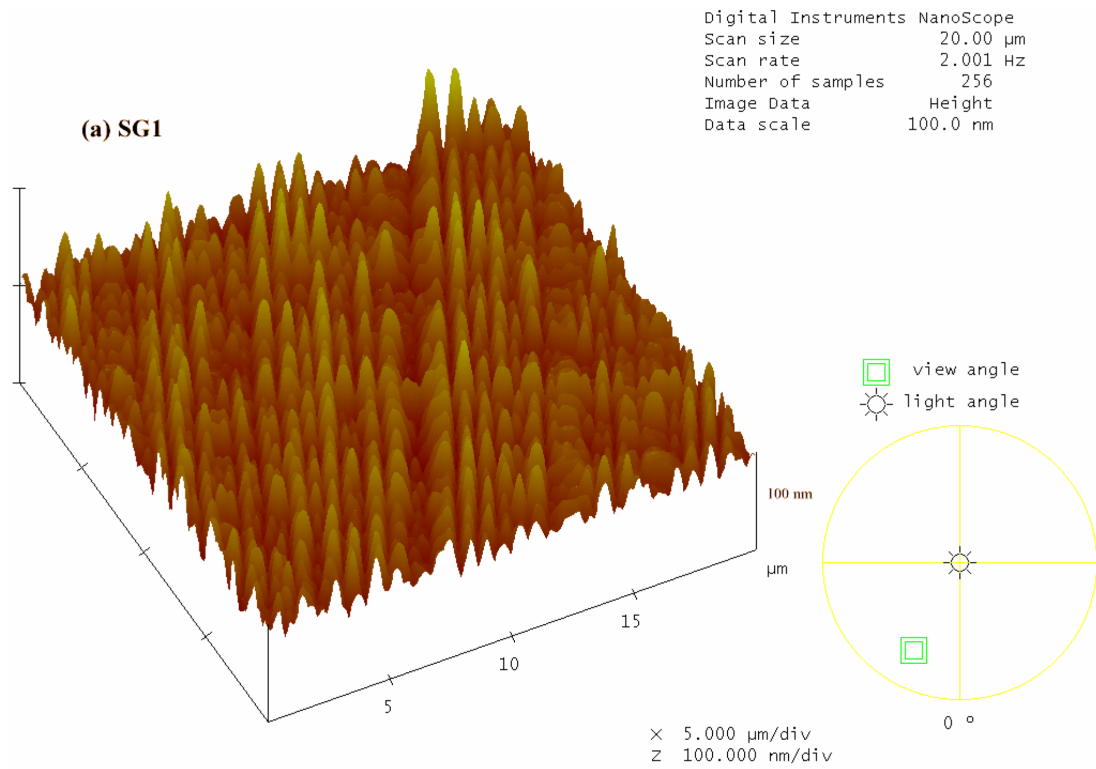
Figure 4.9 EDS profile of sample LG1 (a) and sample TG1 (b).

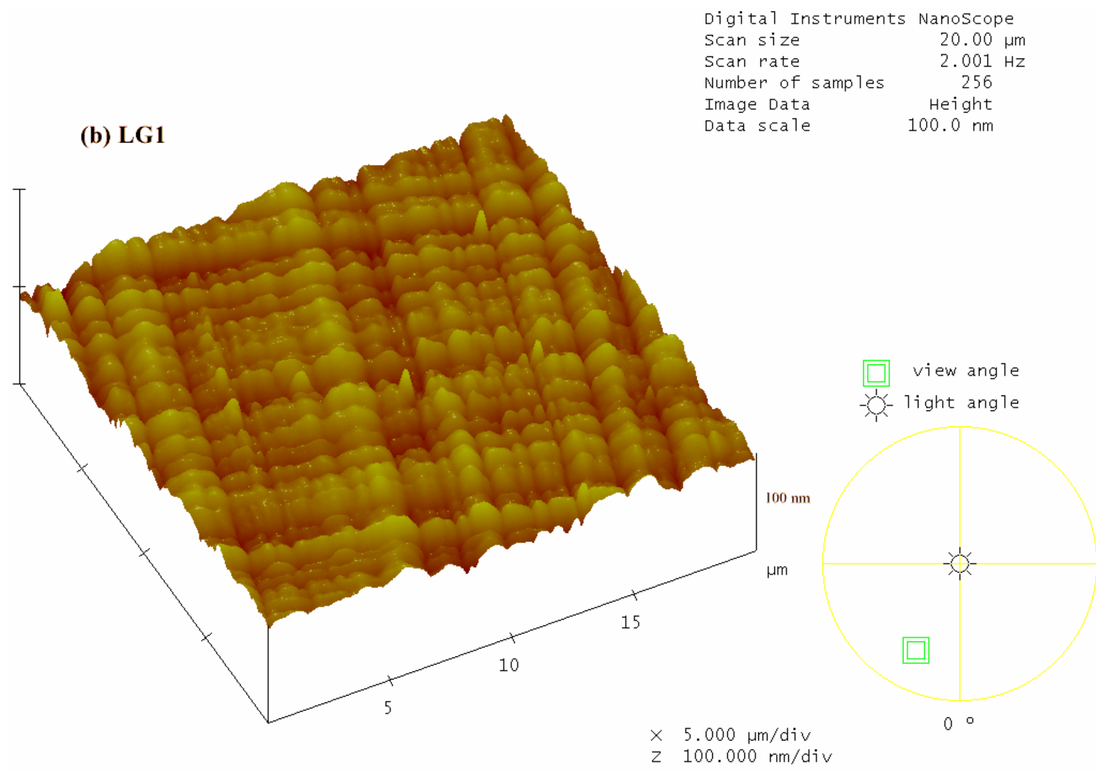
The EDS confirms the lower regions and the SIMS the higher regions. In addition, the X-ray results show equally spaced peaks of the same height, indicating that the uniform layers are present and of similar thickness, and the XTEM shows the delineation of the graded and uniform layers. Since it is unlikely that any inaccuracies in either technique would artificially produce by chance uniform regions in exactly the places that were intended, it can fairly be assumed that the entire composition profile has been successfully achieved.

The SIMS for the linearly graded wafer in Figure 4.8 shows that the Ge profile is close to the desired linear grade, although it must be noted that the profile has a slightly convex shape. This convex shape might be due to the uneven erosion rates as discussed above. One feature that stands out is that final Ge composition is noticeably lower than the terraced wafer in agreement with XRD, and is given by this analysis as $x = 0.45$ instead of the intended 0.50. The EDS profile shown in Figure 4.9 shows a more linear grade of the Ge composition, but again the final composition is lower than was intended ($x = 0.40$). Clearly, the values of x measured using these techniques do not agree, but the XRD is known to be the most accurate for absolute compositional determination. Since the linearly graded wafer was grown as a benchmark against which the terraced wafer could be compared, the fact that it is clearly lower in Ge composition is not a problem. The lower composition of the linearly graded wafer should make the virtual substrate of higher quality, since the lower grading rate will decrease the chances of dislocation interactions, which cause pinning of the threading dislocations. Hence, any improvement of the terraced wafer compared to this linearly graded wafer would imply an even greater improvement against a linear grade to $x = 0.50$. The SIMS of sample SG1 shows that the profile has abrupt steps as was intended, but measures the composition higher than XRD.

4.2.4 AFM

Atomic Force Microscopy (AFM) was carried out on the three samples LG1, TG1 and SG1 using a Digital Instruments Nanoscope III AFM in order to compare the surface morphology of these three approaches to virtual substrate growth. The 3-D profiles of the three samples are shown in Figure 4.10 with a typical line profile of each sample shown in Figure 4.11. The step graded sample (SG1) shows an extremely rough surface in the 3-D plot, with large peaks aligned along the <110> directions. The line profile in Figure 4.11(a) shows the scale of these peaks that have an RMS roughness of 14.4 nm taken over an area of 20 μm × 20 μm. The expected crosshatch pattern usually associated with virtual substrates relaxed by the MFR mechanism can be seen in sample LG1. The line profile in Figure 4.11(b) shows the regular array of troughs and peaks with a periodicity of approximately 1 μm. The RMS roughness of the 20 μm × 20 μm scan of LG1 is 6.9 nm. This is a reasonable value for a $Si_{0.50}Ge_{0.50}$ virtual substrate of only 2 μm thickness. The well spaced troughs and peaks of this sample is due to the many dislocations that have been formed on the same atomic glide planes as seen in Figure 4.5(b). In section 2.5.4 it was explained how these dislocations lead to large undulations on the surface of the virtual substrate. The AFM line profiles of sample TG1 shown in Figure 4.11(c) shows a clear contrast to the step graded sample SG1 in (a) and the linearly graded sample LG1 in (b). There are still troughs and peaks associated with the MFR mechanism, but the amplitude is now much reduced. This is due to the dislocation networks now being more evenly distributed since the uniform layers have separated the dislocations from the different multiplication sites. The surface morphology is an average of the crosshatch pattern expected for each of the five graded layers and is consequently much smaller than in sample LG1 where no such separation occurs.


Digital Instruments NanoScope
Scan size 20.00 µm
Scan rate 2.001 Hz
Number of samples 256
Image Data Height
Data scale 100.0 nm
(a) SG1
view angle
light angle
100 nm
µm
5
10
15
0 °
X 5.000 µm/div
Z 100.000 nm/div
70013.000


Digital Instruments NanoScope
Scan size 20.00 µm
Scan rate 2.001 Hz
Number of samples 256
Image Data Height
Data scale 100.0 nm
(b) LG1
view angle
light angle
100 nm
µm
5
10
15
0 °
X 5.000 µm/div
Z 100.000 nm/div
67045-20um.001

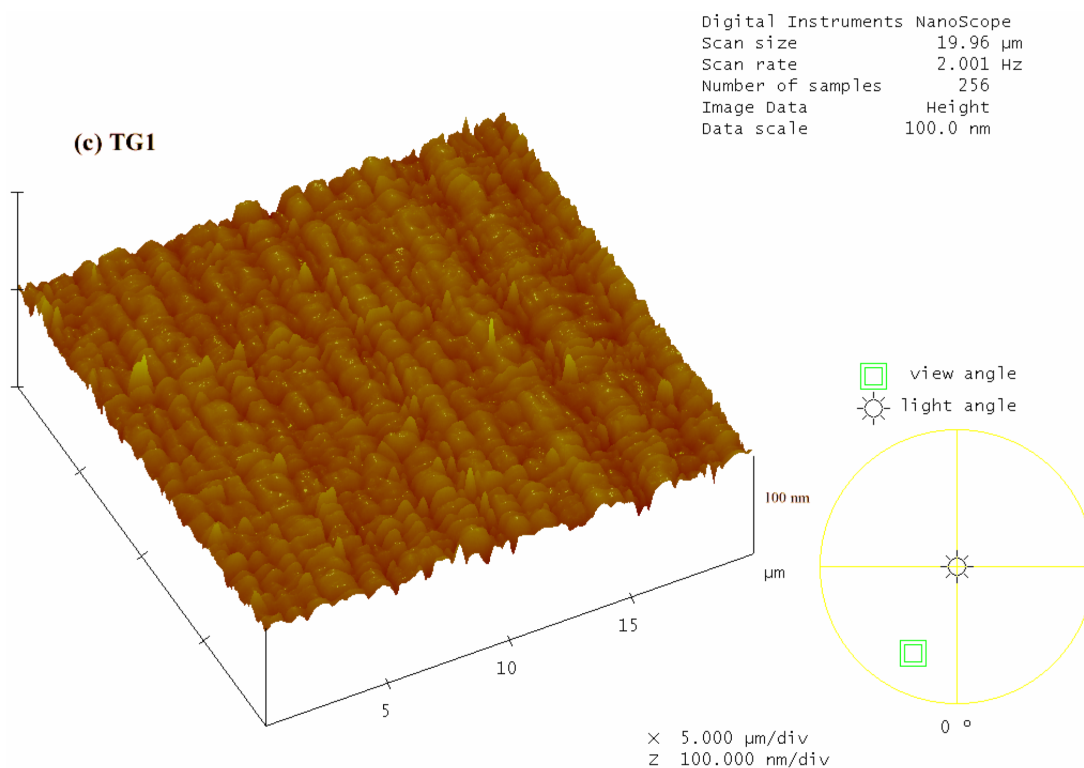


Figure 4.10 3-D AFM surface plots of sample SG1 (a), LG1 (b) and TG1 (c) all to the same scale.

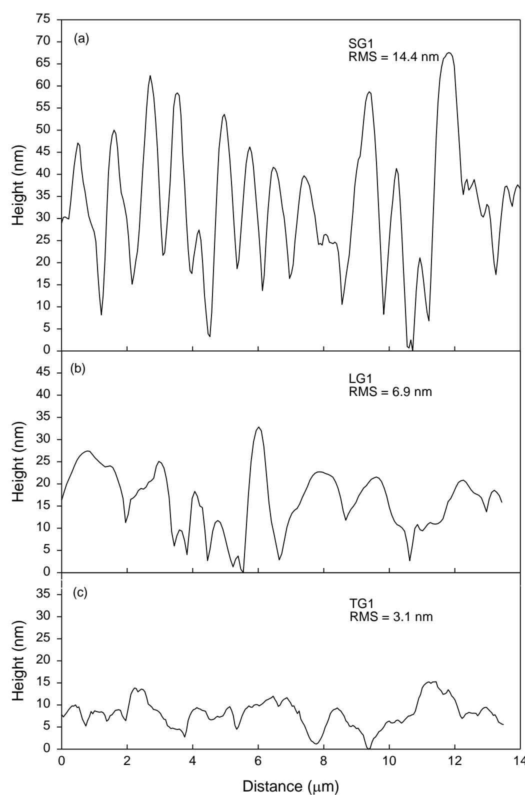


Figure 4.11 AFM line profiles taken from a 10µm *x* 10 µm scan of sample SG1 (a), LG1 (b) and TG1 (c) drawn to the same scale.

The RMS roughness for this sample taken from a 20 μm × 20 μm scan is 3.1 nm, less than half that for sample LG1 and less than a quarter of sample SG1. It is clear that the effect of the terrace grading is to smooth the surface compared to the samples without this Ge profile.

4.2.5 Threading Dislocation Densities

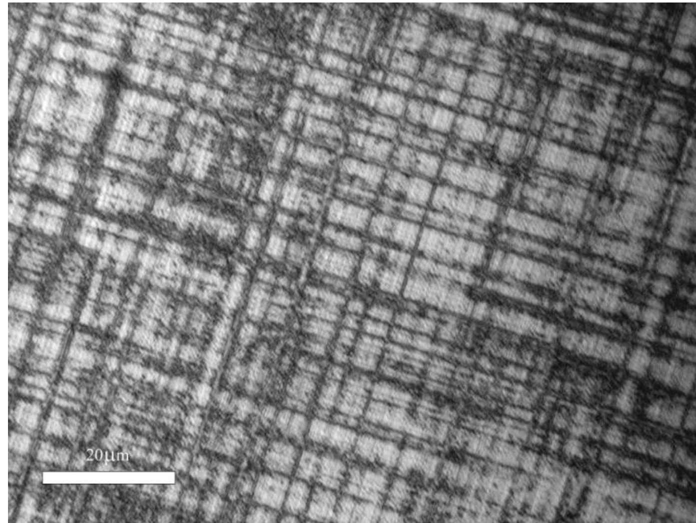
In order to determine the threading dislocation densities of each of the virtual substrates the samples were subjected to a defect etch. In this study a Schimmel etch was used which comprises a mixture of CrO_3(0.75M) : HF(50%) : H_2O in the ratio 2 : 4 : 3. The mixture is an anisotropic etch of SiGe which acts by the oxidation of SiGe by CrO_3 then subsequent removal of the oxide by the HF. Where a threading dislocation penetrates the surface of the sample, the local strain fields causes the etch rate to be enhanced. Consequently, the location of these threading dislocations can be seen in an optical microscope as small pits on the surface of the sample.

Figure 4.12 shows digital images of the etched samples taken using a Nikon DN100 camera attached to a Reichart-Jung microscope. The images were taken in the bright field (without DIC) at a magnification of 50×. The scale bars are 20 μm, giving a total field of view of approximately 100 μm × 80 μm. The etch pits of sample SG1 shows many closely spaced dark bands which is due mainly to the large amplitude of the surface roughness as seen in the AFM of Figure 4.12(a). Since the surface is very uneven the etch rate has also become uneven across the sample, further exaggerating this roughness. Closer inspection of the dark bands reveal that there are many small pits clustered closely together, this shows that the threading dislocations have been pinned and piled-up together at these regularly spaced bands. This is due to the fact that the dislocations can only form on the interface between

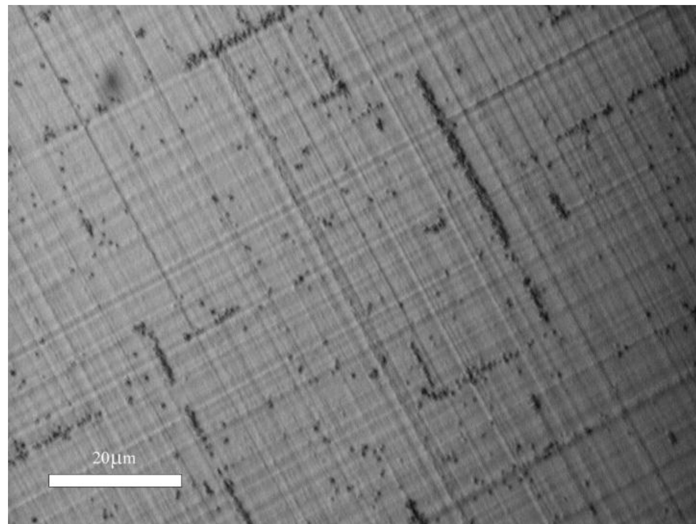
the steps in the germanium profile, so that there is a high probability of these dislocations interacting and being pinned.[64,89] This is further exaggerated by the deep surface troughs which act to prevent the glide of dislocations which cross them.[46] Because of the high density of these etch pits it is impossible to determine the value of the threading dislocation density, although this density is clearly very high making the virtual substrate of poor quality.

The etch pits of sample LG1 is shown in Figure 4.12(b) and the reduced number of threading dislocations is clearly shown. The threading dislocations are much further spaced on the surface of the sample with only a few areas where the dislocations have clustered together in pile-ups. The density of the dislocations that are not part of a pile-up (sometimes known as field dislocations) is of the order $\sim 4\times10^6$ cm^{-2}. However, at the pile-ups the density is too large to measure so that the *total* threading dislocation density is much larger than this and at least $\sim 10^7$ cm^{-2}. The effect of grading the germanium composition has clearly lowered the density of the field dislocations and reduced the occurrences of pile-ups, this is expected since the misfit dislocations are no longer confined to discrete planes but may form anywhere within the graded region. This drastically lowers the probability of gliding dislocations interacting with pre-existing orthogonal dislocation which may pin the gliding threading arm. The threading dislocations have piled-up in some areas due to the formation of many misfit dislocations on the same glide planes as seen in the MFR relaxation mechanism, these act as barriers to dislocation glide since misfit dislocations are stacked on top of each other at these sites leaving no room for orthogonal dislocations to pass. It was the aim of the terrace graded profile to prevent many dislocations forming on top of each other in this way, and so reduce the number of dislocation pile-ups.

(a) SG1



(b) LG1



(c) TG1

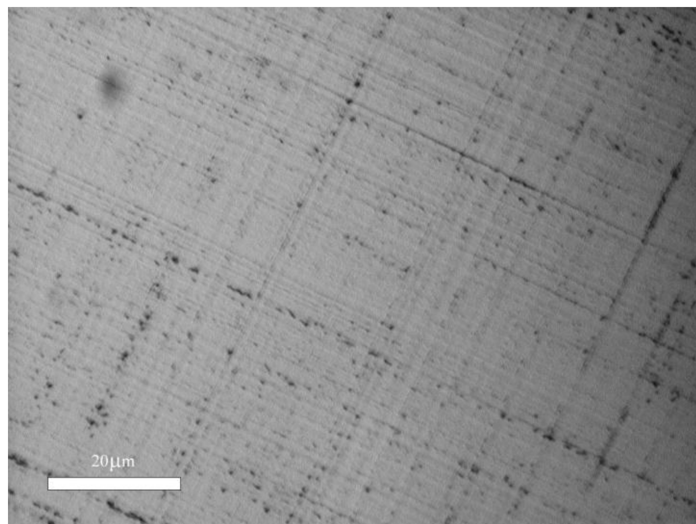


Figure 4.12 Optical micrograph of defect etched samples (a) SG1, (b) LG1 and (c) TG1. The micrographs were taken in the bright field without polarisers. The scale bars are 20 μm.

The etch pits of sample TG1 in Figure 4.12(c) shows that not only have the field dislocations been slightly reduced to ~ 3×10^6 cm^{-2}, but more importantly there are very few areas of dislocation pile-ups. Where dislocations have piled-up the density of dislocation is still much lower than both sample SG1 and LG1 and the *total* threading dislocation density is ~ 4×10^6 cm^{-2}. This is the effect that was hoped for by using this terrace grading technique. Since at each uniform layer the blockade of dislocations from the MFR mechanism is interrupted, threading dislocations are able to pass the orthogonal dislocations at the next graded layer. This frees the dislocations to glide leaving only the threading dislocations that were pinned in the final $x = 0.40 - 0.50$ graded layer still in pile-ups. A further cause for the reduction of the threading dislocation density is the lower amplitude of the surface undulations as shown in Figure 4.11. As explained in section 2.5.4, the surface undulations can act to reduce the available height for dislocations to glide over pre-existing dislocation networks. This can cause the dislocations to get pinned in by dislocations that they would be able to glide over if the undulations were not there. It is clear that the use of terrace grading has considerably improved the density of threading dislocations at the surface of the wafer in a way that was proposed in section 4.1.1.

4.2.6 Discussion

The novel grading profile described in this chapter has shown noticeable improvements over more conventional linearly graded and step graded virtual substrates. By allowing dislocation multiplication by the efficient MFR mechanism, but constraining these to produce only small dislocation pile ups within each linearly graded layer, the quality of the virtual substrate has been improved. These pile-ups are known to produce troughs at the surface of the wafer whose amplitude is related

to the number of dislocations in the pile-ups. By preventing the pile-ups becoming too large, the amplitude of this crosshatch has been reduced. XTEM has clearly shown that the dislocation networks of the terrace graded wafer has no large scale pile-ups, only well separated smaller pile-ups were observed as was hoped. In contrast, the linearly graded virtual substrate has the more familiar large scale pile-up that runs through the entire thickness of the epitaxial layers. AFM has shown that the surface undulations caused by the dislocation network beneath the surface has been reduced in amplitude and become more narrowly spaced. This is important for the processing of devices grown on virtual substrates. Defect etching has shown that the density of threading dislocations has been greatly reduced, in particular there are very few dislocation pile-ups since the uniform layers have separated the dislocation networks caused by the MFR mechanism which act as a barrier for dislocation glide. The lower layers of the terrace graded virtual substrate show signs they have relaxed later than intended in the growth, this may be limiting the quality of the virtual substrate and is to be addressed.

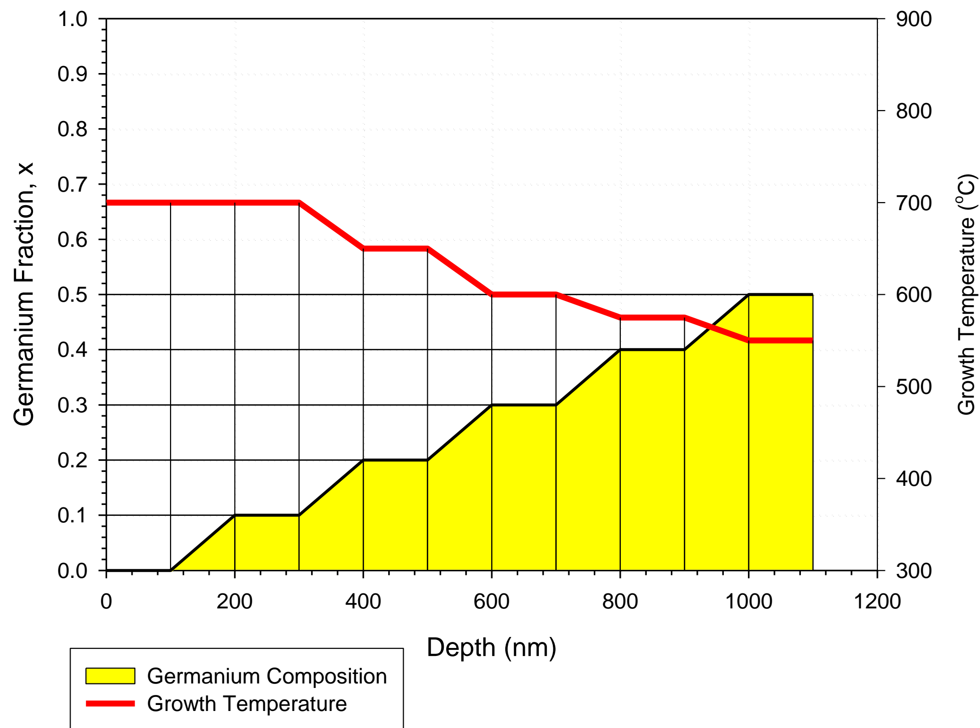
4.3 Thinner Terrace Graded Virtual Substrate

The terrace graded profile of the virtual substrate has been shown to be superior to a conventional linearly graded profile and a step graded profile with the same thickness and grown at the same temperature. The total thickness of the terrace graded virtual substrate was 2 μm. This is thin by comparison to most $x = 0.50$ virtual substrates described in the literature. Fitzgerald *et al*[46] suggest that a grading rate of no more than $x = 0.10$ per μm is needed to produce a high quality virtual substrate. For a $Si_{0.50}Ge_{0.50}$ virtual substrate with capping layer, at least 7 μm is needed. However, in order to ascertain whether the terraced graded virtual substrate

could be grown even thinner, a wafer (**TG2**) was grown with the same specifications as TG1 except that each graded layer and each uniform layer were grown to a thickness of only 100 nm, making the total thickness 1 μm. The temperature was reduced during growth in a similar manner to TG1. The growth specifications are summarized in Figure 4.13 overleaf.

4.3.1 Structural Characterisation

Figure 4.14 shows an XTEM micrograph of this structure. It is evident that at this much reduced thickness, the uniform layers have not separated the dislocation pile-ups associated with the MFR mechanism. Several pile-ups are seen that extend through the entire thickness of the epilayers, terminating in clusters of threading dislocations at the surface. This suggests that at this reduced thickness the uniform layer is not thick enough to prevent the multiplication mechanism sites from relaxing higher layers. This is likely to be due to the uniform layers being still strained as they are grown. This would allow dislocations to glide through these layers enabling the pile-ups to continue to extend into the next graded layer where the strain is again increased. According to Fitzgerald,[46] as a graded layer grows, the relaxation should be able to keep up with the continued increase in strain. This involves a strained layer of approximately the critical thickness under the surface below which dislocations are continually gliding to maintain relaxation. For a grading rate of 10% per μm this thickness is approximately 300 nm, although at higher grading rates this will be lower. In the case of the terrace graded wafers, it is important that the uniform layer is thicker than this critical thickness, so that the lower graded layer has a chance to completely relax before the growth of the next graded layer. Evidently, this has not happened in this case.



Step	Growth Temperature (^{o}C)	Germanium Composition (%)	Thickness (nm)
0	700	0	100
1	700	0 – 10	100
2	700	10	100
3	700 – 650	10 – 20	100
4	650	20	100
5	650 – 600	20 – 30	100
6	600	30	100
7	600 – 575	30 – 40	100
8	575	40	100
9	575 – 550	40 – 50	100
10	550	50	100

Figure 4.13 Growth specifications of sample TG2.

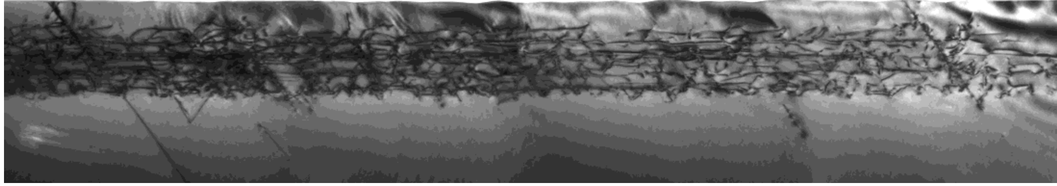


Figure 4.14 Bright field XTEM of sample TG2 with $\boldsymbol{g}$ = (220). Notice the surface roughening with facetted sides.

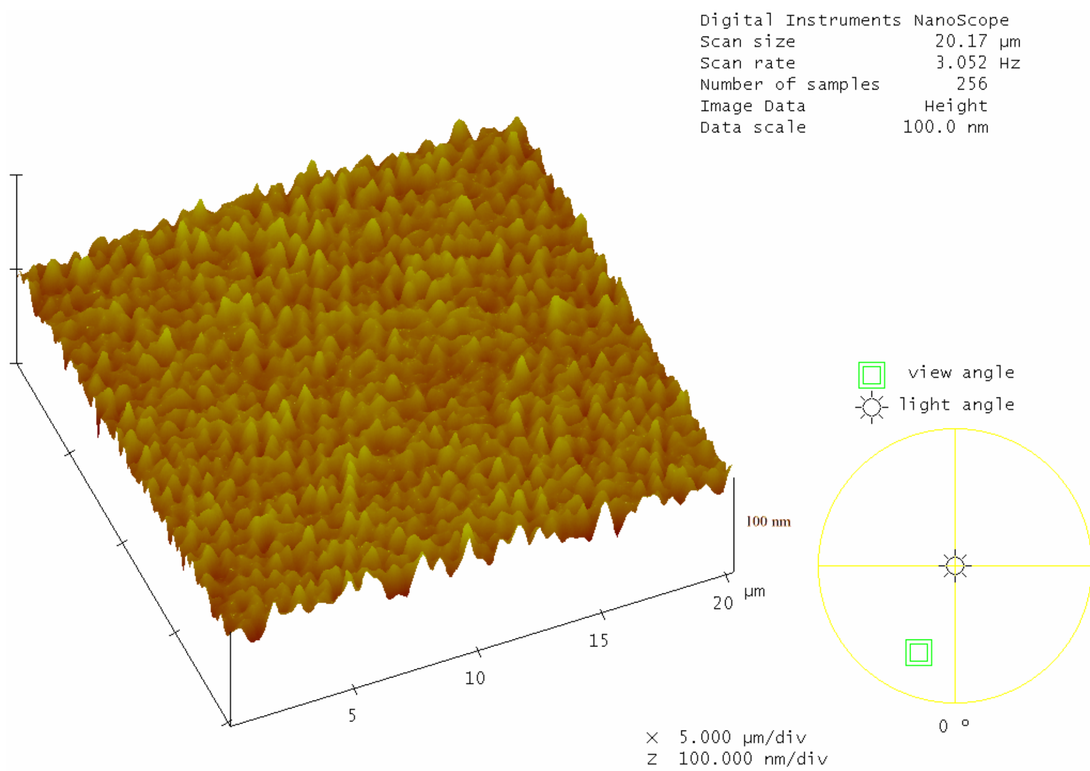


Figure 4.15 3-D AFM surface plot of sample TG2 showing facetted peaks.

It is evident from the XTEM that the surface of this thin terrace graded wafer is inferior to the thicker wafer. 3-D AFM data is shown in Figure 4.15 and confirms the low quality of the surface of this wafer. The surface morphology is dominated by a 3-D ripple with an RMS roughness of approximately 11.7 nm. This ripple is not the same as the crosshatch seen on the 2 μm virtual substrates but has a much smaller wavelength and does not have large troughs which follow the <110> directions. It is likely that at such a smaller thickness, and hence high grading rate, the MFR multiplication mechanism is not the dominant source of strain relaxation. With the decreased thickness of the linearly graded layer, the dislocations needed for plastic relaxation are pushed closer together. This inevitably leads a much higher density of dislocation interactions and consequently there is an increase of dislocation pinning.

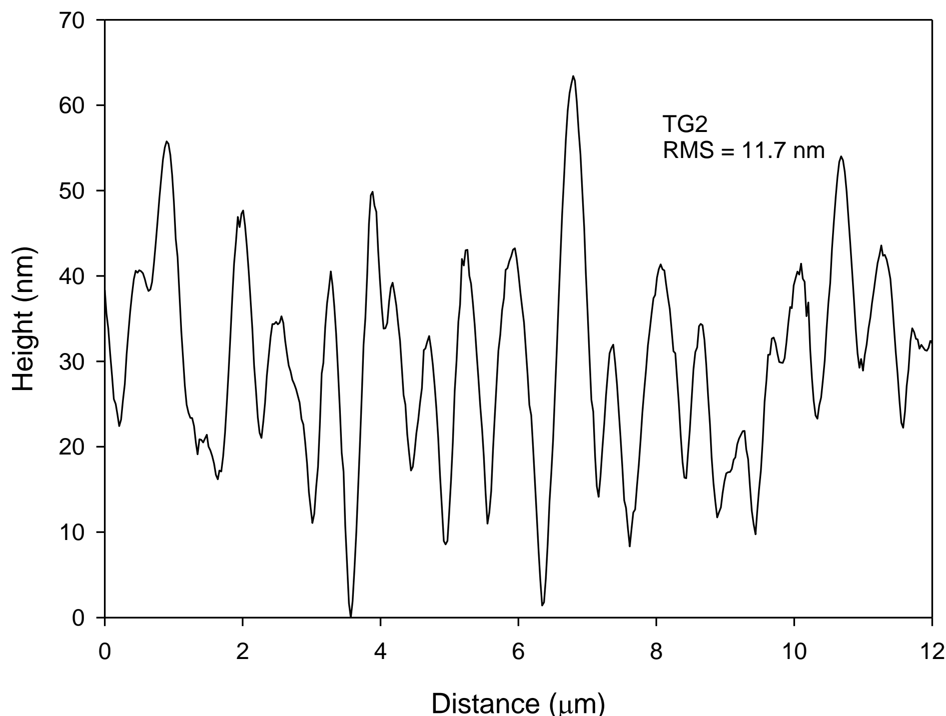


Figure 4.16 AFM line profile of sample TG2.

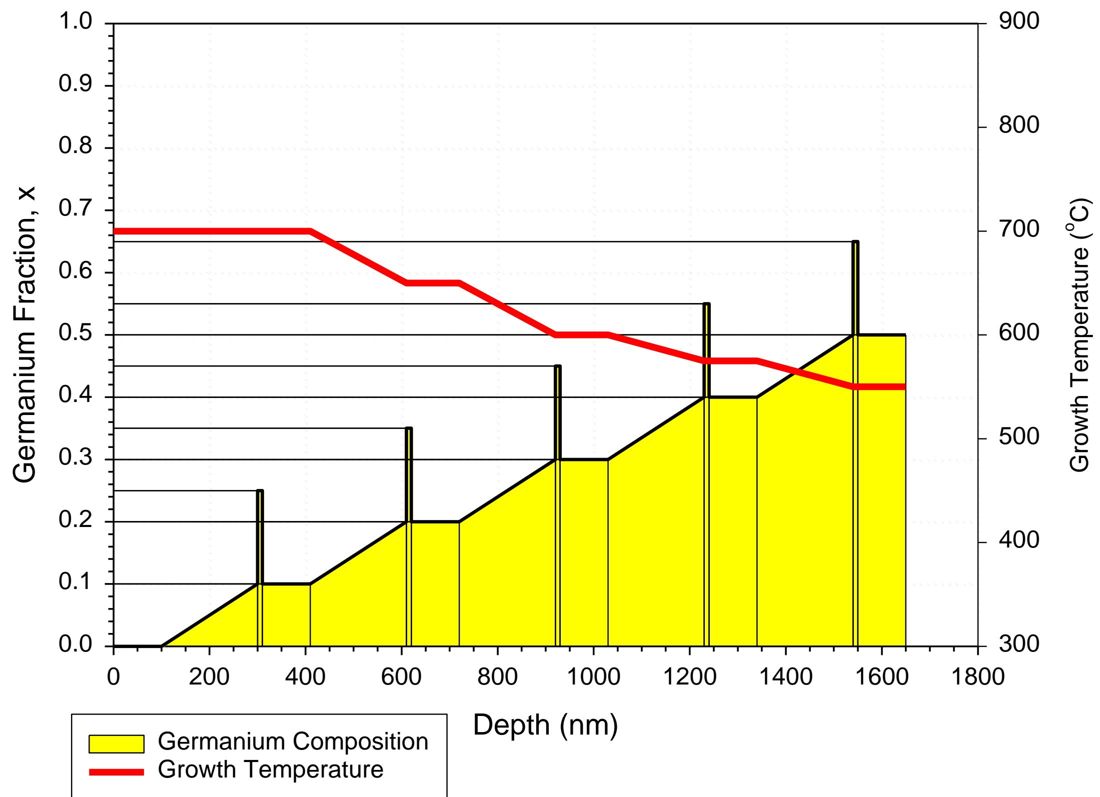
With the dislocations unable to fully relax the graded layers as they grow, the strain may start to relax via a surface roughening mechanism. This would account for the highly facetted nature and short period of the surface undulations. The surface profile shown in Figure 4.16 shows the height of the peaks to be up to 60 nm, over four times larger than for the 2 μm terrace graded buffer (TG1) shown in Figure 4.11(b).

4.3.2 Improving Relaxation using "Stressers"

Since the poor quality of this thinner wafer is believed to be due to the poor state of relaxation as the end of each graded layer is grown, in order to improve this relaxation several ideas were proposed. Firstly, an *in-situ* anneal at the end of each uniform layer should enable dislocation nucleation and glide in order that full relaxation could be achieved. However, since the time needed to increase the temperature for the anneal, followed by about 30 minutes for the anneal and then allowing the temperature to reduce back to the growth temperature would be exceptionally long (although the rapid thermal cycling possible in LPCVD would reduce this considerably), it was decided that any gain in the reduced thickness of the virtual substrate would be offset by the extra time needed for the annealing steps. Consequently, another method for producing thin terrace grade was sought. A second, more novel technique was tried in which at the end of each graded layer a thin layer of higher Ge composition was grown in order to stress the lower layers enough to produce relaxation.

The growth specification for this wafer (**TG3**) was kept similar to the previous terrace graded wafers. The graded regions were kept at 200 nm in order to allow the dislocations to pass each other with minimal interactions but the uniform layers were reduced to 100 nm. In between the graded and uniform layers was grown a 10 nm "stresser" layer at 15% Ge higher than the uniform layers. The higher Ge

content of the stressers should increase the strain exerted on the graded regions and so reduce the need for such a thick uniform layer. The growth specifications are summarized in figure 4.17. The XTEM of this wafer is shown in Figure 4.19(a). It is evident, straightaway, that the stressers have considerably reduced the quality of the virtual substrate. The stressers *have* encouraged relaxation of the lower layers, but rather than forming by a multiplication mechanism, which ultimately reduces the threading dislocations as described in section 2.5.4, the stressers have increased the *nucleation rate*. With a high density of nucleated dislocations, there has been an enormous number of dislocation pinning events. This has caused the threading dislocation density to increase dramatically, an estimate from the XTEM of the threading dislocation density is 10^{12} cm^{-2}. Instead of the stressers increasing the strain energy in the lower layers such that these layers relax, the stressers have exceeded their critical thickness and have relaxed themselves. There are none of the usual signs of the MFR mechanism seen in the XTEM. Since the strain was increased rapidly in these stresser layers, they have been unable to relax by a multiplication mechanism, instead there has been a massive number of nucleation events which have caused the threading dislocation density to increase sharply. These dislocations have quickly become pinned because of their high density. When the subsequent layers have been grown, the high number of threading dislocations have continued to terminate at the growth surface, some have glided so that relaxation is maintained. At each stresser layer more dislocations have been nucleated, leaving the final density extremely high and unsuitable for CMOS device applications.

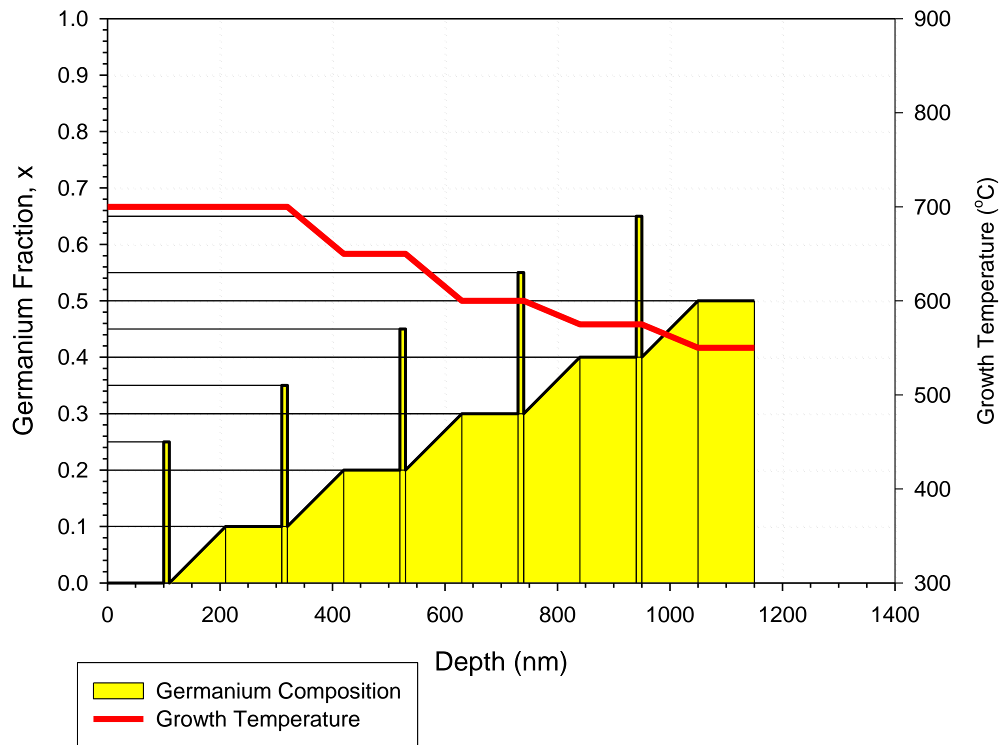


Step	Growth Temperature (°C)	Germanium Composition (%)	Thickness (nm)
0	700	0	100
1	700	0 – 10	200
2	700	25	10
3	700	10	100
4	700 – 650	10 – 20	200
5	650	35	10
6	650	20	100
7	650 – 600	20 – 30	200
8	600	45	10
9	600	30	100
10	600 – 575	30 – 40	200
11	575	55	10
12	575	40	100
13	575 – 550	40 – 50	200
14	550	65	10
10	550	50	100

Figure 4.17 Growth specification for sample TG3.

A second attempt (**TG4**) at using stressers to produce thinner virtual substrates involved the same growth specification as sample TG3, except this time the stressers were grown immediately before the graded layer and the graded layer was reduced to 100 nm, the growth specifications are summarized in Figure 4.18. The idea behind this layer sequence is that the dislocations known to be nucleated by the stresser layers (as determined from the first stresser wafer) could be used to initiate the relaxation. The graded layer could then allow the glide and annihilation of the threading dislocations that have been produced. The uniform layer would then separate the dislocation networks contained within the graded layers. The XTEM of Figure 4.19(b) shows that this approach is equally unsuccessful at producing high quality virtual substrates. Again, the threading dislocations have penetrated through all the layers and terminate at the surface. It would seem that the graded layers do not provide enough opportunity for threading dislocation annihilation and consequently the threading dislocation density remains extremely high. The surface of sample TG4 is extremely rough, much rougher than all the previous samples, including TG3. This is due to the higher grading rate in this sample.

The attempts to produce a $Si_{0.50}Ge_{0.50}$ terrace graded virtual substrate of less than 2 μm thickness have been unsuccessful. The layers of each terraced step are not thick enough to allow full relaxation before subsequent layers are grown. Consequently, the pile-ups of dislocations have not been isolated. The attempts to increase the relaxation of the individual terraced steps growing stresser layers have resulted in a large increase in dislocation nucleation. This in turn has lead to extremely high threading dislocation densities of the order of 10^{12} cm^{-2}. It is possible that lower germanium compositional offsets of the stressers would produce a smaller nucleation rate of dislocations and hence a better quality virtual substrate, or that a



Step	Growth Temperature (°C)	Germanium Composition (%)	Thickness (nm)
0	700	0	100
1	700	25	10
2	700	0 – 10	100
3	700	10	100
4	700	35	10
5	700 – 650	10 – 20	100
6	650	20	100
7	650	45	10
8	650 – 600	20 – 30	100
9	600	30	100
10	600	55	10
11	600 – 575	30 – 40	100
12	575	40	100
13	575	65	10
14	575 – 550	40 – 50	100
15	550	50	100

Figure 4.18 Growth specifications of sample TG4.

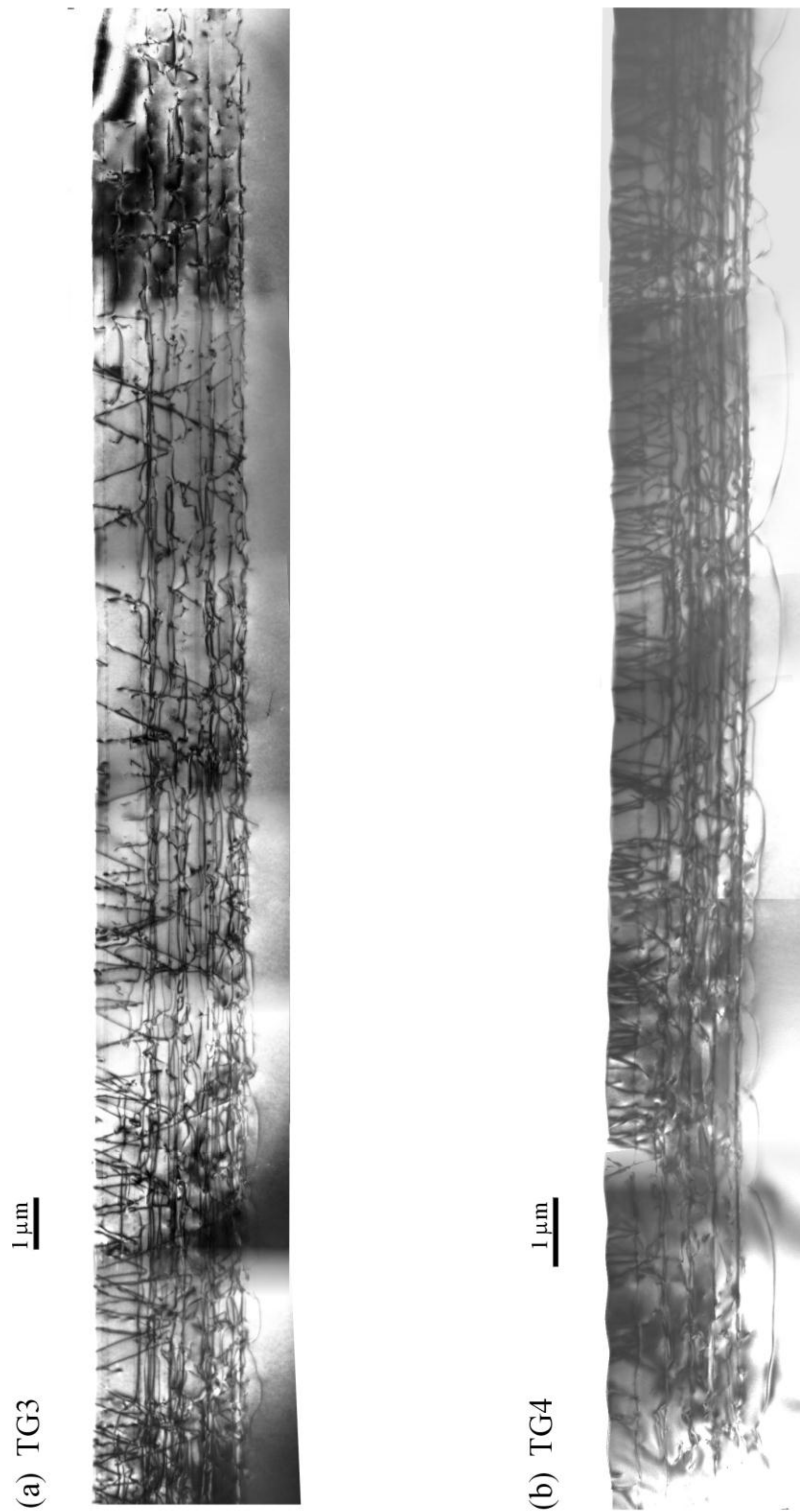


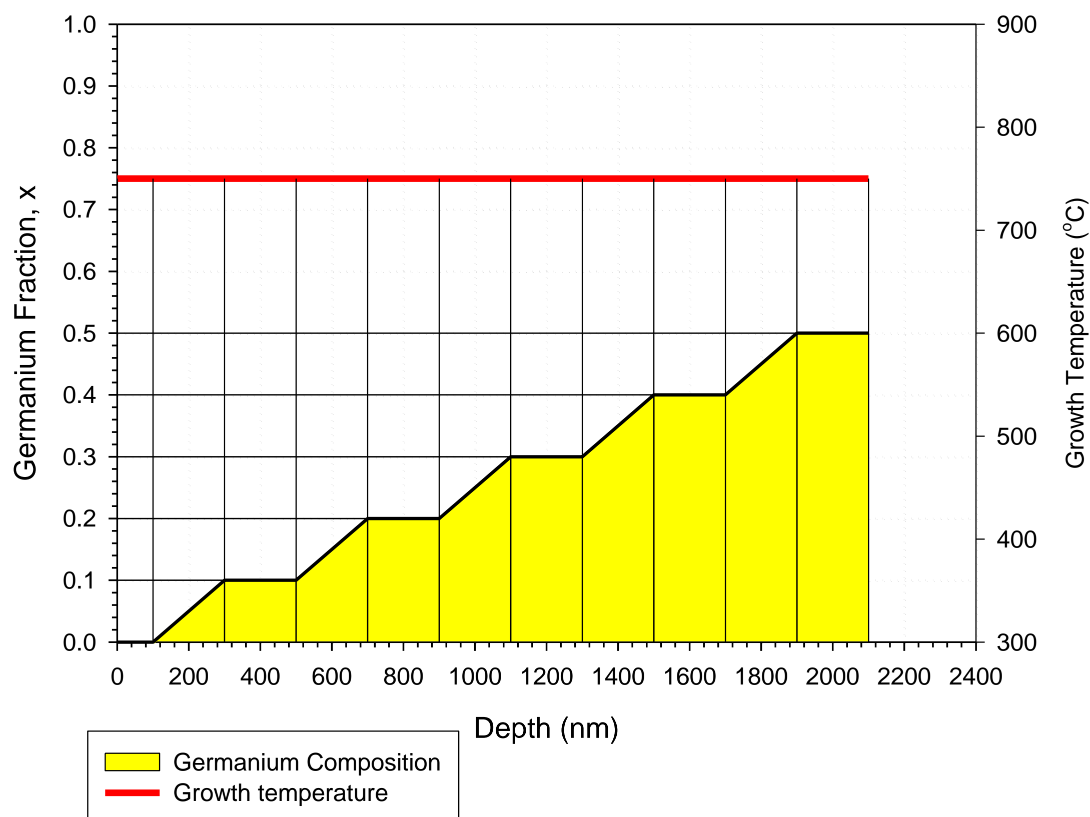
Figure 4.19 Bright field XTEM of sample TG3 (a) and TG4 (b) with ***g*** = (220). Notice the very high density of threading dislocations that penetrate into the final capping layers.

reduced thickness of the stressers such that the critical thickness was not exceeded would produce the required relaxation without further dislocation nucleation. But, since this series of experiments to produce thinner virtual substrates has proven unsuccessful and the wafers utilizing stressers have produced virtual substrates of such poor quality, it was decided to stop any further attempts and concentrate on the more successful results obtained with thicknesses of 2 μm.

4.4 High Temperature Growth

Before any electrical characterisation of channels grown on virtual substrates using this new germanium profile, one more attempt was made to try and improve the quality of the virtual substrate. Since it is believed that the limiting factor of the virtual substrate is the poor relaxation of the lower levels a higher growth temperature was chosen in the hope that this would encourage earlier relaxation of the lower layers. At a higher growth temperature more nucleation sources of the dislocations necessary for the strain relaxation should be active and the glide velocity of these dislocations should be higher. The growth specifications of this sample (**TG5**) are the same as the sample TG1, with the exception that the growth temperature was maintained at 750 °C throughout. The growth specifications are summarized in Figure 4.20.

The XTEM of this sample is shown in Figure 4.21 using the (220) diffraction condition for high dislocation contrast. The XTEM shows that there is a high level of relaxation throughout the thickness of the virtual substrate due to the high density of dislocations. The density of the dislocations is more evenly distributed throughout the graded layers compared with sample TG1, indicating that strain relaxation



Step	Growth Temperature (°C)	Germanium Composition (%)	Thickness (nm)
0	700	0	100
1	700	0 – 10	200
2	700	10	200
3	700 – 650	10 – 20	200
4	650	20	200
5	650 – 600	20 – 30	200
6	600	30	200
7	600 – 575	30 – 40	200
8	575	40	200
9	575 – 550	40 – 50	200
10	550	50	200

Figure 4.20 Growth specifications of sample TG5.

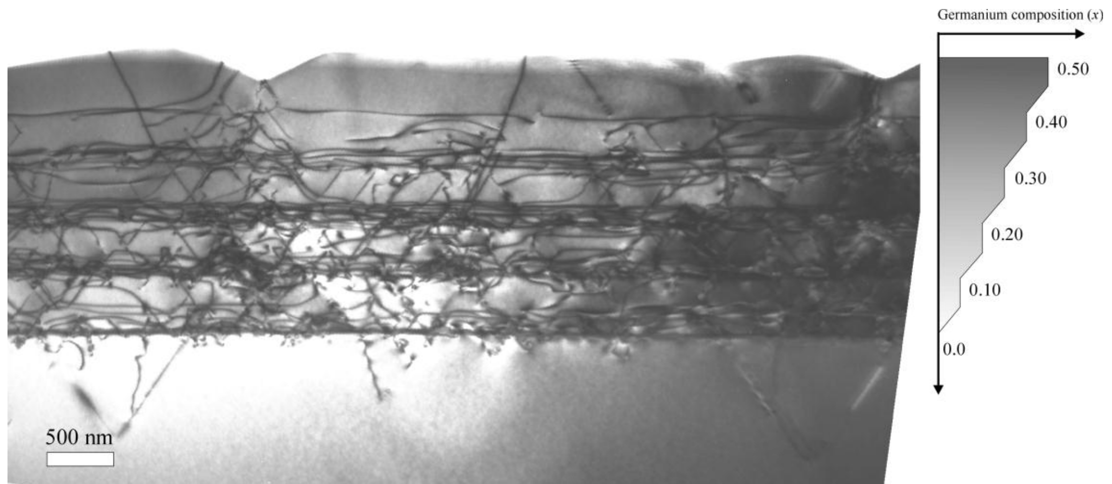


Figure 4.21 Bright field XTEM of sample TG5 with ***g*** = (220).

has probably occurred at a lower level. However, there is less delineation between the graded layers and the uniform germanium composition layer showing that some dislocations have still been pushed down from sources operating in higher layers. The uniform layers are designed to keep the dislocation sources in each of the graded layers separated, but the presence of dislocations within these layers indicates that this has not happened. After the $x = 0.20$ layer there is evidence of dislocation pile-ups that have penetrated each of the layers above, some of which have penetrated the surface as threading dislocations. The surface of the layer is seen to be highly non-planar having deep troughs of ~ 200 nm depth and spaced at around 2 - 3 μm. The sides of these troughs show a facetted nature and there are threading dislocations trapped at the base of most of these troughs. It is likely that the presence of these troughs has locally lowered the activation energy for the nucleation of dislocations, causing preferential nucleation at these sites.[55] Also, the troughs can act as barriers to dislocation glide which would lead to pinning of the dislocations as described in section 2.5.4.

The AFM surface plot and line profile are shown in Figure 4.22 and show the extent of the peaks and troughs seen in the XTEM. The surface has deep troughs orientated along the <110> directions as usually seen in virtual substrates, but in between the troughs can be seen steep facetted peaks indicating the onset of 3-D growth. This is usually due to the growth temperature being too high, allowing the surface atoms to move to relieve strain locally by forming low energy facets. It is for this reason the growth temperature was reduced during growth in sample TG1. The RMS roughness of the surface is 36 nm, over ten times greater than the sample with a graded growth temperature profile.

It is clear that the higher growth temperature has not improved the quality of the virtual substrate. There are threading dislocations seen in the final capping layer and the surface roughness has increased by over a factor of ten. The surface roughening could be reduced by grading the growth temperature in the same manner as sample TG1, but since the $x = 0.10$ layer has still got threading dislocations in the uniform composition it is clear that the higher growth temperature in this sample has not improved the relaxation of these lower layers.

It has come to the attention of the author that during the course of this study another group (Mizushima *et al* [90]) have studied a similar terrace graded structure grown using LPCVD to a final composition of $x = 0.30$ over 0.75 μm (although the SIMS analysis suggests that this is closer to 2 μm). The results presented are minimal, so a full comparison between their approach and the one presented here cannot be made. However, they did find that this terrace grading was an improvement over conventional linear grading, and made a study of the optimum number of steps (linear grade plus uniform layer), keeping the total thickness of the virtual substrate constant. The results of their study is summarised in Figure 4.23

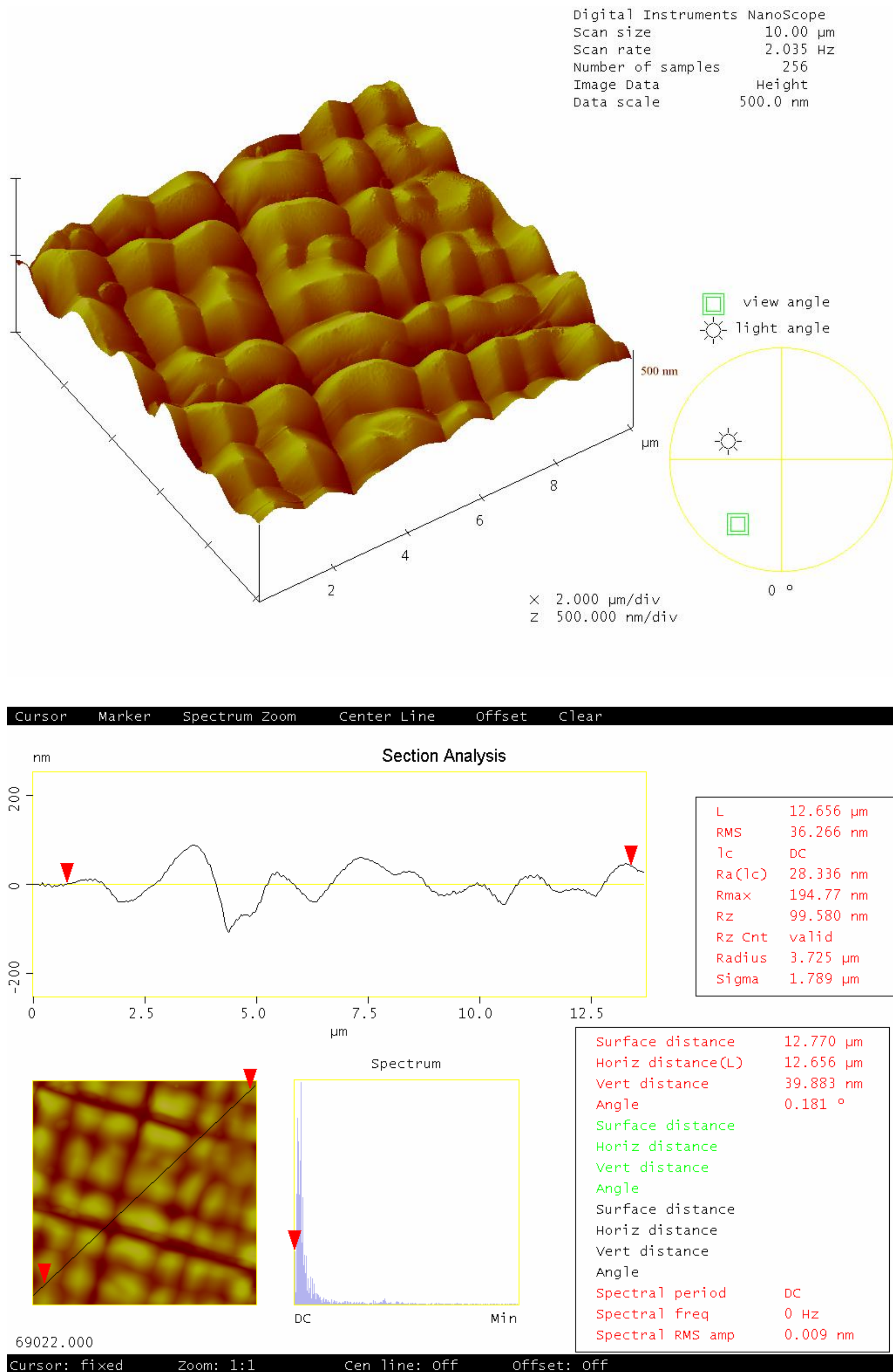


Figure 4.22 AFM analysis of sample TG5.

which shows that the threading dislocation density decreases with the number of steps in the graded layer, having a minimum at 5 – 6 steps. The RMS roughness shows a distinct minimum at 3 – 4 steps. This suggest that the compositional offset of each step should be in the range $x = 0.06 – 0.10$ for this total thickness. However, the actual thickness of the virtual substrates is not clear from the paper.

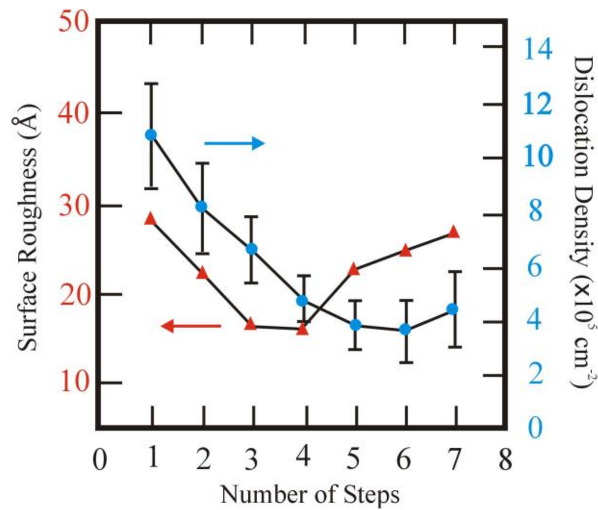


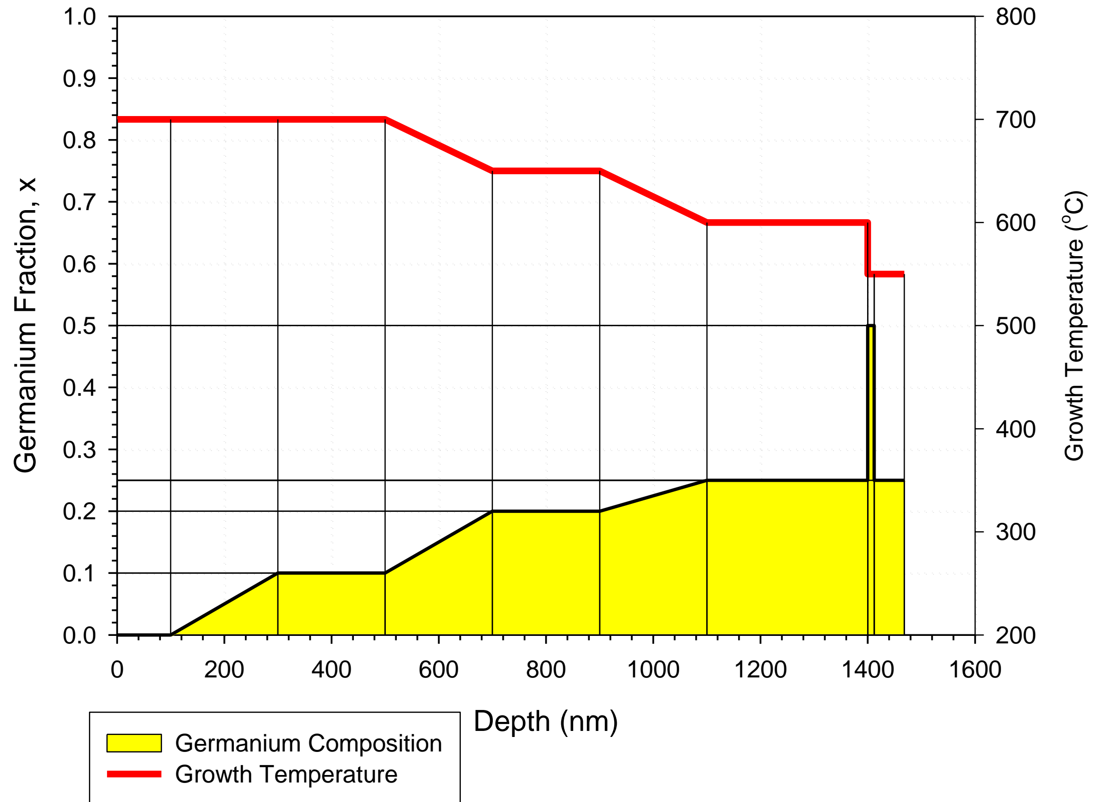
Figure 4.23 Graph showing the effect of the number of steps on the RMS roughness and threading dislocation density of a terrace graded layer of final composition $x = 0.30$. (Taken from Mizushima *et al*[90]).

5 Electrical Characterisation

After it had been shown that the new terrace graded approach to the growth of $Si_{1-x}Ge_x$ virtual substrates is superior to both linearly graded and step graded techniques, the quality of p-type electrical channels grown on these virtual substrates was studied. The technique of growing these substrates by a series of terraces consisting of a linearly graded layer followed by a uniform layer in order to control the dislocation networks needed for strain relaxation should be applicable to any final germanium composition. This is because the dislocation networks and the size of the dislocation pile-ups is controlled by the germanium grading in each step rather than the virtual substrate as a whole. This necessarily involves the full relaxation of each terrace before the growth of the subsequent terraces, so that the layers relax independently. In this chapter virtual substrates with different final germanium compositions and the electrical characteristics of p-type modulation doped channels grown on them have been studied.

5.1 $Si_{0.50}Ge_{0.50}$ Channels on $Si_{0.75}Ge_{0.25}$ Virtual Substrates

The first wafer to be studied is a $Si_{0.75}Ge_{0.25}$ virtual substrate with a $Si_{0.50}Ge_{0.50}$ channel. The same terrace graded approach was applied as sample TG1 in chapter 4. The germanium composition was increased linearly over 200 nm to $x = 0.10$ with a 200 nm uniform layer grown on top, this was then repeated to $x = 0.20$ by another linearly graded layer followed by a uniform layer, both of 200 nm thickness. The final layer of the virtual substrate was graded from $x = 0.20$ – 0.25 over 200 nm with a final uniform layer of $x = 0.25$ of 300 nm thickness. A summary of the growth specification of the virtual substrate is shown in Figure 5.1.



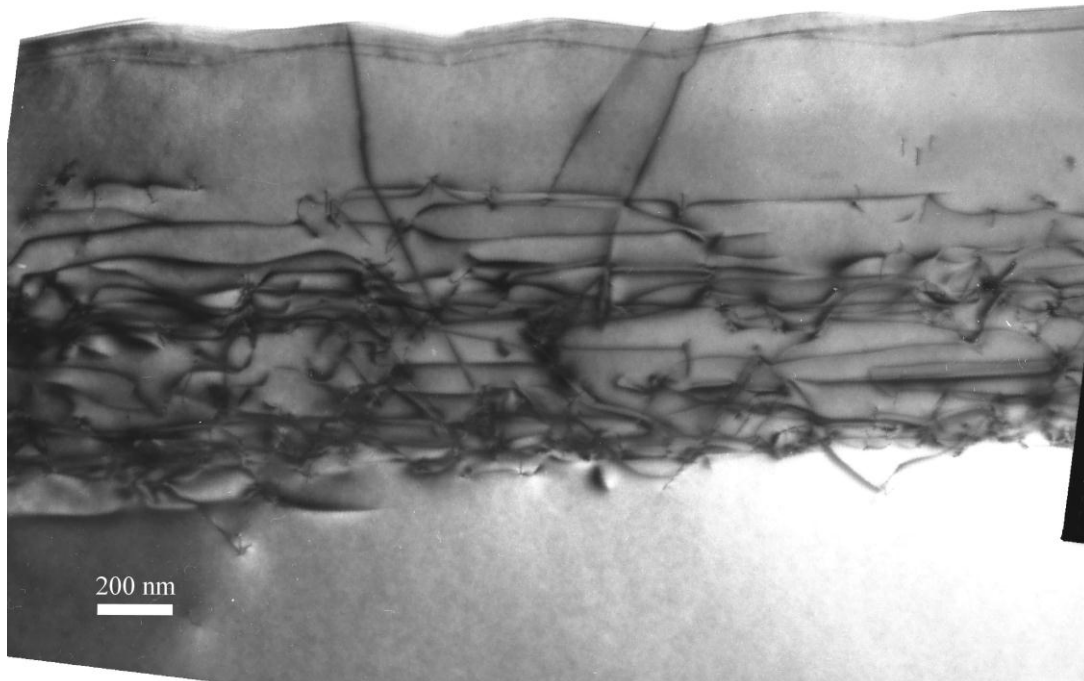
Step	Growth Temperature (°C)	Germanium Composition (%)	Thickness (nm)	Boron Concentration (cm^{-3})
0	700	0	100	-
1	700	0 – 10	200	-
2	700	10	200	-
3	700 – 650	10 – 20	200	-
4	650	20	200	-
5	650 – 600	20 – 25	200	-
6	600	25	300	-
7	550	50	6	-
8	550	25	10	-
9	550	25	50	2.0×10^{18}

Figure 5.1 Growth specifications of sample TGC1.

The electrically active layers were grown on top of this virtual substrate to the following specifications. An 8 nm channel of $x = 0.50$ was grown followed by a 10 nm spacer layer of $x = 0.25$. Following this, a 50 nm doping layer was grown with $x = 0.25$ and a concentration of boron (p-type dopant) of 2.0×10^{18} cm^{-2} in order to populate the channel with p-type carriers (holes). The whole of the electrically active region was grown at a temperature of 550 °C. The thickness of 8 nm for the electrical channel was chosen to be well below the MB critical thickness described in section 2.4.2. This was to ensure that there would be no relaxation of the channel during further thermal processes which would mask any effect of post-growth thermal treatments. It should be noted therefore, that this thickness of channel was not chosen to give the best electrical results.

The bright field XTEM micrograph of this sample designated **TGC1** (Terrace Grade plus Channel) is shown in Figure 5.2(a) using the (220) diffraction for high dislocation contrast. As can be seen, the dislocations are not confined to the graded layers as would be expected using this terrace graded technique. Instead, there are dislocations uniformly distributed throughout the whole of the virtual substrate layers. The uniform layers have not separated the dislocation pile-ups as was discussed in chapter 4, and pile-ups that extend the entire thickness of the wafer can be seen. Some of the dislocations have penetrated the electrical channel and terminated at the surface as threading dislocations. These threading dislocations are thought to be detrimental to the performance of electrical channels and are likely to cause poor results for this sample. It was discussed in chapter 4 that the lower layers of sample TG1 were believed to have relaxed after subsequent layers were grown, leading to incomplete separation of the dislocation pile-ups and this could explain the dislocation structure of this sample. The surface of the wafer shows high amplitude

(a)



(b)

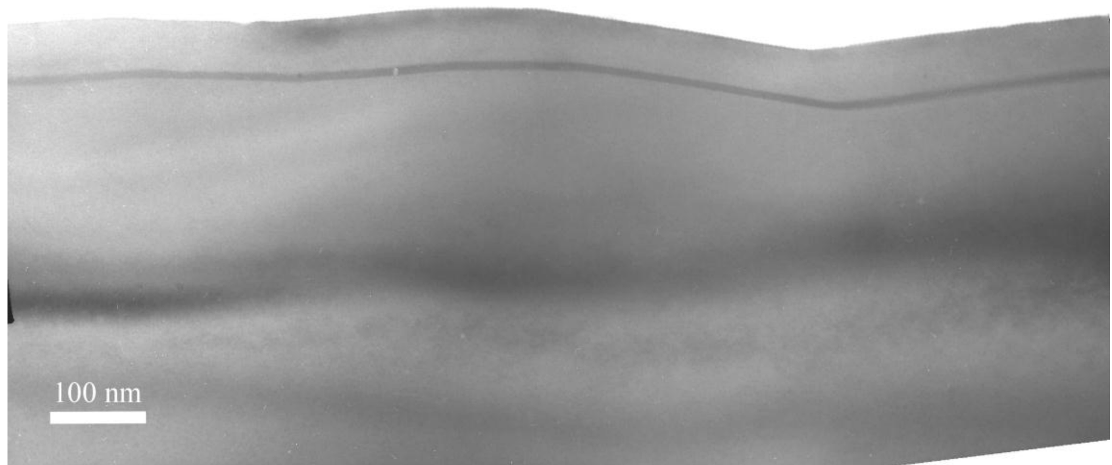
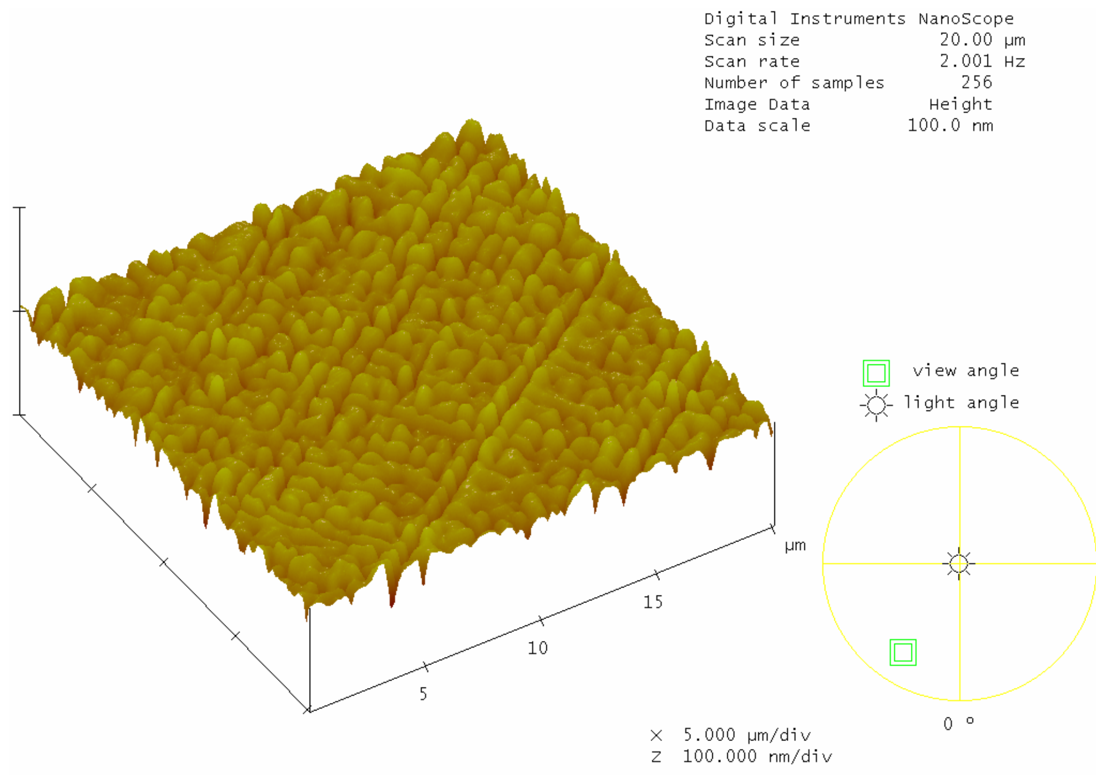


Figure 5.2 Bright field XTEM micrograph of sample TGC1 with $\boldsymbol{g} = (220)$ to image the dislocations (a) and $\boldsymbol{g} = (004)$ to image the channel (b).

undulations with a period of ~ 200 – 400 nm. This is consistent with the relaxation process of surface roughening described in section 2.4.4. In order to observe the channel the (004) diffraction condition was used, it was explained in section 3.1.13 that this diffraction condition gives strong contrast to different states of strain and produces no distortion in the vertical direction and hence is the most useful condition to image the channel. Figure 5.2(b) shows the XTEM micrograph using this diffraction condition. It can be seen that the undulations seen at the surface are also present in the channel, this may also degrade the electrical properties.[57]

The surface of the sample has been studied using AFM in order that the extent of the undulations can be determined. The 3-D surface plots and corresponding line profile are shown in Figure 5.3. As can be seen, the AFM shows a series of large peaks randomly distributed on the surface. In between the peaks can be seen the familiar <110> orientated troughs that are characteristic of the crosshatch associated with the MFR relaxation mechanism, although these are dwarfed by the tall 3-D peaks that are indicative of surface roughening. The RMS roughening of the surface of the sample is 5.7 nm, this is higher than the $Si_{0.50}Ge_{0.50}$ terrace graded virtual substrate TG1. The RMS roughening of a *linearly* graded sample has been shown to increase strongly with terminating germanium composition,[84] so this $Si_{0.75}Ge_{0.25}$ virtual substrate would be expected to have a lower RMS value rather than this larger value, indicating the poor quality of this substrate.

From the XTEM and AFM it is clear that the upper layers of the sample have relaxed by the onset of surface roughening. The virtual substrate should be 100% relaxed before the growth of the electrical layers, and the channel was chosen to be under the MB critical thickness. If this was the case there should be no driving force



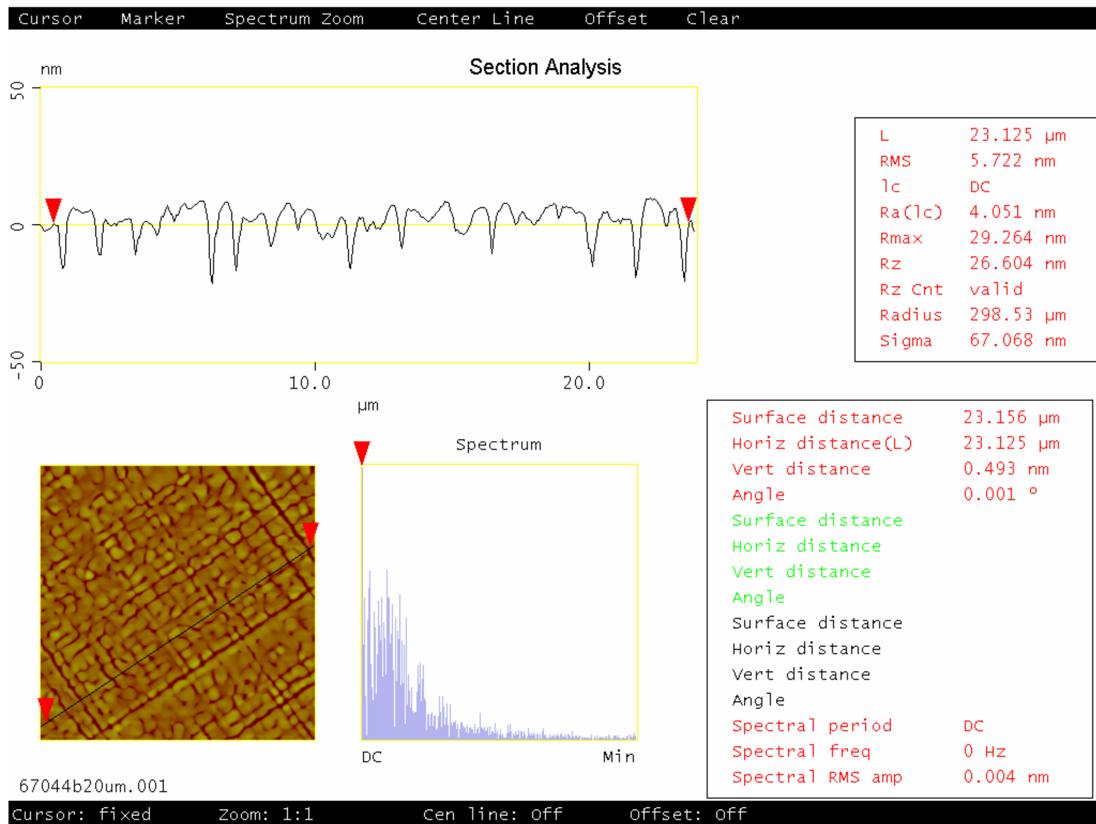


Figure 5.3 AFM of sample TGC1.

for surface roughening, hence it is apparent that the virtual substrate was *not* fully relaxed before subsequent growth. It was proposed in chapter 4 that the lowest layers of the terrace graded sample TG1 did not relax until the upper layer had been grown. Since this sample (TGC1) has a lower final composition there is less strain energy to force relaxation, this could lead to a residual strain as the electrical layers are grown. This residual strain could then relax by the onset of roughening of the channel and the surface of the sample causing the deep undulations seen in the TEM and AFM.

The electrical properties of the modulation doped p-channel was characterised by the van der Pauw (vdP) method described in section 3.4. The sample was subject to a series of *ex-situ* anneals in a dry N_2 ambient atmosphere to enhance the crystalline quality of the channel. As expected, due to the poor structural quality of the sample, the electrical results are poor. Only the low temperature mobility could be measured and the results are summarised in Figure 5.4

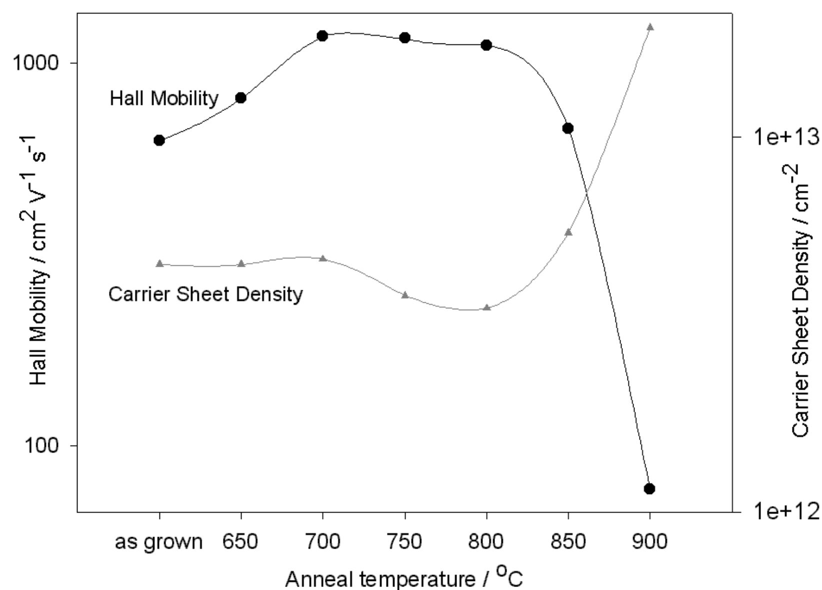


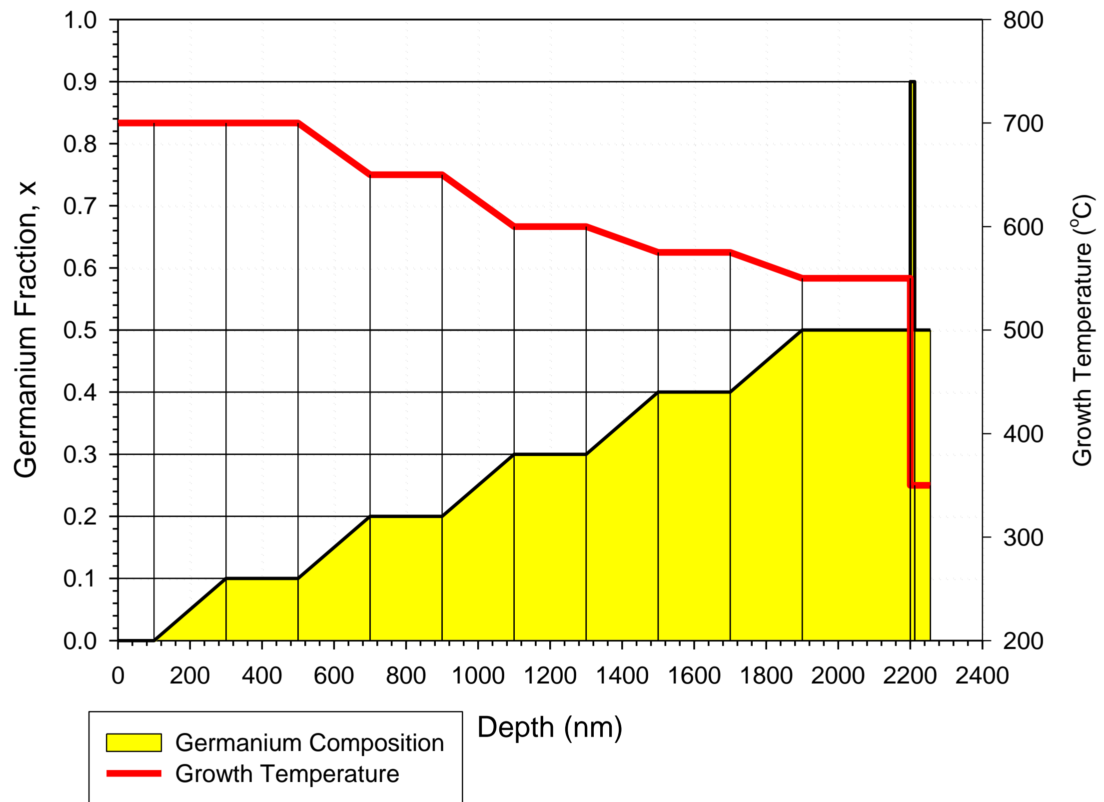
Figure 5.4 Low temperature (10 K) electrical properties of sample TGC1 with different annealing temperatures.

5.2 $Si_{0.10}Ge_{0.90}$ Channels on $Si_{0.50}Ge_{0.50}$ Virtual Substrates

The quality of the $Si_{0.75}Ge_{0.25}$ virtual substrate was found to be of poor structural quality, the dislocation pile-ups were not separated by the uniform layers and the surface and channel had a pronounced roughening. This was attributed to the virtual substrate not fully relaxing by the end of the growth in a similar manner to the *lower* layers of the $Si_{0.50}Ge_{0.50}$ virtual substrate TG1. For this reason it was decided to characterise a p-type channel grown on a virtual substrate grown identically to sample TG1. This virtual substrate has been shown to be fully relaxed from the X-ray analysis of section 4.2.2 so the problem of poor relaxation can be certain to be avoided.

A virtual substrate (designated **TGC2**) was grown with the same specifications as sample TG1 which was summarised in Figure 4.2, on top of this fully relaxed virtual substrate an electrically active channel was grown with the following specifications. A 100 nm $x = 0.50$ layer ensures the channel is kept away from the dislocation networks of the virtual substrate, the channel was then grown to a thickness of 6 nm with $x = 0.90$. A spacer layer was then grown with $x = 0.50$ with a thickness of 15 nm before a 50 nm $x = 0.50$ layer, doped with boron to a concentration of 2.0×10^{18} cm^{-2}. The growth temperature was maintained at 350 °C throughout the electrical layers in order to suppress surface roughening, a summary of these growth specifications is given in Figure 5.5.

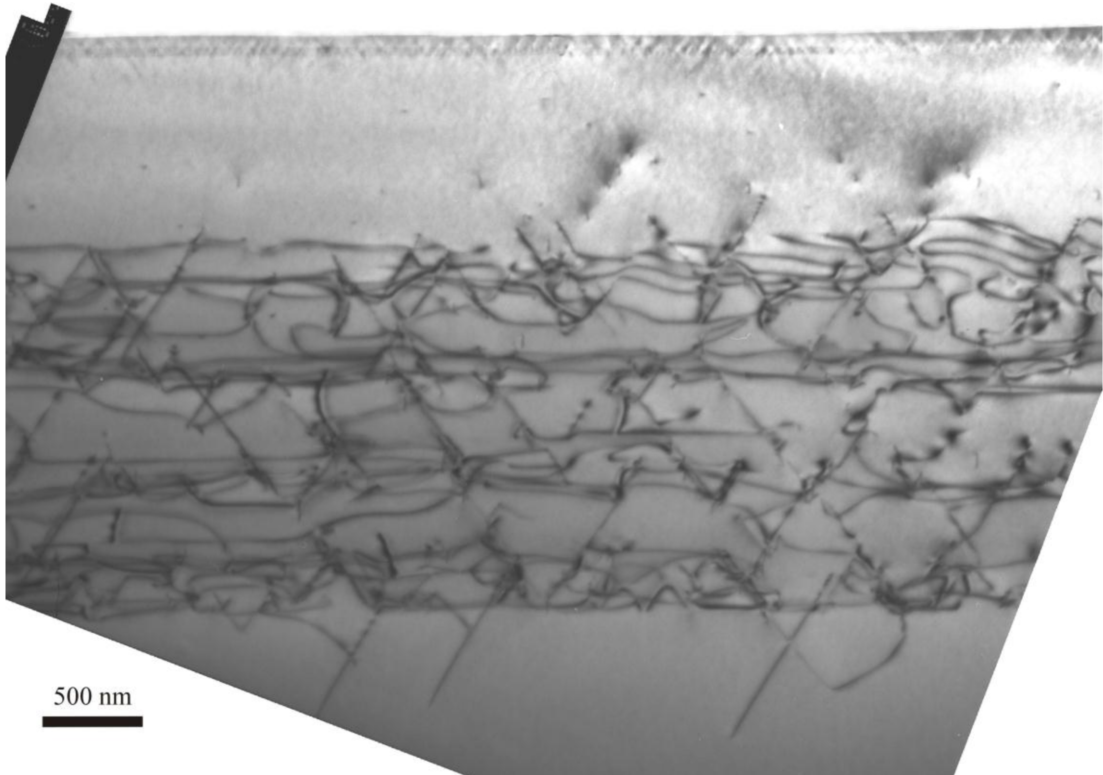
The bright field XTEM of this sample is shown in Figure 5.6(a) using the (220) diffraction condition for high dislocation contrast. The dislocation structure is similar to sample TG1 of Figure 4.6(c) with most misfit dislocations confined to the graded germanium layers. The uniform regions have separated the dislocation pile-ups from these graded layers, ensuring that no pile-ups extend throughout the



Step	Growth Temperature (°C)	Germanium Composition (%)	Thickness (nm)	Boron Concentration (cm^{-3})
0	700	0	100	-
1	700	0 – 10	200	-
2	700	10	200	-
3	700 – 650	10 – 20	200	-
4	650	20	200	-
5	650 – 600	20 – 30	200	-
6	600	30	200	-
7	600 – 575	30 – 40	200	-
8	575	40	200	-
9	575 – 550	40 – 50	200	-
10	600	50	300	-
12	350	90	6	
13	350	50	10	-
14	350	25	50	2.0×10^{18}

Figure 5.5 Growth specifications of sample TGC2.

(a)



(b)

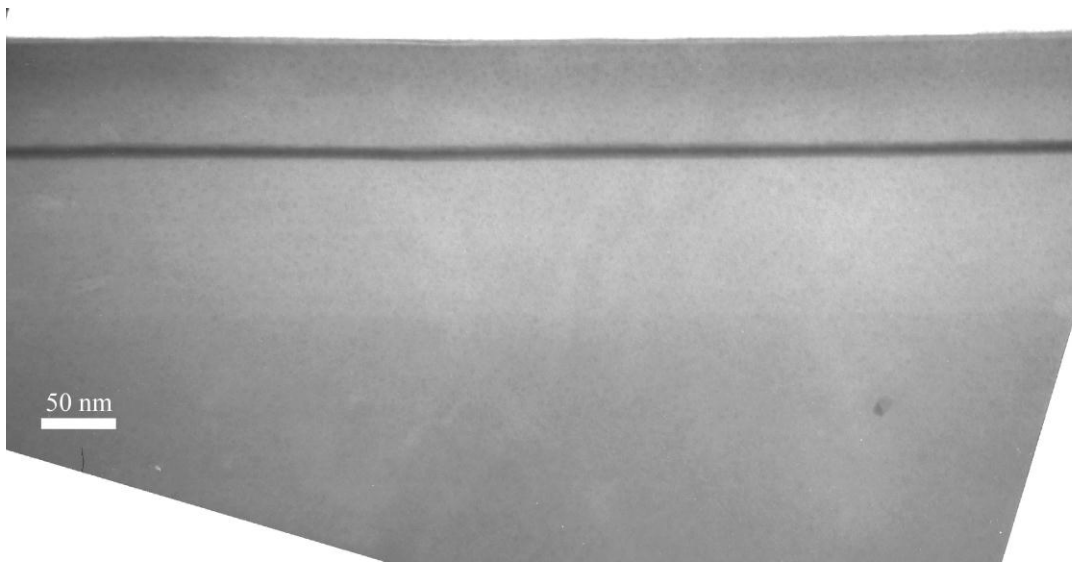


Figure 5.6 Bright field XTEM micrograph of sample TGC2 with ***g*** = (220) for dislocation contrast (a) and ***g*** = (004) to image the channel (b).

sample. As with sample TG1 the lower layers have a higher density of dislocations which is attributed to the late onset of relaxation. The whole of the electrically active layers show no evidence of dislocations, this is a sign that the virtual substrate is of high quality. In order to examine the channel and the surface of the electrical layers, high magnification XTEM was conducted in the (004) diffraction condition. Figure 5.6(b) shows the XTEM micrograph and clearly shows that the channel is highly planar and of uniform thickness. This is a remarkably high quality channel to be grown on a virtual substrate of germanium composition as high as $x = 0.50$ and is evidence of the high quality of the underlying virtual substrate using this terrace grading profile. The surface of the sample can also be seen to be extremely planar, this is important in the subsequent fabrication of electronic devices on such layers.

The quality of the surface was further studied using AFM, the 3-D surface plot and line profile are shown in Figure 5.7. The highly planar surface is immediately evidence in the 3-D surface plot which is drawn to the same scale as the AFM of the virtual substrates discussed in chapter 4. The surface shows evidence of the crosshatch pattern normally attributed to the MFR relaxation mechanism, but the amplitude is very low. The line profile shows the amplitude of this crosshatch, which has an RMS roughness of only 2.1 nm. The RMS roughness is even lower than the sample TG1 (RMS = 3.1 nm) which has the same structure but without the electrically active layers. The effect of the compressive strain in the channel is normally to *increase* the roughness of the surface in order to relieve some of this strain, however this sample has had a decrease in the roughness. This is attributed to the extra thickness of the final uniform layer, before the growth of the channel, causing the effect of the strain fields associated with the dislocation networks to be reduced at the surface.

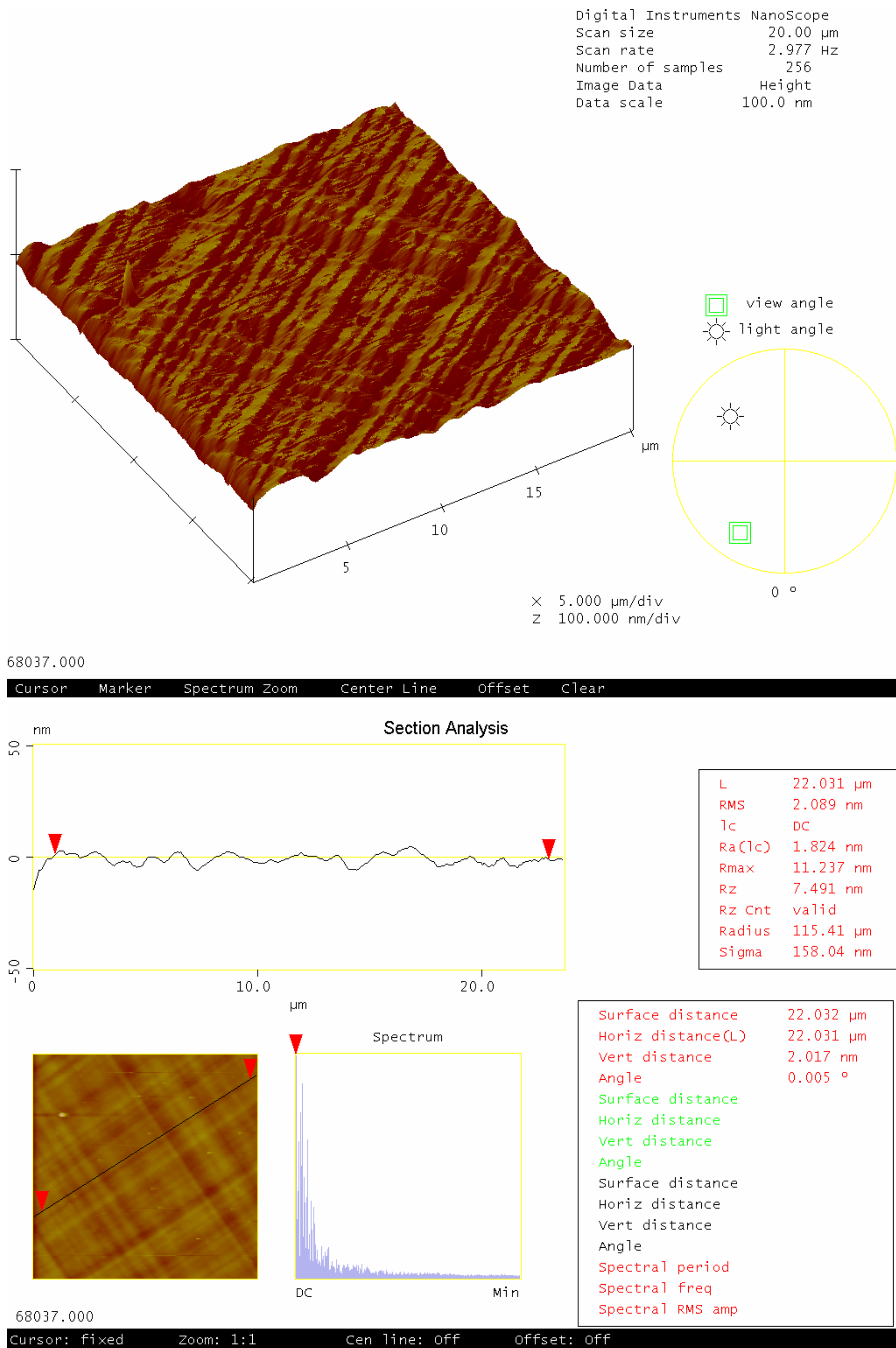


Figure 5.7 AFM data of sample TGC2.

The electrical properties of the $Si_{0.10}Ge_{0.90}$ channel grown on the $Si_{0.50}Ge_{0.50}$ terrace graded virtual substrate, which show such high structural quality, was determined by the vdP method described in section 3.4. The sample was subject to a series of *ex-situ* anneals in a dry N_2 ambient atmosphere. The purpose of the anneals was to enhance the crystalline quality of the channel (by the removal of point defects) and to ensure full activation of the dopants. Samples from wafer TGC2 were annealed for 30 minutes at temperatures of 550, 600, 650 and 750 °C. Hall mobility and carrier concentration was determined as a function of temperature in the range 10 – 300 K, the results are shown in Figure 5.8. At each of the anneal temperatures the hole mobility stays constant up to a temperature of approximately 100 K, this is an indication that the carriers have been successfully confined to a 2-dimensional layer within the channel region, since this flat profile is usually attributed to interface scattering. The as-grown sample clearly has the lowest mobility, this is because of the inherent point defect concentration at the relatively low growth temperature of 350 °C.[91]

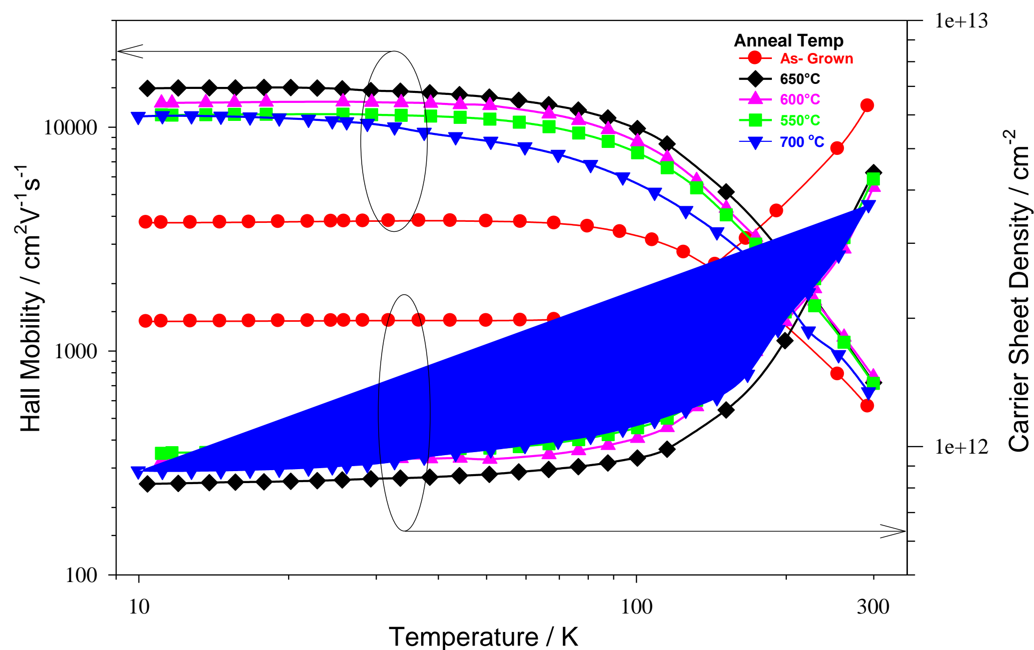


Figure 5.8 Electrical properties of sample TGC2 with different annealing temperatures.

All of the annealed samples show a distinct improvement on the as-grown layers, with the anneal temperature of 650 °C having the greatest hole mobility of 15000 $cm^2V^{-1}s^{-1}$ at a temperature of 10 K. The carrier concentration at this temperature was high at 1.0×10^{12} cm^{-2}. All of the samples show a dramatic reduction in mobility at the temperature rises above 10 K, this is due to the carriers in the doping layer becoming active and forming a parasitic conduction channel as can be seen by the rapid rise in the carrier concentration. These carriers have a significantly reduced mobility due to ionised impurity scattering within the doping layer and reduce the measured Hall mobility. It is possible to eliminate this effect by ensuring that the doping concentration is only high enough to populate the channel, leading to a fully depleted doping layer. In this way there is no parasitic channel to reduce the measured mobility, allowing the accurate measurement of the mobility of carriers in the channel only. Despite this, the room temperature hole mobility has a value of 760 $cm^2V^{-1}s^{-1}$ after an anneal of 600 °C, this shows an improvement over silicon MOS devices with similar effective fields which have hole mobilities of 100 - 200 $cm^2V^{-1}s^{-1}$.[4]

The hole mobility of the sample is quite high considering the specifications of the active layers were not optimised for electrical results. The thicknesses of the sample and spacer layers were chosen so that electrical measurements were guaranteed to be possible. The channel thickness of 6 nm is very low at this compositional offset, but was chosen to eliminate any possibility of partial relaxation of the channel. Work carried out by R.J.H. Morris[92] at the University of Warwick on virtual substrates with similar compositional offsets showed that the optimum thickness of the channel to be between 15 - 20 nm. These showed a very large improvement over channel thickness of 6 nm as was chosen for this study. In further

work the specifications of the active layers should be investigated to optimise the electrical results, it is expected that substantial improvements can be made.

5.3 Discussion

The $x = 0.25$ SiGe virtual substrate (TGC1) has proved to be of poor quality, with large undulation of the channel and surface. Threading dislocations are seen penetrating through the channel which explains the poor electrical results. It was expected that the terrace grading technique should be applicable to any final germanium composition, but this $x = 0.25$ substrate seems to contradict this belief. However, the poor quality at this germanium composition has been attributed to the poor relaxation of the individual terraced layers that was also seen in the $x = 0.50$ virtual substrate TG1. If the full relaxation could be achieved for this lower value of x, then the quality of the virtual substrate should be much higher. This problem of poor relaxation in the lower layers is the subject of the next chapter.

The $x = 0.50$ virtual substrate (TGC2) has proven to be of much higher quality, the relaxation of the layers was already known to be complete from the X-ray analysis of chapter 4. The greater thickness and higher misfit strain allow the higher strain energy to be relaxed through misfit dislocation formation. This has lead to a remarkably smooth channel and surface. The RMS roughness of the virtual substrate with the electrically active layers is actually lower than the virtual substrate alone. This is despite the tendency for a compressively strained channel to cause an increase in surface roughening. This lower RMS roughness is attributed to the extra thickness of the final uniform layer of the virtual substrate, this has reduced the effect of the strain fluctuations from the misfit dislocation networks at the surface. Further work should include the effect of the thickness of this capping layer on the surface RMS

roughness, it is quite probable that further improvement in surface morphology are achievable. Ismail[93] has shown improvements in mobility with capping layers up to 500 nm (above this thickness no further improvement was found). Since the aim of this work was to produce high quality virtual substrates that are comparatively thin compared to conventional grading techniques, the effect of increasing the thickness of the capping layer has not been carried out here.

The electrical results of the $x = 0.90$ channel are particularly encouraging, the thickness of the channel was chosen to make absolutely certain that the layer was below the MB critical thickness and was only 6 nm. At this low thickness the charge carriers are poorly confined and are pushed up against the interface of the virtual substrate and channel. For thick channels the potential well has a triangular form and the carriers are only pushed against one interface, however thinner channels have a more rectangular form and hence the carriers are scattered by both interfaces. This interface roughness leads to scattering of the charge carriers, reducing the mobility.[94] The more these charge carriers are pushed up against these interfaces, the greater the effect of this interface scattering. Another effect is the sensitivity of the energy levels within the well to small variations in the well thickness. For small thickness these energy variations are much greater giving a larger scattering potential. Further work should concentrate on optimising the electrical layers, in particular concentrating on channel thicknesses of 15 – 20 nm where large improvements are expected to be made.

6 Relaxation Processes in Terrace Graded Virtual Substrates

6.1 Introduction

In order to optimise the growth schedule of the new terrace graded structure it was necessary to study the mechanism by which the SiGe layers relax. The main mechanisms by which relaxation usually occurs in SiGe layers was explained in section 2.5. For the terrace graded profile, the most efficient relaxation occurs by the MFR mechanism in each of the graded layers. The uniform layers provide further strain energy to ensure that the graded layer is completely relaxed and also gives the threading arms a chance to annihilate as described in section 2.5.4. The subsequent layers can then be grown and relax in a similar manner. The uniform layer also ensures that the dislocation pile-ups caused by the MFR mechanism in one layer do not extend into the others. In order that the most efficient relaxation can occur it is important that as each layer grows, it is fully relaxed before commencing growth of the subsequent layers. If the growth continues before the lower layers are fully relaxed, then when these layers do relax, the dislocations will be pushed down from higher up, rather than nucleating within the layer. This increases the pile-ups height since a greater amount of strain must be relieved by each multiplication site. The increased size of the pile-ups will cause a greater amplitude of the surface undulations since the amplitude is proportional to the number of dislocations on each glide plane.[69] Also, since more strain needs to be relieved, the chances of dislocation nucleation and the associated density of threading dislocations are increased.

It was shown in section 4.2.2 that all the layers in the complete terrace graded structures are fully relaxed. There was also evidence that the lowest layers have

become over-relaxed, most likely from an over-saturation of dislocations that have been pushed into these layers from MFR mechanisms operating in later layers. This is probably because these later layers were grown *before* the lowest layers were relaxed. The strain in these lower layers would enable the dislocations to punch through the uniform regions and relax the graded layers beneath. Once dislocations have been able to penetrate the lower layers by dislocation injection from multiplication sites higher up, then it is possible that this efficient dislocation source could lead to an over-saturation of dislocations.

6.2 Study of the Relaxation of the x = 0 .10 Layers

In order to determine the exact point at which each layer in the terrace graded structures relax, two test wafers were grown. The first wafer was grown containing only the first step of the terrace grade, this consisted of a 200 nm graded layer to $x = 0.10$ followed by 200 nm of uniform $x = 0.10$ SiGe. The growth temperature was maintained at 700 ^{o}C. To check the effectiveness of an *in-situ* anneal on the relaxation of this $x = 0.10$ layer, a second wafer was grown with exactly the same specification but followed by a 30 minute anneal at 900 ^{o}C. This anneal should ensure that the layer is completely relaxed.

6.2.1 XTEM

In order to study the relaxation of these $x = 0.10$ layers the samples were prepared for XTEM and studied in a JEOL 2000FX TEM as described in section 3.1. The micrographs of the two samples are shown in Figure 6.1(a,b). It can be seen that the sample without the *in-situ* anneal has very few dislocations. This is consistent

with the hypothesis of chapter 4 that the first layers of the terrace graded wafers do not relax until subsequent layers have been grown. The dislocations that can be seen are very long and only occur at the bottom of the graded layer. If the step had been completely relaxed then the dislocations should be uniformly distributed throughout the linear grade (the first 200 nm of the epilayers). Since with such a lower density of dislocations, the chance of dislocation pinning is much reduced, the threading dislocation density should be low. The reduced dislocation interactions, which are necessary to produce multiplication mechanisms, explains the low relaxation that is evident from the XTEM. In contrast, the sample annealed at 900 ^{o}C shows a much denser distribution of dislocations. The layer is obviously in a much higher state of relaxation than the un-annealed wafer. There is evidence of dislocation pile-ups characteristic of the MFR mechanism that was also seen in the terrace graded wafer. There are several pile-ups of about 3 - 4 dislocations and a few dislocation have been pushed into the substrate. As explained in section 2.5.4 this is the signature of the MFR multiplication mechanism.

The un-annealed sample was given an *ex-situ* anneal in a N_2 ambient atmosphere at 900 °C for 30 minutes to compare the effect of *ex-situ* and *in-situ* annealing. It was expected that both the samples subject to the anneals would show similar characteristics, but as the XTEM in Figure 6.1(c) shows, there is a marked difference. The XTEM shows that the *ex-situ* annealed sample has become much more relaxed, with a greater dislocation density which is distributed throughout the graded layer. However, instead of there being well spaced multiplication sites, like the *in-situ* annealed sample, the dislocations are much more closely packed and have been pushed deep into the substrate, much further than the *in-situ* annealed sample.

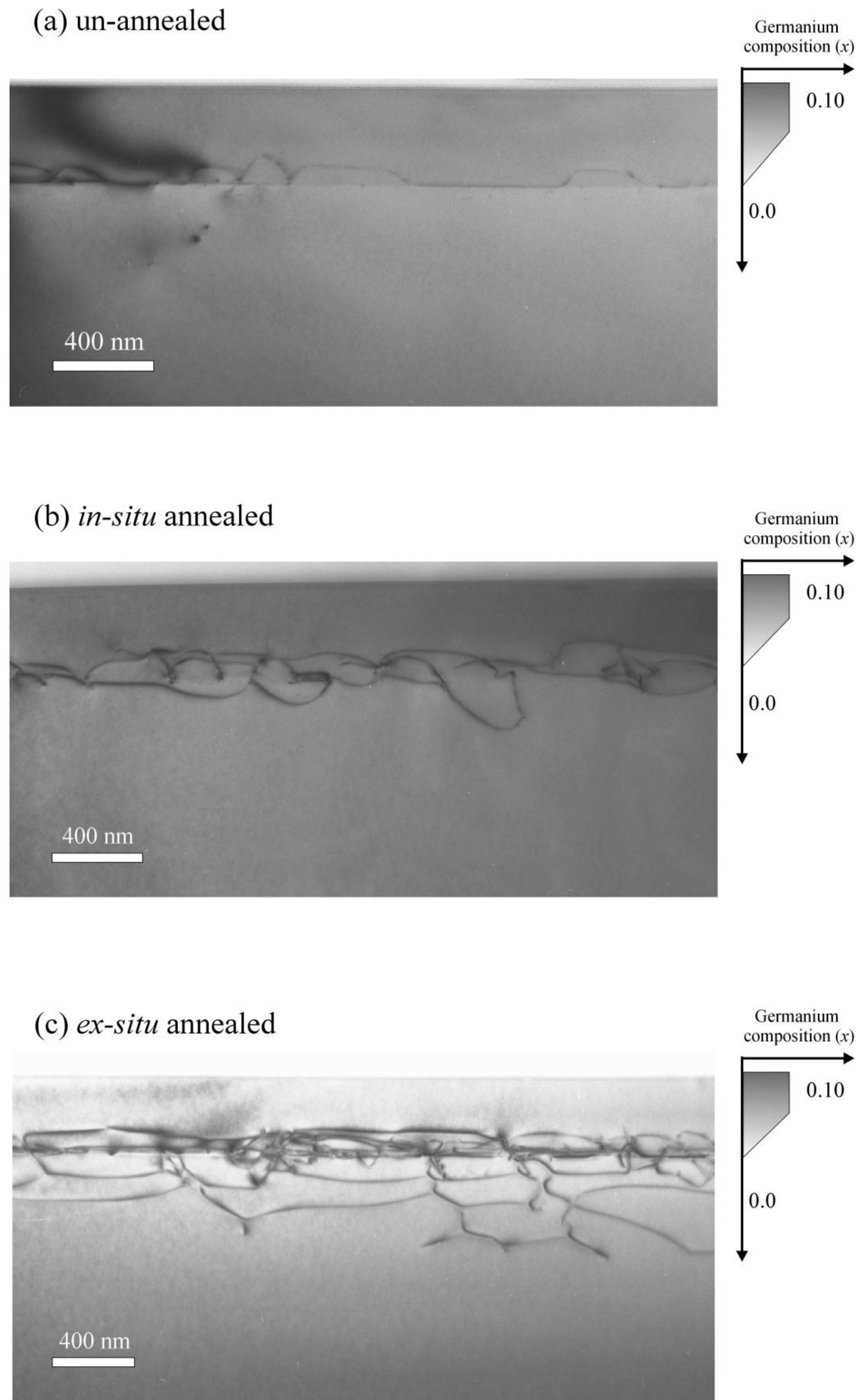


Figure 6.1 Bright field XTEM of the $x = 0.10$ layers using the $\boldsymbol{g} = (220)$ diffraction condition. (a) shows the un-annealed sample, (b) shows the *in-situ* annealed sample and (c) shows the *ex-situ* annealed sample.

One possible reason that the *in-situ* and *ex-situ* anneals have such different effects on the dislocation structure of the $x = 0.10$ layer may be due to the difference in temperature ramping. The *in-situ* annealed sample was slowly heated from the growth temperature to the annealing temperature under UHV conditions. After the annealing time had passed the sample was allowed to cool slowly until it reached room temperature. In contrast, the *ex-situ* annealed sample was pushed into a furnace already at the annealing temperature and consequently the temperature of the sample was raised very rapidly. After the annealing time had expired the sample was removed from the furnace and allowed to cool, since the sample has a small heat capacity its rate of cooling was also extremely high and reached room temperature in only a few seconds. The effects of these different temperature cycles could be the cause of the difference in the dislocation structures. The slow thermal cycle of the *in-situ* anneal would allow the dislocations to keep up with the excess strain caused by the increased temperature, encouraging long misfit dislocations. Also, since the process of glide has a lower activation energy than dislocation nucleation or multiplication, the glide process will be thermally activated at lower temperatures than these mechanisms. Consequently, as dislocations are formed, they may glide to keep up with the strain relaxation, the nucleation of the dislocations acting as the bottleneck. This will encourage long dislocation lengths, and consequently less multiplication or nucleation sites. In contrast, the rapid temperature ramp of the *ex-situ* anneal will activate the nucleation and multiplication mechanisms concurrently with the process of glide, this will remove the bottleneck of dislocation formation and favour strain relaxation by these mechanisms. The consequence of this is that the misfit dislocation lengths are shorter and formed at more closely packed sources.

Another difference that must be considered is the conditions under which the samples were annealed. The *ex-situ* anneal was carried out in a N_2 atmosphere and would have a native oxide on the surface. In contrast, the *in-situ* anneal was carried out under UHV conditions, and would not have any oxide at all on the surface. It has been shown that the dislocation velocity is strongly influenced by the surface conditions and presence of an oxide and so this may account for the differences in relaxation mechanisms between the two samples.[95] Further work could clarify the effect of the temperature ramping rates and surface conditions. For example, to study the effect of the temperature ramping rate, an *ex-situ* anneal with a ramp rate comparable to that of the *in*-situ anneal could be performed. The effect of the native oxide could be studied by subjecting the sample to a dilute HF dip prior to an *ex-situ* anneal. The HF removes any oxide from the surface of the sample and terminates it with hydrogen, preventing re-oxidation. The sample could then be *ex-situ* annealed without the effect of the oxide on the surface.

6.2.2 AFM

In order to see how the different annealing procedures affect the surface topography, the samples were examined using AFM. The 3-D surface plots and corresponding line profiles with roughness data of the un-annealed, *in-situ* annealed and *ex-situ* annealed sample are shown in Figure 6.2, Figure 6.3 and Figure 6.4, respectively. The un-annealed sample has undulations on the surface that is similar to the crosshatch pattern expected from the MFR mechanism. It has troughs and peaks with a spacing ~ 1 μm, although the surface morphology does show very small islands associated with the start of 3-D growth. This is likely to be due to the

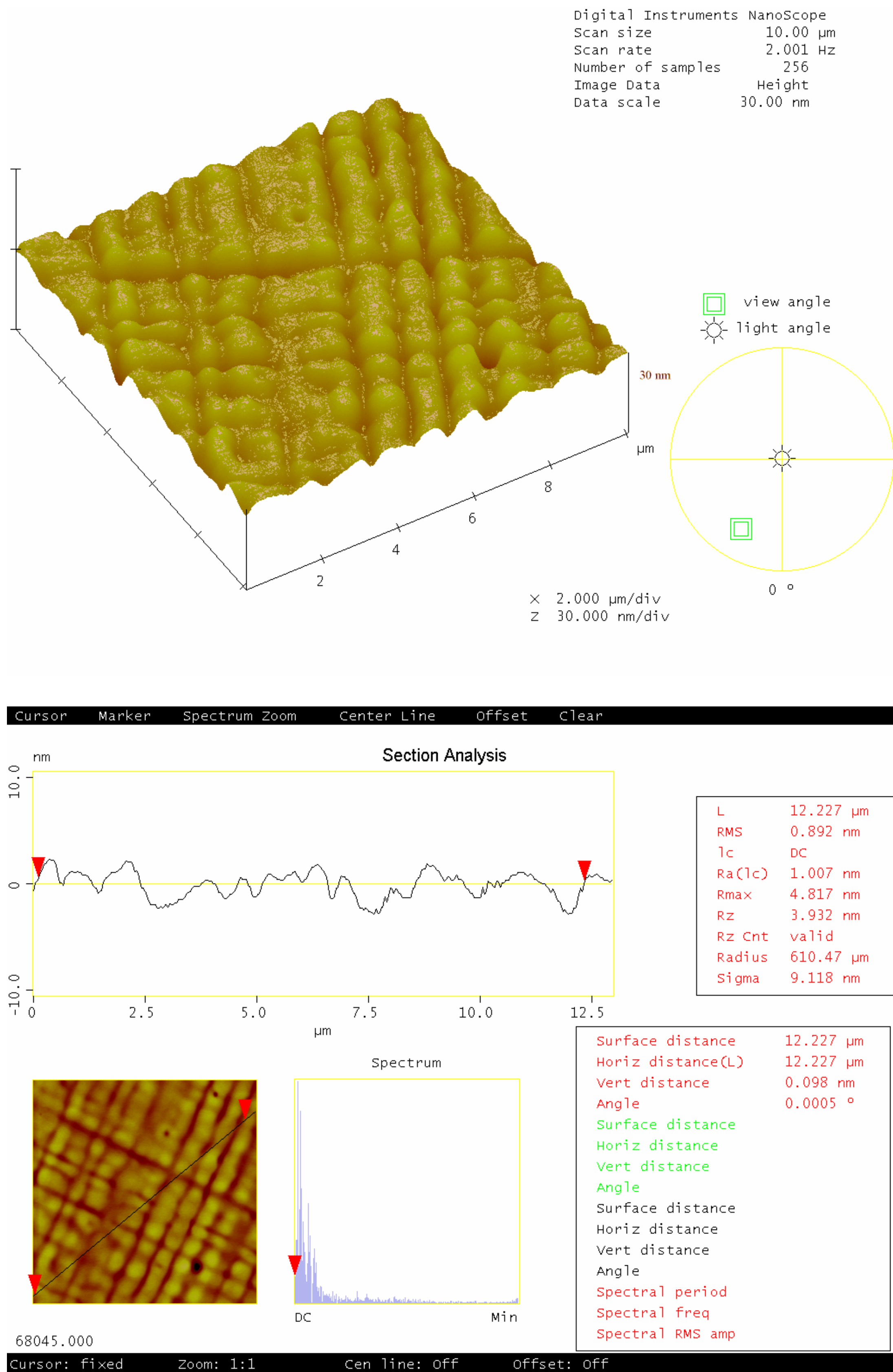


Figure 6.2 AFM 3-D surface plot and line profile of the un-annealed sample.

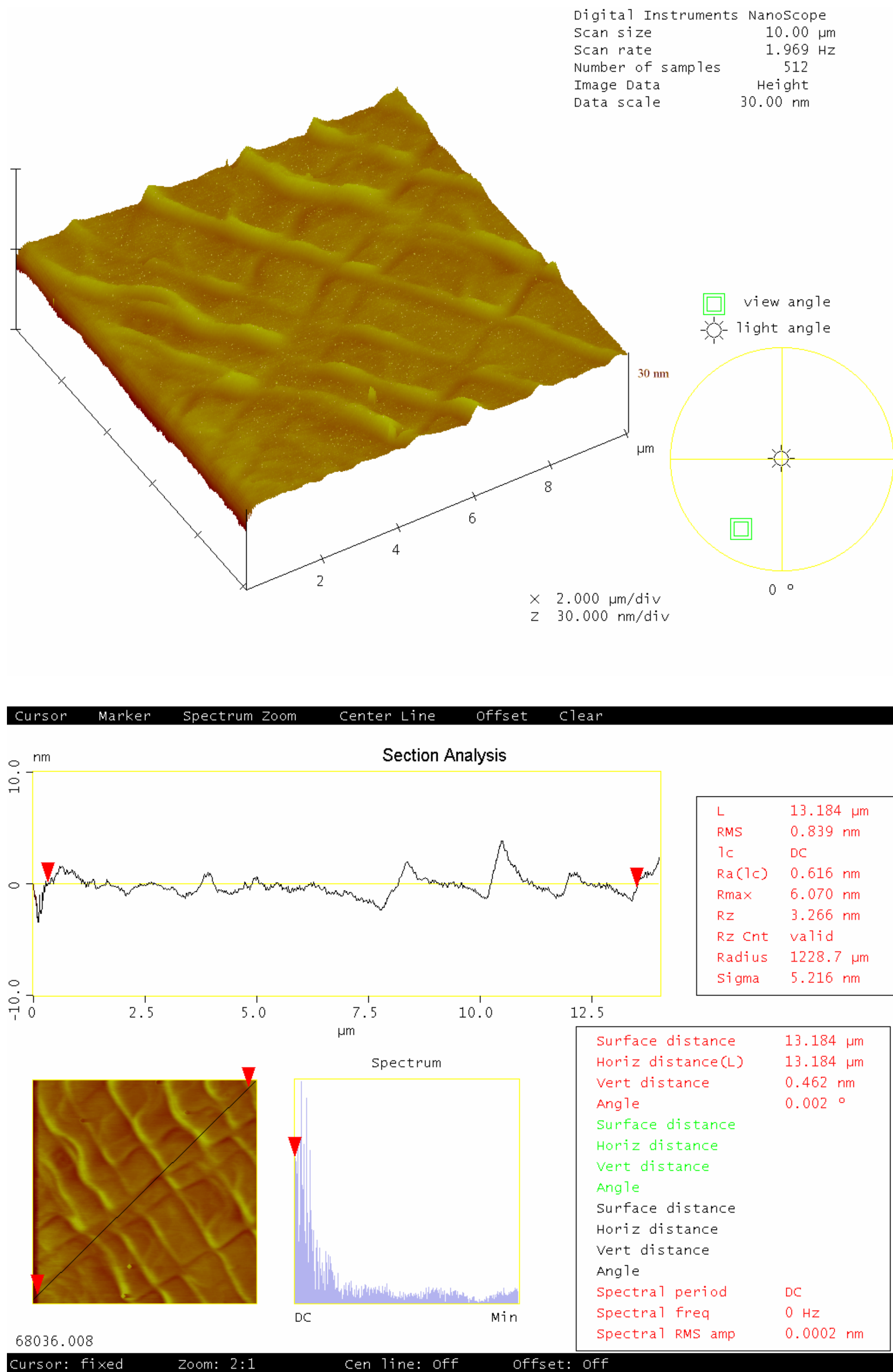


Figure 6.3 AFM 3-D surface plot and line profile of the *in-situ* annealed sample.

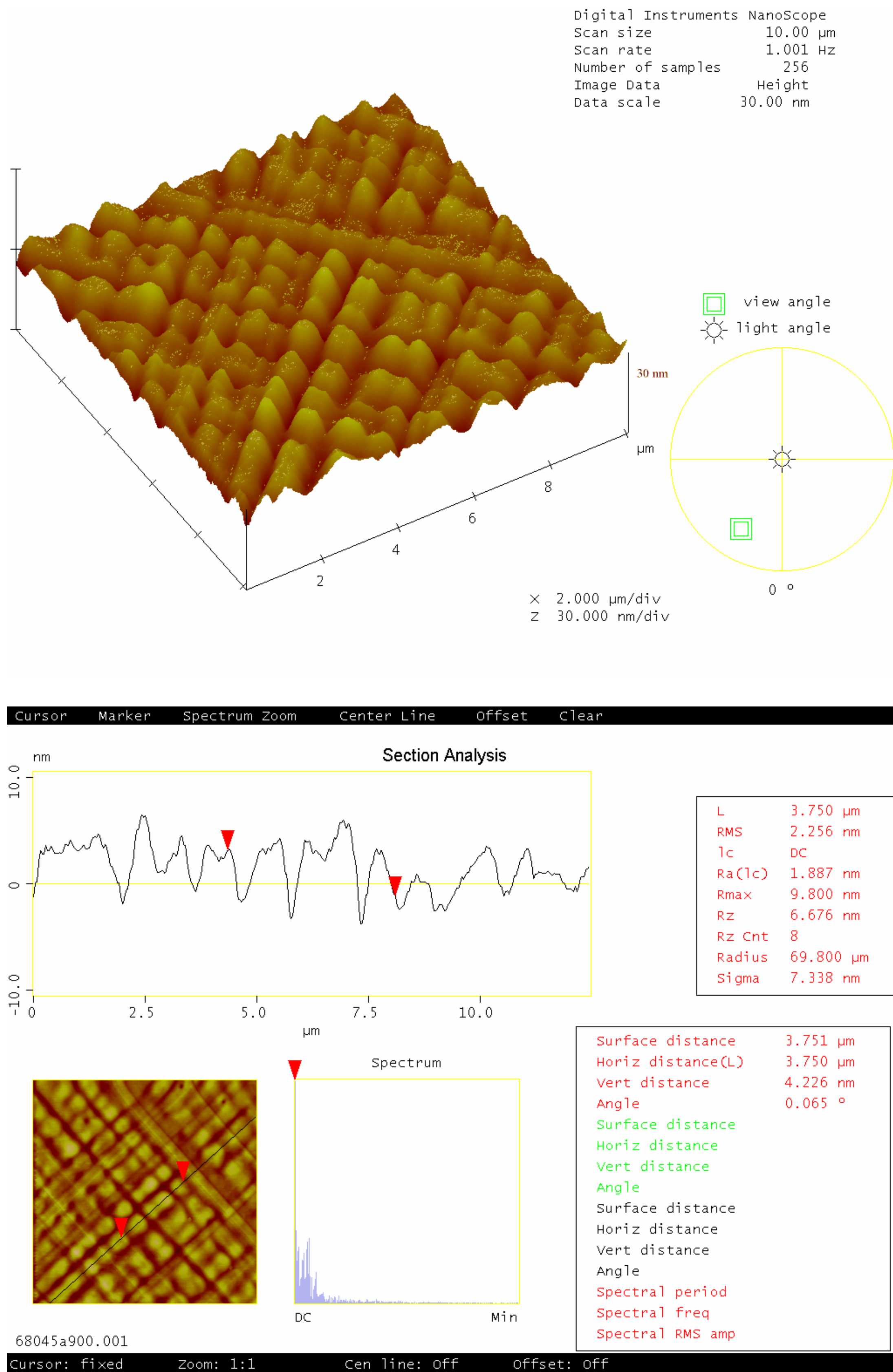


Figure 6.4 AFM 3-D surface plot and line profile of the *ex-situ* annealed sample.

there being insufficient dislocations present to allow relaxation, and so the surface has corrugated in order that some strain relaxation can occur from this mechanism. The RMS roughness of this sample is shown to be 0.9 nm.

The AFM of the *in-situ* annealed sample has a very different surface morphology. The 3-D growth of the un-annealed sample has given way to a series of very well defined peaks and troughs that is expected from the MFR mechanism. In between these troughs the surface is highly planar. Comparison of the line profiles of the un-annealed and *in-situ* annealed samples clearly shows this difference in topography. The *in-situ* annealed sample has distinct and high peaks that can be attributed to multiple dislocations on the same atomic glide plane, this is in accordance with Lutz *et al*[69]. The line profile of the un-annealed sample has many more peaks but with a lower amplitude. The RMS roughness of the *in-situ* annealed sample is shown to be 0.8 nm, which is actually lower than the un-annealed sample, although not significantly.

The *ex-situ* annealed sample has a significantly larger RMS roughness of 2.3 nm, the line profile shown in Figure 6.4 shows that the peaks and troughs have been enlarged considerably over the un-annealed sample. This is a consequence of the strain relaxation under high annealing temperature and is supported by the XTEM of Figure 6.1. The 3-D surface plot shows that there are still islands on the surface of the sample, but the troughs and peaks associated with the underlying dislocation network are much more evident, since now there is much greater density of dislocations.

6.2.3 X-ray

XRD was carried out on the *in-situ*, *ex-situ* and un-annealed samples to ascertain their state of relaxation. The results are summarised in Table 6.1. It is clear that the *in-situ* annealed sample has close to $x = 0.10$ as intended and is highly relaxed. However, the un-annealed sample seems to show unexpected results. The analysis of the X-ray rocking curves according to section 3.2 suggest that the germanium composition is $x = 0.18$, much higher than was intended and the relaxation of 85% is much higher than the XTEM suggests. This is not a rogue result, but a short falling of the analysis technique used. As stated in section 3.2 the composition and relaxation can only be accurately determined by this method if the relaxation is high. This is because the rocking curve only samples the line in reciprocal space that relaxed layers fall on. So this method can only determine the exact position of the SiGe layer peak if it lies near to this fully relaxed locus. The fact that the analysis from this sample gives clearly erroneous results is indicative that the layer is far from the fully relaxed state. To accurately determine the relaxation it would be necessary to take a full reciprocal space map of the silicon and SiGe peaks. Unfortunately, this takes a very long time to obtain and consequently this is the subject of further work. Despite this, it is fair to say from the XTEM and the XRD, that the un-annealed sample is significantly strained, but impossible to say exactly to what extent. With the addition of an *ex-situ* anneal the relaxation is increased and the measured value of x falls. Since this is *exactly* the same sample that

	un-annealed	*in-situ* annealed	*ex-situ* annealed
Composition, x	0.181	0.115	0.155
Relaxation (%)	85.3	90.1	87.6

Table 6.1 XRD results of the un-annealed, *in-situ* annealed and *ex-situ* annealed $x = 0.10$ layers.

without the anneal, the effect of the poor relaxation on the measured value of x is was measured clear, since both samples must have the same germanium composition. However, even *with* this *ex-situ* anneal, the sample is not as highly relaxed as the *in-situ* annealed sample.

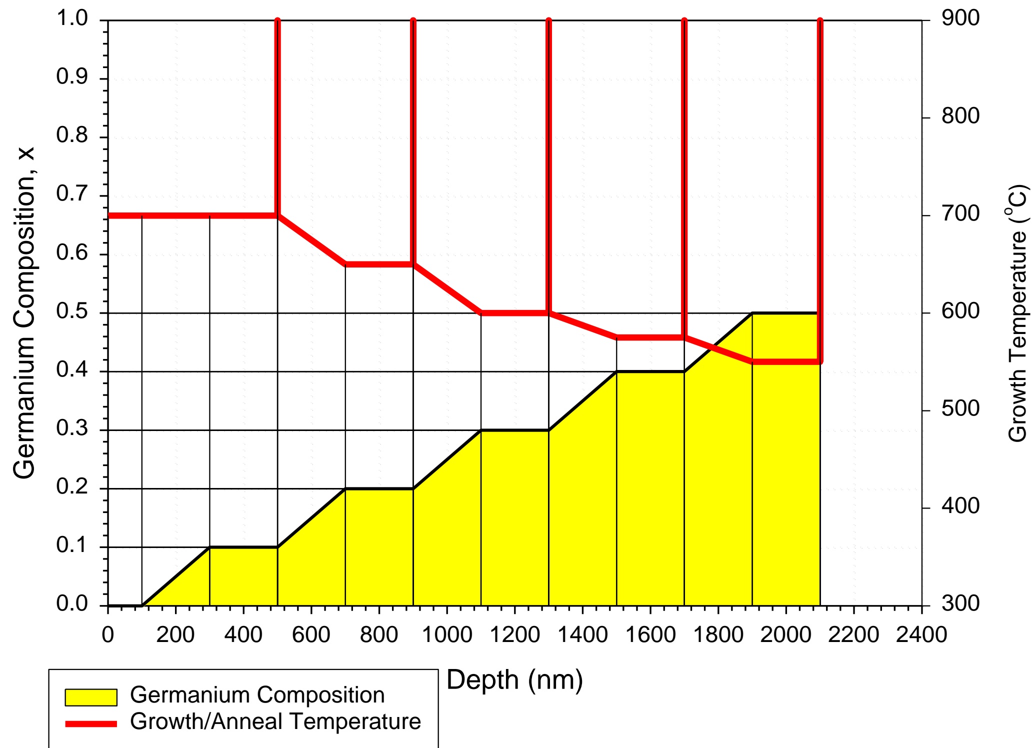
6.2.4 Conclusion

It has been shown that the first layers of the terrace graded virtual substrates are indeed still strained before the second layer is grown. This was postulated in chapter 4 due to the higher density of dislocations in the XTEM and the over-relaxation of these layers. It was also explained how this could adversely effect the quality of the surface of the virtual substrate, since the multiplication sources which act to relax each of the graded layers will not be separated as intended by this germanium grading profile. Importantly, this strain can be relieved by a 900 °C *in-situ* anneal for 30 minutes leading to a dislocation network that is well spaced and completely contained within the graded layer. The surface topography of the sample after the anneal shows the expected crosshatch pattern with a low RMS roughness of less than 1 nm, indicating a high quality layer. If each layer in the $Si_{0.50}Ge_{0.50}$ terrace graded virtual substrate can be relieved in the same way, then the surface roughness of this wafer should have a value that is not much greater than the first $x = 0.10$ layer.

6.3 Study of the Effectiveness of High Temperature Anneals During the Growth of the Terrace Graded Virtual Substrates

Now that it has been established that the lower layers of the terrace graded virtual substrates are not fully relaxed before subsequent growth, and that a 900 °C anneal will relieve this residual strain, further attempts to improve the terrace graded wafers were tried. A terrace graded wafer was grown with the same specification of TG1 in chapter 4 but with the addition of a 30 minute anneal at 900 °C after the growth of each of the uniform layers. The growth specification of this wafer, designated **TGA1** (Terrace Grade plus Anneal), is shown in Figure 6.5. It was hoped that the anneal after each uniform layer would ensure that the underlying layers would relax and that subsequent growth could proceed independently of the underlying dislocation networks.

The XTEM of this sample is shown in Figure 6.6, it shows that the *lower* layers have a much higher quality dislocation network than sample TG1 shown in Figure 4.6(c) which did not have the strain relieving anneals. The $x = 0$ .10, 0.20 and 0.30 layers have well separated dislocation structures in each of the graded layers. These misfit dislocations are very long and have only a few dislocation pile-ups as was hoped. The uniform layers are nearly free of dislocations, indicating the complete separation of the dislocation networks from each of the terrace graded layers. The density of the dislocations in the graded layers appear to be the same in each terraced step, this fact indicates that each layer has relaxed completely before subsequent growth since no dislocations have been pushed down from higher layers to complete the relaxation of the lower layers. This is in clear contrast to the XTEM of sample TG1 which shows that the uniform layers have many dislocations,



Step	Growth Temperature (°C)	Germanium Composition (%)	Thickness (nm)
0	700	0	100
1	700	0 – 10	200
2	700	10	200
3	900	Anneal	-
4	700 – 650	10 – 20	200
5	650	20	200
6	900	Anneal	-
7	650 – 600	20 – 30	200
9	600	30	200
8	900	Anneal	-
10	600 – 575	30 – 40	200
12	575	40	200
11	900	Anneal	-
13	575 – 550	40 – 50	200
14	550	50	200
15	900	Anneal	-

Figure 6.5 Growth specifications of sample TGA1.

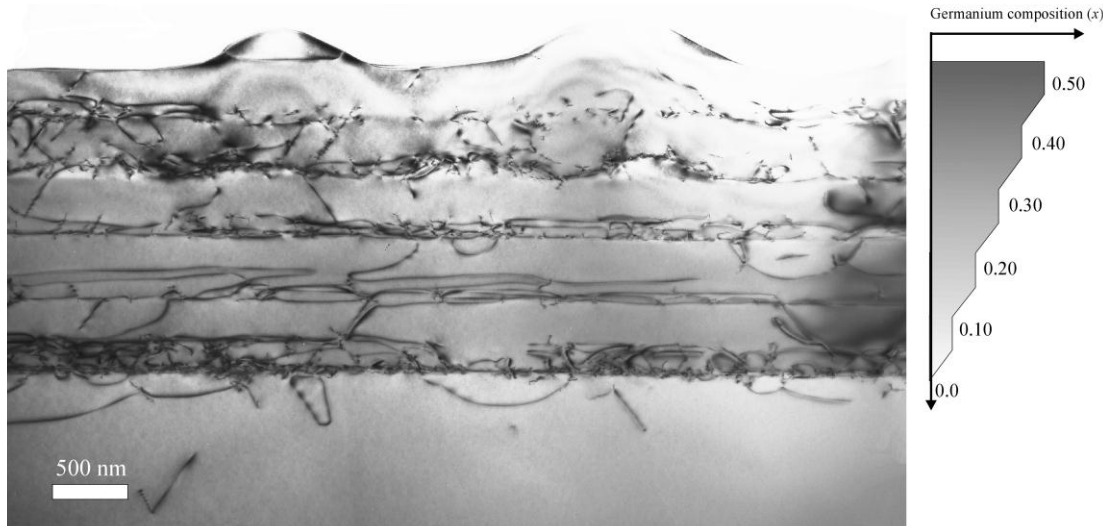


Figure 6.6 Bright field XTEM of sample TGA1 using the $\boldsymbol{g}$ = (220) diffraction condition.

and that the lower layers have a much higher dislocation density than the higher layers.. This form of strain relaxation was the intended aim of the annealing steps. However, after the growth of the $x = 0.30$ layers there is a sudden change of character of the terraced layers. Instead of the highly planar and well separated dislocation networks, the growth of the layers has pronounced 3-dimensional features. Large rounded mounds of approximately 500 nm in width and 200 nm height can be seen at the interface between the $x = 0.30$ uniform layer and the $x = 0.30 - 0.40$ graded layer, with a spacing of approximately 2 μm. A higher magnification of one of these features is shown in Figure 6.7. In this micrograph it can be seen that the presence of these features dramatically changes the character of the dislocations. The mounds themselves do not appear to contain any dislocations, but at the edge of these features there is evidence of a significant increase in the number of dislocations nucleation events. The presence of these mounds is likely to have caused a considerable increase in the local strain of the layers above, causing early dislocation nucleation in order to relieve this strain. The surface morphology of

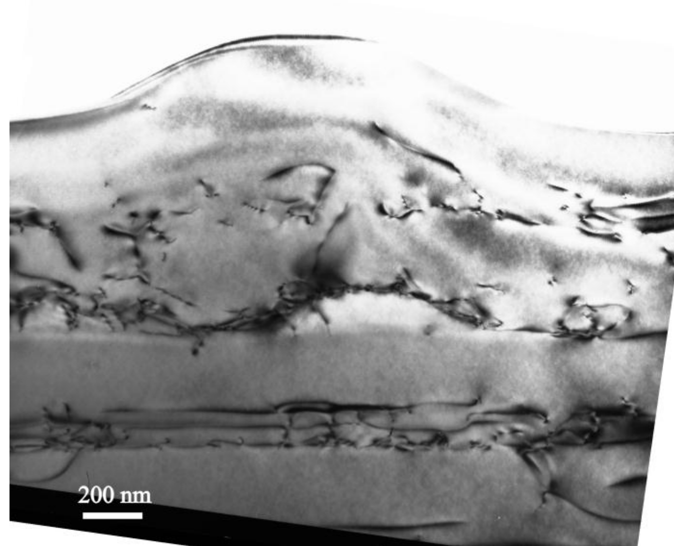


Figure 6.7 Higher magnification of the mounds seen in Figure 6.6. Dislocations can be seen to have nucleated around the outside of the mound.

this sample has clearly been affected by the presence of these mounds by forming even larger features directly above them. These large surface features are to be expected with overgrowth over such large non-uniformities. As well as the overlying layers being forced to conform to these irregular shapes, the strain fields at the surface are also likely to be strongly affected, this would further exaggerate the height of the mounds due to local differences in growth rates.

Figure 6.8 shows an XTEM micrograph of the same sample, but this time in the (004) diffraction condition. As explained in section 3.1.13, under this diffraction condition strong contrast occurs for different strained states, rather than the strong dislocation contrast seen in the (220) diffraction condition. It can be seen that the lower layers have a very even contrast, further showing the uniform state of relaxation in these layers. However, the mounds have a very light appearance in this condition which is an indication that there is a much higher silicon concentration in these mounds than the surrounding material. In order that this assertion could be tested, the X-ray spectrum of the sample was taken. In this technique the electron

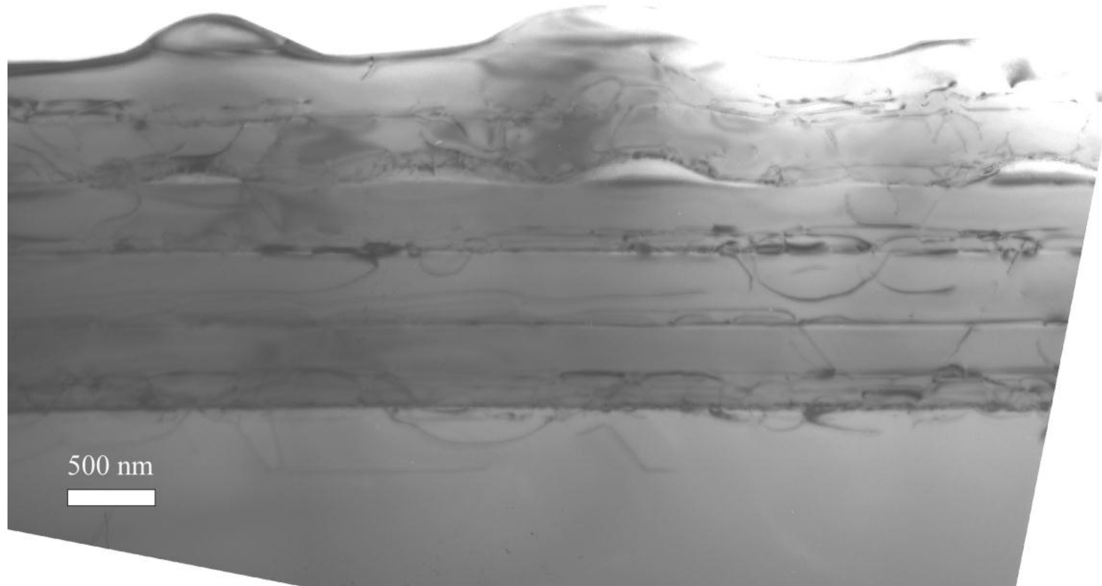
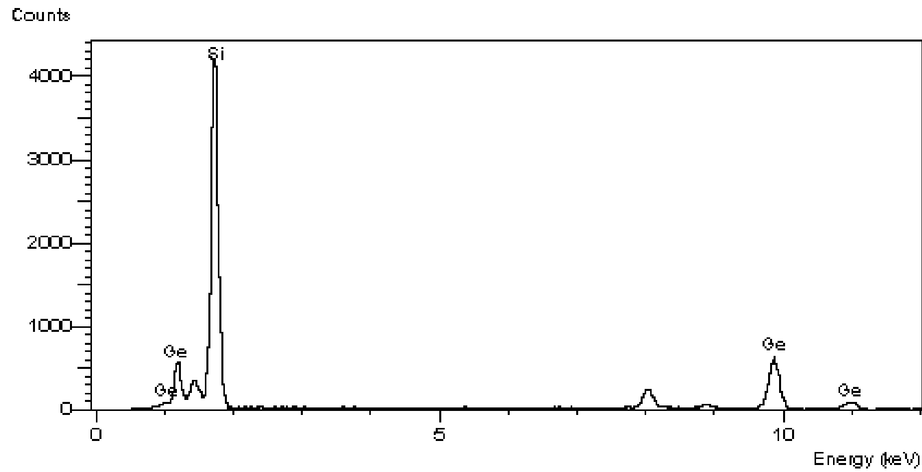


Figure 6.8 Bright field XTEM micrograph using the ***g*** = (004) diffraction condition of sample TGA1. In this diffraction condition strong contrast is shown for different states of relaxation

beam in the TEM was focused into a narrow spot, the characteristic X-ray spectra of the atoms hit by the electron beam was collected, and specialised software is used to identify the elements which are present in the sample at this point. The spectra taken whilst the spot was focused at the centre of a mound and just to the side are shown in Figure 6.9(a) and (b) respectively. As can be seen, at the side of the mound the germanium peak has a high magnitude, indicating the presence of germanium in a reasonably large quantity as is expected for this layer. However, at the centre of the mound the germanium peak is significantly reduced, almost to a negligible level. This clearly shows that the mound is a large cluster of silicon atoms that has somehow formed during the growth of this layer.

The AFM of this sample is shown in Figure 6.10, the 3-D surface plot shows the surface mounds very clearly. These mounds have a significant detail in their shape, having small grooves orientated along the <110> directions. These grooves are the remnants of the crosshatch pattern that would have formed prior to the

(a) Centre of a mound



(b) To the side of a mound

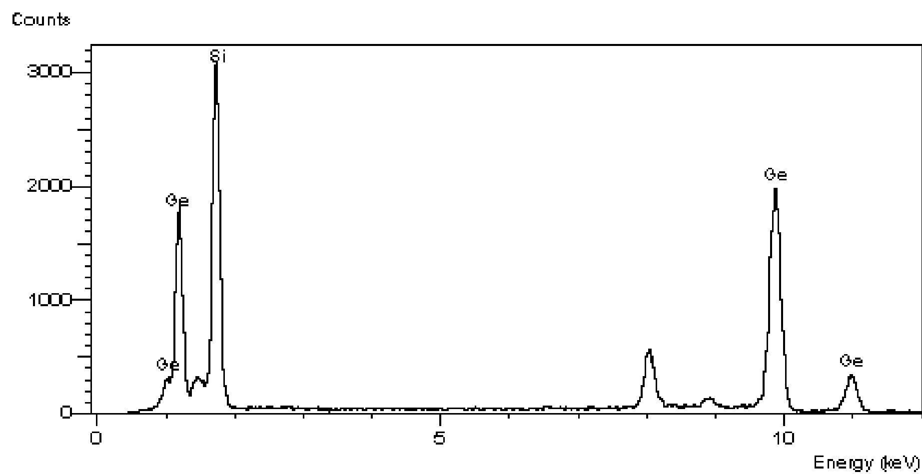


Figure 6.9 X-ray analysis in the TEM of the mounds seen in sample TGA1. (a) shows the spectra taken whilst the spot was focussed at the centre of a mound showing very little presence of germanium. (b) shows the spectra taken to one side of the mound in the same layer, the Ge peaks are much more pronounced, indicating a higher concentration of germanium.

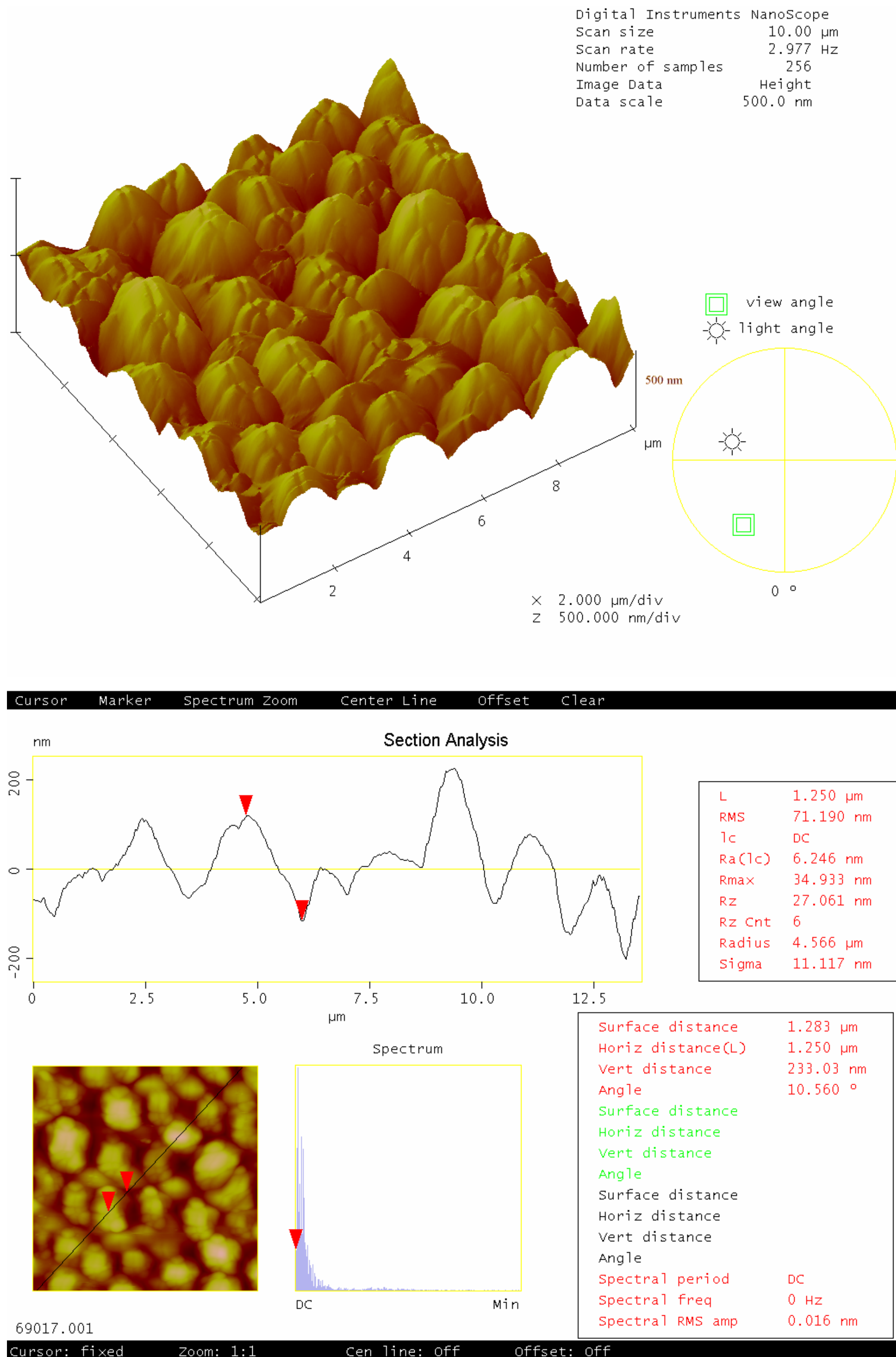


Figure 6.10 AFM 3-D surface plot and line profile of sample TGA1.

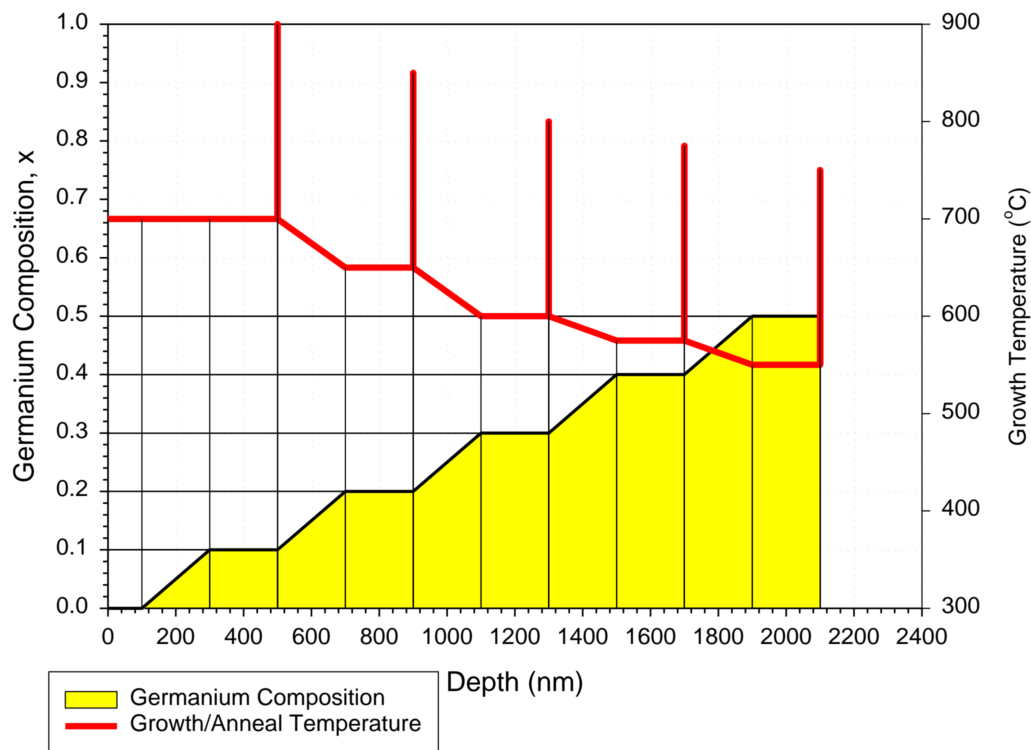
transition to 3-D growth and is now dwarfed by the size of the surface mounds. The line profile of Figure 6.10 shows that strong faceting that has occurred, the height and angle of one of these facets is indicated in the accompanying data and the angle of these facets is approximately ~11° which is the same as the {105} facets. This faceting is an indication that roughening during growth of the upper layers has not been kinetically limited. Instead, the adatoms have been able to arrange themselves to form the energetically favourable facets.

The formation of *germanium* clusters is not uncommon in strained SiGe systems, these islands form in order to minimise the free energy of the system as described in section 2.2.3. However, this strain relieving mechanism does not explain the formation of the silicon cluster seen in this study. It is proposed that the formation of the large silicon clusters between the $x = 0.30$ and 0.40 layers is due to the high temperature anneals after each of the uniform germanium layers. Germanium forms weaker bonds than silicon, and the anneals are at approximately 900 °C, which is close to the melting point of pure bulk germanium (938 °C). At these high anneal temperatures the Ge-Ge bonds will be easily broken, whereas the Si-Si bonds will be largely unaffected at this temperature (silicon having a melting point of 1412 °C). The higher energy of the Si-Si bonds could act as the driving force for the rearrangement of the atoms to form the energetically favoured Si-Si bonds, leading to clustering of the silicon atoms. The range of this ordering is likely to be quite low, so this effect will only be seen in the higher germanium composition layers where the low energy Ge-Ge bonds are more prevalent.

6.4 The Use of Progressively Lower Anneals During the Growth of Terrace Graded Virtual Substrates.

Since it has been shown by the growth of the $x = 0.10$ layers of the terrace graded structure that the relaxation of these layers in the $Si_{0.50}Ge_{0.50}$ virtual substrates is incomplete before subsequent growth, and furthermore that these layers can be forced to relax by an *in-situ* anneal, this technique of terrace grading with an *in-situ* anneal after each layer was deemed an appropriate way forward in producing high quality virtual substrates. However, as shown above the anneal temperature suitable for relaxing the $x = 0.10$ layers (900 °C) is too high for layers with $x > 0.30$, leading to formation of silicon clusters which severely degrade the quality of the virtual substrates. In order that this approach could be successfully applied, it was necessary to reduce the temperature of these anneals in relation to the increased germanium composition. For this reason a second attempt at growing a terrace graded virtual substrate with *in-situ* anneals was attempted where the anneal temperature was reduced at each layer. This wafer (designated **TGA2**) was grown with the same specification as TGA1 with the exception that the annealing temperature was maintained at 200 °C above the growth temperature (which was reduced in a manner similar to TG1 and TGA1). The growth specifications for this sample is shown in Figure 6.11.

The XTEM of this sample is shown in Figure 6.12 using the $\boldsymbol{g} = 220$ diffraction condition. It can be seen that the dislocation structure is superior to the original terrace graded virtual substrate without the anneals (TG1). The dislocations in each layer show remarkable uniformity, indicating that the relaxation has occurred at the same point for each of the terraced layers. The misfit dislocations are long and confined almost entirely to the graded sections of each layer. There is no evidence of



Step	Growth Temperature (°C)	Germanium Composition (%)	Thickness (nm)
0	700	0	100
1	700	0 – 10	200
2	700	10	200
3	900	Anneal	-
4	700 – 650	10 – 20	200
5	650	20	200
6	850	Anneal	-
7	650 – 600	20 – 30	200
8	600	30	200
9	800	Anneal	-
10	600 – 575	30 – 40	200
11	575	40	200
12	775	Anneal	-
13	575 – 550	40 – 50	200
14	550	50	200
15	750	Anneal	-

Figure 6.11 Growth specifications of sample TGA2.

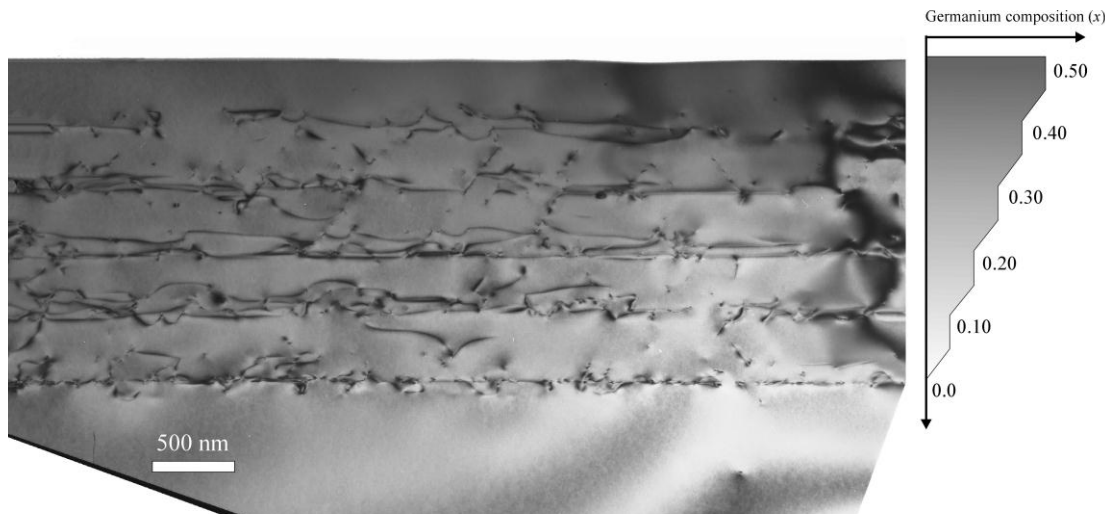


Figure 6.12 Bright field XTEM micrograph of sample TGA2 using the ***g*** = (220) diffraction condition. The dislocations evenly spread are confined to the graded layers.

any dislocation pile-ups that have extended beyond one layer, this is in contrast to TG1 where dislocation pile-ups have extended through several of the layers. This shows that the graded layers have completely decoupled the dislocations networks of each graded layer, so that relaxation of each step occurs independently of the dislocation network in the lower levels. There is no evidence of the silicon clustering found in sample TGA1, where the anneal temperature was proposed to be too high for the higher germanium layers. There is no evidence in the TEM that threading dislocations have penetrated the final capping layer of the structure, indicating a high quality virtual substrate. The full relaxation of each layer can be certain since even in sample TG1, XRD showed that each of the layers had fully relaxed by the end of the growth. The *in-situ* anneals have been shown to relax the $x = 0\ .10$ layer (the hardest layer to relax due to the lower dislocation glide velocity of dislocations of layers with low x, as described in section 2.5.1) so with the addition of these anneals there can be no doubt that each layer has attained full relaxation before the subsequent layers have been grown.

The surface of this sample has been studied in the AFM to ascertain the effect of these anneals on the surface morphology. Figure 6.13 shows a 3-D surface plot of the sample showing a well defined crosshatch pattern with troughs spaced at approximately 2 μm. In between these troughs the surface has a smaller amplitude undulation with a much smaller spacing. A line profile is shown in Figure 6.14 drawn to the same scale as for sample TG1 shown in Figure 4.11. This line profile shows that sample TGA2 has less roughening than both conventional techniques for growing virtual substrates (LG1 and SG1). However, the sample is slightly rougher than terrace graded sample without the anneals (TG1). The RMS roughness of the sample is 4.7 nm, in comparison the RMS roughness of sample TG1 is only 3.1 nm.

It is possible that the annealing temperature for the uppermost layers is still too high, leading to roughening of the surface due to the higher mobility of the surface atoms. If this is the case, by lowering the annealing temperature of the upper layer the RMS roughness of the sample should decrease further, it is even possible that with the enhanced dislocation glide velocity of the high germanium layers, that there is no need for an anneal after the final layer. This is supported by the fact that the top layer of TG1 was shown to be nearly fully relaxed without annealing and without any higher layers to increase the strain. This is in contrast to the $x = 0.10$ sample of the previous section which had the same jump in composition (10% Ge) but was still strained. Another factor making the higher layers easier to relax is the presence of threading dislocations that have already nucleated in the previous layers. This removes the barrier of nucleation, described in section 2.4.1, that needs to be overcome in the case of the first $x = 0.10$ layer.

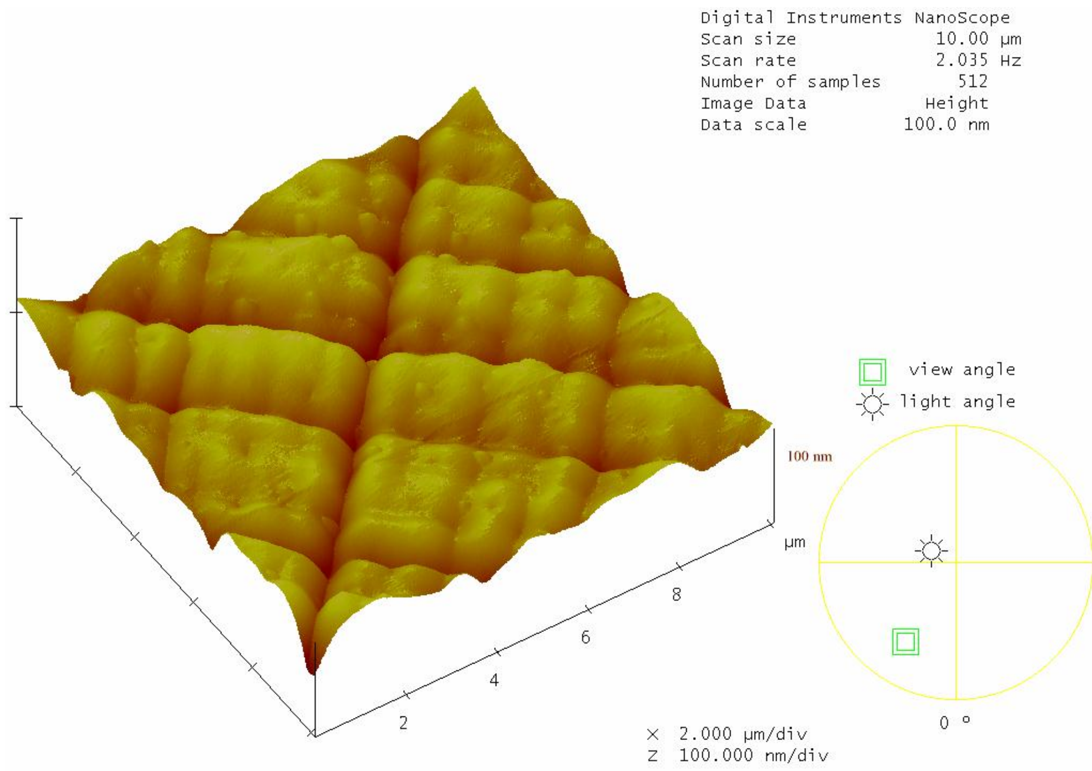


Figure 6.13 AFM 3-D surface plot of sample TGA2

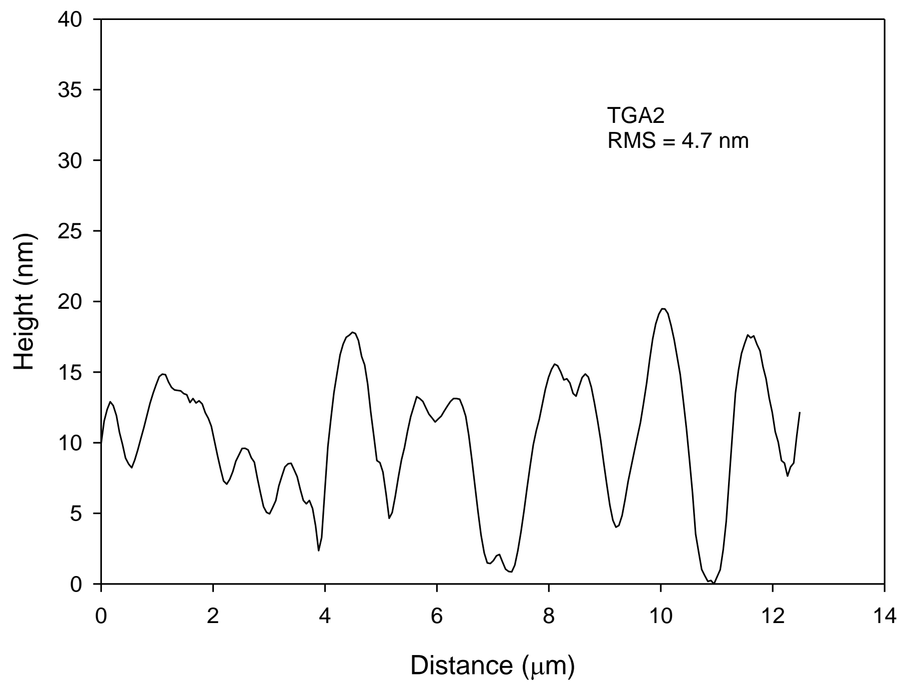


Figure 6.14 AFM line profile of sample TGA2

The threading dislocation density was studied by using a Schimmel defect etch as described in section 4.2.5. The optical micrograph (without DIC) is shown below in Figure 6.15. By comparison with the etch pit densities of Figure 4.12, it is clear straight away that the effect of the *in-situ* anneals is to reduce the density of threading dislocations. The most important feature is that there is now hardly any evidence of dislocation pile-ups that is usually present in virtual substrates, the linearly graded sample (LG1) shown in Figure 4.12(b) being a good example of this. With the pile-ups being so small the *total* threading dislocation density is hardly effected by their presence and is estimated to be 2.0×10^{6} cm^{-2}. According to Fitzgerald *et al*[46] there is a need for a *mobile* threading dislocation density of 10^{5} - 10^{6} cm^{-2} to allow relaxation by the formation of misfit dislocations. But dislocations that have become pinned as part of a pile-up cannot contribution to this

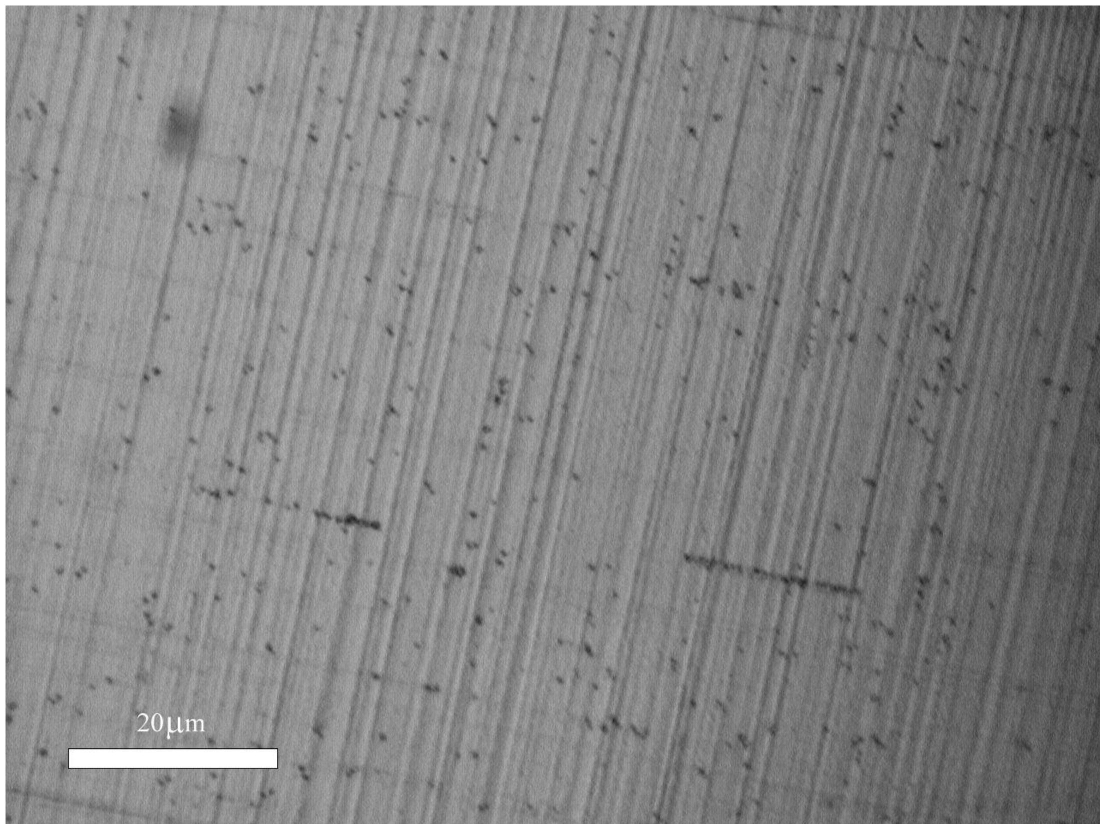


Figure 6.15 Optical micrograph of defect etched sample TGA2.

strain relieving mechanism and so have no useful purpose. It is important therefore that the dislocation pile-ups are reduced to a minimum in order to produce high quality virtual substrates. It is clear that the effect of using a terrace grading profile with a strain relieving anneal after the growth of each uniform layer is to reduce the number of threading dislocations that have become pinned.

6.5 Discussion

It has been shown that the lower layer of the terrace graded virtual substrates have not fully relaxed before the following layers have been grown. This was first hypothesized in chapter 4 due to the high density of dislocations in these lower layers and the over-relaxation determined from XRD and now confirmed from studies of just the first $x = 0.10$ layers. However, it has been shown that an *in-situ* anneal has the effect of fully relieving this residual strain leaving a highly planar surface interrupted with low amplitude, well spaced crosshatch. The RMS roughness was not affected as a result of this anneal. Growth of a full $Si_{0.50}Ge_{0.50}$ terrace graded virtual substrate with a 30 minute anneal at 900 °C after each uniform layer led to a better spaced dislocation network in the lower layers, with fewer dislocations penetrating the uniform layers. However, after the $x = 0.30$ layers large silicon clusters were formed leading to dislocations nucleation and a highly non-planar surface which is completely unsuitable for device applications.

The formation of these clusters was proposed to be due to the annealing temperature being too high for the upper layers. In order to improve on this sample TGA2 was grown with progressively lower annealing temperatures. This had the effect of preventing the formation of the silicon clusters, but retained the superior

dislocation network throughout the sample. The surface of this wafer was found to be rougher than the terrace graded sample without the *in-situ* anneals, although this was still considerably smoother than the virtual substrates grown by either of the conventional methods. The cause of this rougher surface is likely to be the final anneal of the sample after the growth of the $x = 0.50$ uniform layer allowing the surface to corrugate. Further work should include studying the effect of these final anneals to determine if they can be eliminated, or reduced in temperature in order that the surface remains flat. It has been shown that it is necessary to anneal the $x = 0.10$ layers, but no study has been made to discover which, if any, of the subsequent layers need to be annealed. It would be beneficial to reduce the number and duration of the annealing stages in order to reduce contamination. Also, if this technique is used in an industrial setting, the time necessary for the anneals will significantly reduce the throughput and hence increase the total cost.

It is now possible to describe the mechanism by which the terrace graded virtual substrates relax, and the reason for the improved quality over conventional step-graded and linearly graded wafers. As the first graded layer is grown the misfit stress gradually builds up until it becomes energetically favourable for dislocations to form at the interface to relieve the strain energy. However, since there are no pre-existing threading dislocations the formation of misfit dislocations is limited by the rate of nucleation. As the uniform layer grows there is a further increase of strain in the system and the first dislocation will start to nucleate at the surface by half loop formation, and then glide to the graded layer to relieve the misfit stress. It is known that once the uniform layer has been grown there are insufficient dislocations to fully relieve the strain, this being due to the high dislocation nucleation barriers and low glide velocity. A 30 minute *in-situ* anneal at 900 °C ensures that there are sufficient

dislocations formed to relieve the strain completely. This is most likely due to the MFR multiplication mechanism producing long dislocations that are stacked on top of each other on the same glide planes, formed from a few well-spaced multiplication sites. The effect of these stacked dislocations is two-fold. Firstly it acts as a blockade for dislocations gliding in an orthogonal direction. When dislocations encounter this plane of dislocations they are unable to continue, since there are dislocations blocking their glide throughout the layer. The second effect is a large undulation at the surface of the wafer caused by the presence of the piled-up dislocations underneath. This undulation (crosshatch) is due to the additive effect of the atomic steps formed by the presence of each dislocation on the glide plane, followed by the local lowering of the growth rate at these sites. These undulations also tend to pin threading dislocations and lead to a vicious circle which increases both the threading dislocation density and RMS roughness. The final threading dislocation density and RMS roughness increases sharply with terminating germanium composition. This is because as the number of dislocations needed for strain relaxation increases, the number of dislocations sourced from each MFR site increases and consequently there are more dislocations on each atomic plane leading to more blocking and larger surface undulations. In the terrace graded approach, the increase in germanium composition of each step is kept low so that there is only a small density of threading dislocations that become pinned, and additionally keeping the RMS roughness to a minimum.

After the growth and relaxation of the first $x = 0.10$ step the next graded layer is grown. Since there are already threading dislocations that have nucleated in the lower layers present, there is no need to nucleate any more. The second graded layer will start to relax at an earlier stage than the first since there is no need for

further nucleation and the glide velocity is increased at this higher germanium composition. Any dislocations that did become pinned in the first layers will now be able to continue gliding, reducing the need for further nucleation to replace them. As the graded layer grows then the process of dislocation blocking and multiplication repeats, but the sites of these multiplication mechanisms will be independent to the sites that relaxed the first graded layer. Because of this, the surface undulations that are necessarily formed will have no correlation with the undulations formed by the lower layers, and so will not add to their amplitude. This process of relaxation will occur at each step of graded layer followed by uniform layer, the *in-situ* anneal at the end of each uniform layer making sure that the strain is fully relieved before the growth of the subsequent layers. The final density of threading dislocations should be the same as an $x = 0.10$ linearly graded layer, since the dislocations are freed after each terraced step, and the surface RMS roughness should be similarly low.

There are still many factors that can be investigated to try to improve further the quality of virtual substrates grown using this terrace grading technique. Now that the problem of relaxation in the lower layer has been overcome by using an *in-situ* anneal after the uniform layers, reductions in the thickness of the virtual substrate may be possible. The thicknesses of the graded regions and uniform layers have been kept constant at 200 nm throughout the course of this study for the sake of consistency. However, further work should investigate the effect of varying these thicknesses in order to keep the total thickness as low as possible without compromising the quality of the virtual substrate. Since the layers with a higher germanium composition have higher dislocation velocities, it is possible that the thicknesses of these layers may be reduced, and so reducing the thickness of the virtual substrate as a whole.

7 Conclusion

A new compositional grading profile for the growth of SiGe virtual substrates has been proposed, whereby the germanium composition has been increase by a combination of linearly graded and uniform composition layers, forming a bi-layer (or terrace) which is repeated in steps of $x = 0.10$ to the desired final composition. It was proposed, that by interrupting the grading of germanium, by a series of uniform layers, that the dislocation pile-ups associated with the dislocation multiplication mechanisms will be interrupted. Since the dislocation pile-ups are responsible for high densities of threading dislocations and large surface roughness, the reduction of the height of the pile-ups should improve the quality of the virtual substrate. In order to compare this new terrace grading profile with conventional grading techniques (linear grading and step grading), three virtual substrates were grown, SG1, LG1 and TG1 having a step graded, linear graded and terrace graded germanium profile, respectively. XTEM revealed that the terrace graded substrate had far less dislocation pile-ups as was expected, with no threading dislocation seen in the TEM. The surface of each sample was studied using AFM and the roughness of the terrace graded sample was found to be much lower than either the step graded or linearly graded samples, having an RMS roughness of only 3.1 nm. Schimmel etching revealed the threading dislocations at the surface of each sample. The step graded wafer had a very large number of dislocations that was too high to measure. The linearly graded sample showed a much reduced density of threading dislocations, but large clusters of dislocation formed pile-ups due to pinning by the underlying dislocation networks. The terrace graded sample, however, had very few clusters of dislocations, this was

attributed to the pinned dislocations being released at each of the uniform layers, leading to an overall reduction in dislocations that have become pinned.

The quality of electrically active SiGe p-channels grown on this new terrace graded virtual substrate was investigated. Growth of a $Si_{0.50}Ge_{0.50}$ channel on a $Si_{0.75}Ge_{0.25}$ virtual substrate was found to have a poor structural quality with large dislocation pile-ups penetrating the entire epilayer. Large amplitude undulations were clearly seen by XTEM in the channel and on the surface, with dislocations seemingly being nucleated at the cusps of the surface roughening. The poor structural quality was attributed to poor relaxation of the underlying buffer layer. The electrical properties of this sample were very poor, with only low temperature measurements being possible. The growth of a $Si_{0.10}Ge_{0.90}$ channel on a $Si_{0.50}Ge_{0.50}$ virtual substrate was shown to be far more successful. The XTEM revealed a highly planar surface and channel, with no dislocation seen penetrating the upper layers. The dislocation pile-ups were reduced in size, as was seen in the terrace graded sample of chapter 4. AFM revealed that the surface roughness was actually reduced in comparison to the terrace graded sample of chapter 4, which was grown to the same specifications. The RMS roughness was only 2.1 nm. The reduction in surface roughness was attributed to the large capping layer grown as part of the electrically active layers, reducing the effect of the strain field from the underlying dislocation networks. High hole mobilities were demonstrated, despite the electrically active layers not been optimised for electrical properties.

The relaxation mechanism of the terrace graded sample was studied in chapter 6. Two sample were grown consisting of only the $x = 0.10$ layer of the terrace graded profile. One of the samples was subject to a 900 °C *in-situ* anneal for 30 minutes, the other was not. XTEM of the two structures showed that the misfit

dislocation networks (needed for strain relief) were only *just* starting to form in the un-annealed sample, indicating a poor state of relaxation. In contrast, the *in-situ* annealed sample was found to have a well developed network of misfit dislocations, indicating a much higher degree of relaxation. The un-annealed sample was subjected to an *ex-situ* anneal at 900 °C under a N_2 atmosphere for 30 minutes. The XTEM showed that this encouraged the formation of dislocations, but these were seen to have been pushed deep into the substrate. AFM of the sample showed that the un-annealed sample and *in-situ* annealed sample both had a similar surface RMS roughness of < 1 nm. The *ex-situ* annealed sample, however, showed a higher degree of roughening, with an RMS roughness of 2.3 nm. The differences between the *in-situ* and *ex-situ* anneals are not fully understood, but was thought that it may be due to the different temperature ramp rates for the two anneals, or the effect of the surface oxide formed on the *ex-situ* annealed sample. XRD showed that the *in-situ* annealed sample was indeed in a higher state of relaxation (90%) in comparison to the un-annealed sample. Due to the limits of the XRD technique, the exact state of relaxation of the un-annealed sample was not determined. It was apparent, however, that this sample was in a higher state of strain than the *in-situ* annealed sample.

The effect of an *in-situ* anneal during growth of a terrace graded virtual substrate was investigated. It was found that by annealing at 900 °C after the growth of each of the uniform layers, that the dislocation structure of the lower layers was improved, having less dislocation pile-ups in the uniform layers. However after the $x = 0.30$ uniform layer large silicon clusters were seen, with a height of approximately 200 nm. These clusters caused a large disruption of the surface of the sample, leaving the virtual substrate completely unsuitable for MOS device applications. It is not understood why the silicon clusters have formed, but it was

proposed to be due to a surface diffusion effect in the absence of strain. A final terrace graded substrate was grown with an *in-situ* anneal after each uniform layer, only this time, the temperature of the anneal was reduced with increasing germanium concentration. The XTEM of this sample showed that the dislocation pile-ups had been reduced and were confined to the individual graded layers, the uniform layers successfully prevented the pile-ups penetrating the higher layers. The dislocation density in the XTEM was seen to be extremely uniform, indicating that the relaxation has been able to keep up with the growth. AFM revealed that the surface RMS roughness was slightly higher than the sample TG1 grown without the *in-situ* anneals. This was unexpected, and has been attributed to the anneal temperature of the upper layers being too high. The density of threading dislocation, revealed by Schimmel etching was shown to be much reduced, and most importantly, the pile-ups of trapped dislocations were almost eliminated.

In conclusion, a method of grading the germanium composition, by a repeated sequence of linearly graded and uniform layers, has been shown to produce a high quality SiGe virtual substrates. Threading dislocation densities have been reduced compared to conventional grading techniques of the same thickness and RMS roughness reduced to ~ 3 nm, less than half that of the conventional techniques. The growth of a $Si_{0.10}Ge_{0.90}$ channel on a $Si_{0.50}Ge_{0.50}$ virtual substrate has been demonstrated with highly planar interfaces and high hole mobilities. It has further been shown, that the introduction of a high temperature anneal after the uniform layers ensures that the relaxation keeps up with growth, leading to a much reduced density of threading dislocations that have become pinned at dislocation pile-ups.

7.1 Further Work

There is much further work that can be carried out to understand and improve this new grading technique for SiGe Virtual substrates. The use of *in-situ* anneals has been shown to greatly increase the relaxation of each graded layer, but is time consuming and needs to be reduced to a minimum. The use of an annealing step after the first graded layer is clearly necessary from the relaxation studies of section 6.2, but once dislocations have been nucleated it is not clear whether further annealing steps are needed. This needs to be studied in order that the most efficient growth schedule can be determined, especially since it is believed that the anneals of the higher germanium composition layers is detrimental to the RMS roughness of the surface. A study of the effect of different temperature ramp rates needs to be carried out in order that the mechanism of the strain relaxation can be understood. It was suggested in section 6.2.1 that a fast temperature ramp failed to force high relaxation in the $x = 0.10$ layer and produced a rough surface. However, since this high temperature ramp was carried out *ex-situ* under a N_2 atmosphere, with the inevitable native oxide on the surface, it is unclear whether the high ramp rate is the cause of the poor relaxation. This is academic in the case of growth by MBE, since only a slow ramp rate is possible. However, in other systems (notably LPCVD) the temperature can be changed very rapidly and it would be beneficial to raise the temperature as quickly as possible in order to reduce the length of the annealing step. Also, the anneals in this study were carried out at 900 °C, but it may be beneficial to increase this temperature in order that the annealing time may be reduced.

The grading rate in this study was kept constant in order to control the many variables that may affect the quality of the virtual substrate. It was suggested in section 2.5.2 that a high initial grading rate may be beneficial in order to encourage

dislocation formation in the lower germanium layers (due to the lower glide velocity). The effect of increasing the initial grading rate needs to be studied if the most efficient relaxation of the virtual substrate is to be found. One interesting idea is to *over-grade* the linearly graded layers to a higher composition than the following uniform layer. This has been carried out by Mooney *et al*[52] in step graded buffer layers. The reason behind this over-grading is that there always exists a thin strained layer, of height h^*, in which dislocation formation is energetically unfavoured.[46] Consequently, the upper limit to the relaxation of a graded virtual substrate is ~ 95% (as seen in the final layers in this study). If the next uniform layer is reduced in germanium composition in such a way that its *fully* relaxed lattice parameter is matched to the ~ 95% relaxed lattice parameter of the preceding graded layer, then the uniform layer should be able to attain full relaxation. Since the uniform layers will have no residual strain, there will be no driving force for dislocations in these layers, making them more effective at separating the dislocations in the linearly graded regions.

The effectiveness of the terrace grading profile should not be limited to growth by MBE. Other growth techniques should be studied using this grading profile in order that this technique can be carried over to industrial standard (CVD) growth systems. The higher growth rates possible in these systems (~ 5 nms^{-1} by LEPCVD) will allow for thicker layers, and lower grading rates to be studied in order that the optimum conditions can be found. MBE is limited by low growth rates (~ 1 $Ås^{-1}$) and limited source material, to producing relatively thin virtual substrates. Although it was the intention of this work to keep the thickness of the layers low, the ability to fully characterise the effectiveness of this new grading technique would be invaluable.

The incorporation of other techniques currently been studied for the production of high quality virtual substrates with the terrace graded profile could give even further improvements. The CMP step employed by Currie *et al*[71] has been shown to release pinned threading dislocations and reduce the surface roughness. It would be interesting to see if such a CMP step might further increase the quality of the terrace grading profile. The most likely point to remove the sample for the polishing stage is after the growth of the uniform layers (if an anneal is used, the polishing step should be after this). The growth of the following linearly graded layers will then be able to proceed on a flat surface, free from crosshatch. Another technique that has shown a lot of promise is that of using surfactant mediated growth. Liu *et al*[76] has reduced the threading dislocation density to as low as 10^4 cm^{-2} on a linearly graded layer with the same specifications as the sample LG1, used for comparison in section 4.1.1 of this work. Since both terrace grading and Sb surfactants have improved the quality of SiGe virtual substrates, it would be most interesting to incorporate both techniques

References

[1] Paul D. J. Advanced Materials **11** (1999) 191

[2] Landolt-Börnstein. *Numerical Data and Functional Relationships in Science and technology,* New Series, Group III (Springer-Verlag, Berlin) vol. 17a (1982) and vol. 22a (1989)

[3] Kittel C. *Introduction to Solid State Physics* 7th Edition, (J. Wiley, New York, 1996)

[4] Parker E. H. C and Whall T. E. Solid-State Electronics **43** (1999) 1497

[5] Lang D. V, People R, Bean J. C and Sergent A. M. Appl. Phys. Lett **47** (1985) 1333

[6] Van de Walle C. G and Martin R. M. Phys Rev B **34** (8) 5621 (1986)

[7] Schäffler F. Semicond. Sci. Technol. **12** (1997) 1515

[8] Basaran E., Kubiak R. A., Whall T. E., and Parker E. H. C., Appl. Phys. Lett **64** (1994) 3470

[9] Xie Y. H, Monroe D, Fitzgerald E. A, Silverman P. J, Thiel F. A, and Watson G. P. Appl. Phys. Lett. **63** (1993) 2263

[10] Hock G, Hackbarth T, Erben U, Kohn E, and Konig U. Electron Lett. **34** (1998) 1888

[11] Sugii N, Nakagawa K, Kimura Y, Yamaguchi S and Miyao M. Semicond. Sci. Technol **13** (1998) A140

[12] Ismail K, Nelson S. F, Chu J. O and Meyerson B. S. Appli. Phys. Lett. **63** (1993) 660

[13] EE Times August 16, 2002

[14] M. T. Currie, C. W. Leitz, Langdo T. A, G. Taraschi and E. A. Fitzgerald J.Vac.Sci.Technol.B **19** (2001) 2268

[15] Leitz C. W, Currie M. T, Lee M. L, Cheng Z.-Y, Antoniadis D. A, and Fitzgerald E. A. J. Appl. Phys. **92**, (2002) 3745

[16] Windish D and Becker P. Phys. Status. Solidi A (Germany) vol. **118** (1990) p.379

[17] Baker J. F. C and Hart M. Acta Crystallogr. A (Denmark) vol. **31** (1975) p.364

[18] Oleinski R. W and Abbaschian G. J *Bull. Alloy Phase Diagrams* (USA) vol. **5** (1984) p.180

[19] Dismukes J. P, Ekstrom R. J and Paff R. J. J. Phys. Chem. Vol. **68** (1964) 3021

[20] Kasper E, Schuh A, Bauer G, Höllander B and Kibbel H. J. Cryst. Growth **157** (1995) 68

[21] Nijs J. F. A. *Advanced Silicon & Semiconducting Silicon-Based Materials and Devices*. (IOP Publishing, 1993) p. 141

[22] www.ASM.com

[23] Meyerson B. S. Appl. Phys. Lett. **48** (1986) 797

[24] Reif R. J. Vac. Sci. Technol. **A2**, (1984) 429

[25] Gravesteijn D. J. Van de Walle G. F. A, Pruijmboom A and Van Gorkum A. A. Mat. Res. Soc. Symp. Proc. **220** (1991) 3

[26] Rosenblad C, von Känel H, Kummer M, Dommann A, Müller E. Appl. Phys. Lett **76** (2000) 427

[27] Regolini J. L, Bensahel D, Scheid E. and Mercier J. Appl. Phys. Lett **54** (1989) 658

[28] Kasper E. and Bean J. C. *Silicon Molecular Beam Epitaxy* (Boca Raton, CRC Press, 1988)

[29] Tabor D. *Gases, Liquids and Solids (and other states of matter)* 3rd Edition.(Cambridge University Press, 1991) p. 100

[30] Kern. W. *Handbook of semiconductor wafer cleaning technology.* (Noyes Publications, 1992)

[31] Attard G and Barnes C. *Oxford Chemistry Primers: Surfaces*. (Oxford University Press,1998)

[32] Frank F.C. and Van der Merwe, Proc. R. Soc. London Ser. A **198** (1949) 205

[33] Volmer M and Weber A, Z. Phys. Chem. **119** (1926) 277

[34] Stranski J. N. and Krastanov L. Ber. Akad. Wiss. Wien **146** (1938) 797

[35] Kasper E. and Jorke H. J. Vac. Sci. Technol. A **10** (1992) 1927

[36] Cottrell A. H *Theory of Crystal Dislocations*.(Blackie and Son Ltd, 1964)

[37] Peierls R. E. Proc. Phys. Soc. Vol. **52** (1940) p.23

[38] Hull R. and Bean J. C. J. Vac. Sci. Technol. A **7** (1989) 2580

[39] Hirthe J. P and Lothe J. *Theory of Dislocations* 2nd Edition(John Wiley, New York. 1982)

[40] Hull R, Bean J. C, Peticolas L. J, Bahnck D, Weir B. E and Feldman L. C. Appl. Phys. Lett **61** (1992) 2802

[41] Frank F. C, Van der Merwe J. H. Proc. R. Soc. Lond. A (UK) vol.**198** and vol. **200** (1949)

[42] Matthews J. W and Blakeslee A. F. J. Cryst. Growth **27**, (1974) 118

[43] People R. and Bean J. C. Appl. Phys. Lett **47** (3) (1985) 322

[44] Chidambarrao D, Srinivasan G. R, Cunningham B and Murthy C. S. Appl. Phys. Lett **57** (1990) 1001

[45] Willis J. R, Jain S. C and Bullough R. Phil. Mag. A **62**, (1990) 115

[46] Fitzgerald E. A, Currie M. T, Samavedam S. B, Langdo T. A, Taraschi G, Yang V, Leitz C. W and Bulsara M. T. Phys. Stat. Sol (a) **171**, (1999) 227

[47] Hull R. *Properties of Strained and Relaxed Silicon-Germanium*. Emis Datareviews No. **12** (IEE, INSPEC, London. 1994) p.28

[48] Hagen W and Strunk H. Appl. Phys. Lett. **17** (1978) 85

[49] Eaglesham D. J, Kvam E. P, Maher D. M, Humphreys C. J and Bean J. C. Phil Mag. A **59** (1989) 1059

[50] LeGoues F.K, Meyerson B. S, Morar J. F and Kirchner P. D. J. Appl. Phys. **71** (1992) 4230

[51] Tersoff J. LeGoues F. K. Phys Rev. Lett. **72** (1994) 3570

[52] Mooney P.M. Materials Science and Engineering, **R17** (1996) 105

[53] Cullis A. G, Robbins D. J, Barnett S. J and Pidduck A. J. J. Vac. Sci. Technol. A **12** (1994) 1924

[54] Jesson D. E, Chen K. M, and Pennycook S. J. MRS Bulletin (April, 1996) p. 31

[55] Cullis A. G. MRS Bulletin (April, 1996) p. 21

[56] Fitzgerald E. A, Xie Y. –H, Monroe D, Silverman P. J, Kuo J. M, Kortan A. R, Thiel F. A and Weir B. E. J. Vac. Sci. technol. B **10** (1992) 1807

[57] Rowell N. L. and Houghton D. C, Mater. Res. Soc. Symp. Proc. **198**, (1990) 509

[58] Fukuda Y, Kohama Y, Seki M and Ohmachi Y. Jap. J. Appl. Phys. **28** (1989) L19

[59] Tersoff J. Appl. Phys. Lett. **62** (1993) 693

[60] Herzog H. –J, Hackbarth T, Höck G, Zeuner M and König U. Thin Solid Films **380** (2000) 36

[61] Dodson B. W and Tsao J. Y, Appl. Phys. Lett. **51** (1987) 1325

[62] Tuppen C. G and Gibblings C. J. Appl. Phys. Lett. **68** (1990) 1526

[63] Houghton D. C. Appl. Phys. Lett. **57** (1990) 2124

[64] Stach E. A, Hull R, Tromp R. M, Ross F. M, Reuter M. C and Bean J. C. Phil. Mag. A **80** (2000) 2159

[65] Peach M. O and Koehler J. S. Phys. Rev **80** (1950) 436

[66] Mooney P. M. Mats.Sci. Eng, **R17** (1996) 105

[67] Mooney P. M Jordan-Sweet J. L, Chu J. O and LeGoues F. K. Appl. Phys. Lett. **66** (1995) 3642

[68] Dutartre D. Warren P, Provenier F, Chollet F and Périot A. J. Vac. Sci. Technol. A **12** (1994) 1009

[69] Lutz M. A, Feenstra R. M, LeGoues F. K, Mooney P. M and Chu J. O. Appl. Phys. Lett. **66** (1995) 724

[70] Olsen S. H, O'Neill A. G, Norris D. J, Cullis A. G, Woods N. J, Zhang J, Fobelets K and Kemhadjian H. A. Semicond. Sci. Technol. **17** (2002) 655

[71] Currie M.T, Samavedam S. B, Langdo T. A, Leitz C. W and Fitzgerald E. A. Appl. Phys. Lett **72** (1998) 1718

[72] Kissinger G, Morgenstern T, Morgenster G and Richter H. Appl. Phys. Lett **66** (1995) 2083

[73] Kasper E, Lyutovich K, Bauer M, Oehme M. Thin Solid Films **336** (1998) 319

[74] Bauer M, Lyutovich K, Oehme M, Kasper E, Herzog H. –J, Ernst F. Thin Solid Films **369** (2000) 152

[75] Gaiduk P. I, Nylandsted Larsen A, Lundsgaard Hansen J. Thin Solid Films **367** (2000) 120

[76] Liu J. L, Wang K. L, Moore C. D, Goorsky M. S, Borca-Tascuic T and Chen G. Thin Solid Films **369** (2000) 121

[77] Fitzgerald E. A, Watson G. P, Proano R. E and Ast D. G.J. Appl. Phys. **65** (1989) 2220

[78] Noble D. B, Hoyt J. L, King C. A and Gibbons J. F. Appl. Phys. Lett **56** (1990) 51

[79] Luryi S and Suhir E. Appl. Phys. Lett. **49** (1986) 140

[80] Hammond R, Phillips P. J, Whall T. E and Parker E. H. C. Appl. Phys. Lett **71** (1997) 2517

[81] Wohl G, Kasper E. Hackbarth T, Kibbel H, Klose M and Ernst F. Journal of Material Science: Materials in Electronics **12** (2001) 235

[82] Gasiorowicz, Stephen. *Quantum Physics*, 2nd Edition (New York: Wiley, 1995)

[83] Hecht, Eugene.Optics, *3rd Edition* (Reading, MA: Addison-Wesley, 1998)

[84] Hartmann J. M, Gallas B, Zhang J and Harris J. J. Semicond. Sci. Technol. **15** (2000) 370

[85] Dowsett M. J, Morris R. J. H, Chou P. –F, Corcoran S. F, Kheyrandish H, Cooke G. A, Maul J. L and Patel S. B. *Secondary Ion Mass Spectroscopy SIMS XIII* (Japan, 2001) Proceed 13th Int. Conf. SIMS

[86] van der Pauw L.J. Phillips Res Repts **13** (1958) 1-9

[87] Mooney P. M, Jordan-Sweet J. L, Noyan I. C and Kaldor S. K. Appl. Phys. Lett. **74** (1999) 726

[88] Mooney P.M LeGoues F. K, Chu J. O and Nelson S. F. Appl. Phys. Lett. **62** (1993) 3464

[89] Stach E. A, Schwarz K. W, Hull R, Ross F. M and Tromp R. M. Phys. Rev. Lett. **84** (2000) 947

[90] Mizushima K, Shiona I, Yamaguchi K and Muraki N. *New Group IV (Si-Ge-C) Semiconductors: Control of Properties and Applications to Ultrahigh Speed and Opto-Electronic Devices*. (Sendai, Japan, 2001)

[91] Grasby T. G, Parry C. P, Phillips P. J, McGregor B. M, Morris R. J. H, Hammond R. Knights A. P and Coleman P. G. Appl. Phys. Lett. **74** (1999) 1848

[92] Morris R. J. H. *Ph.D Thesis*, University of Warwick, to be published (2003)

[93] Ismail K. Solid State Phen. 47 – 48 (1996) 409

[94] Feenstra R. M, Lutz M. A. Stern F, Ismail K, Mooney P. M, LeGoues F. K, Stanis C, Chu J. O and Meyerson B. S. J. Vac. Sci Technol. B **13** (1995) 1608

[95] Stach E. A, Hull R, Tromp R. M, Reuter M. C, Copel M, LeGoues F. K and Bean J. C. J. Appl. Phys. **83** (1998) 1931