Interfacial roughness of $Si_{1-x}Ge_x/Si$ multilayer structures on Si(111) probed by x-ray scattering

1997 J. Phys.: Condens. Matter 9 4521

(http://iopscience.iop.org/0953-8984/9/22/005)

J. Phys.: Condens. Matter **9** (1997) 4521–4533. Printed in the UK PII: S0953-8984(97)78651-X

Interfacial roughness of $Si_{1-x}Ge_x/Si$ multilayer structures on Si(111) probed by x-ray scattering

P M Reimer†§, J-H Li†‖, Y Yamaguchi†, O Sakata†, H Hashizume†, N Usami‡ and Y Shiraki‡

† Materials and Structures Laboratory, Tokyo Institute of Technology, Nagatsuta, Midori, Yokohama 226, Japan

‡ Research Centre for Advanced Science and Technology, University of Tokyo, Komaba, Meguro, Tokyo 153, Japan

Received 14 October 1996, in final form 18 February 1997

Abstract. Interface structures in $Si_{0.9}Ge_{0.1}/Si$ heterostructures and superlattices grown on vicinal Si(111) substrates are studied using x-ray reflectivity and diffuse scattering. A set of diffuse scattering data collected from a heterostructure sample were simultaneously fitted to a distorted-wave Born approximation calculation, giving a r.m.s. roughness in good agreement with those from reflectivity data, except for the top oxide layer, with a large out-of-plane correlation length of 460 Å. A modified correlation function was introduced to account for the differential correlation between long- and short-wavelength roughnesses. Strong diffuse peaks are observed in rocking scans on a ten-period $Si_{0.9}Ge_{0.1}/Si$ superlattice, revealing the structure of the miscut substrate to be highly replicated through subsequent overlayer growth of the Si and SiGe layers. The structure consists of surface terraces separated by bunched steps 41 Å high. The correlation of successive layers is slightly misoriented from the surface normal, as is evident from the splitting in the specular and diffuse scattering peaks.

1. Introduction

The roughnesses of $Si_{1-x}Ge_x/Si$ multilayer structures, and the roughness correlations from layer to layer have attracted much attention in the last decade. The system holds promise as a direct-band-gap semiconductor fabricated with well-developed silicon technology, and controlling the roughness at interfaces is crucial for device performance. The system shows some unusual roughening behaviour: while the tops of the SiGe alloy layers are rough, forming islands for high Ge content, some measurements find that this roughness is not replicated in layers of Si grown on top of the alloy [1], but successive SiGe layers show roughness that can be correlated with that of the underlying SiGe layers. Elastic theories have recently been put forward which propose that elastic stress in the intervening Si layer can 'seed' islands in subsequent SiGe layers [2].

X-ray scattering is a powerful technique for probing buried interface structures with atomic resolution. The recent development of scattering theory for rough surfaces [3], and its extension to multilayer interfaces [4, 5], allow correlated structures to be explored in both the in-plane and out-of-plane directions.

§ Present address: Physics Department, Goshen College, Goshen, IN 46526, USA.

‖ Present address: Institute of Physics, Chinese Academy of Sciences, Beijing 100080, People's Republic of China.

0953-8984/97/224521+13$19.50

Much work has focused on SiGe/Si heterostructures grown on Si(001) substrates [6–11]. Recent studies have shown that the interfacial roughness correlations between interfaces in a SiGe/Si multilayer on a vicinal Si(001) substrate can be strongly misaligned from the surface normal [12–14].

In the present paper we investigate roughness structures in $Si_{0.9}Ge_{0.1}$/Si grown on vicinal Si(111) substrates using grazing-angle x-ray scattering techniques. We will propose a useful modification to the correlation functions used to analyse low-angle diffuse scattering, as well as showing the appropriate geometrical corrections, which to our knowledge have not been previously made explicit.

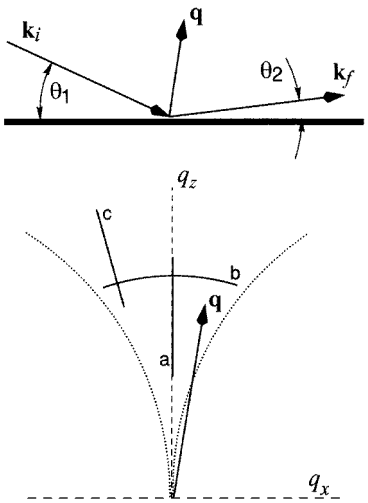

Figure 1. A schematic diagram of the scattering set-up. Top: the incident wave-vector $\boldsymbol{k}_1$ makes an angle θ_1 with the sample surface. The outgoing wave-vector $\boldsymbol{k}_2$ makes an angle θ_2 with the same surface. The scattering angle is $2\theta = \theta_1 + \theta_2$. Bottom: the momentum transfer is $\boldsymbol{q} = \boldsymbol{k}_2 - \boldsymbol{k}_1$, with components q_x in the plane of the surface film, and q_z perpendicular to the surface. Requiring that the wavevectors lie above the surface limits the region of q-space that can be probed to the area above the dotted line in which $\theta_1 > 0$ and $\theta_2 > 0$. Three types of scan are shown: (a) specular rod scans where $\theta_1 = 2\theta/2$, (b) rocking scans where 2θ is held constant while θ_1 varies, and (c) offset scans where $\theta_1 = 2\theta/2 + \Delta$ with Δ a constant.

2. Experiment

In this study three kinds of scan are made which trace different cuts in momentum-transfer space. In terms of the angles defined in figure 1, these are: (a) specular rod scans along the line $\theta_1 = \theta_2$, (b) rocking scans where the scattering angle 2θ ($=\theta_1 + \theta_2$) is held constant while θ_1 varies, and (c) specular offset scans, nearly specular scan, but with a constant offset $\Delta = \theta_1 - 2\theta/2$. Measurements were made on two separate samples: sample A is a multilayer with a few SiGe and Si layers; sample B is a short-period superlattice ('short' is relative to the layer spacing of sample A).

The multilayer sample, A, consists of two $Si_{0.9}Ge_{0.1}$ layers, nominally 60 Å thick, separated by 200 Å of Si, and capped with 250 Å of Si, on a Si(111) substrate. The heterostructure was grown in a gas-source molecular beam epitaxy (MBE) facility using growth rates of 2.3 Å s^{-1} for Si and 0.7 Å s^{-1} for SiGe, on the substrate held at 740 °C.

Post-growth x-ray investigation found that the substrate was miscut by 0.44° in a direction which lies 55° from the [11$\bar{2}$] direction.

X-ray experiments were performed at the Photon Factory synchrotron radiation source, KEK, Tsukuba, Japan, using beamline 20B, on a bending magnet of the 2.5 GeV positron storage ring. The beam is monochromatized with a double-bounce Si(111) monochromator, but there is no focusing mirror, so the resolution is high. Measurements were made at $\lambda = 1.49$ Å, with the sample mounted on the 'BIGDIFF' diffractometer housed inside a large evacuated chamber to reduce air scattering [15]. The crystal components are in the $(+,-,+)$ arrangement. The incident beam was limited by slits to 50 μm in the plane of diffraction, and 3 mm out of the plane. A 100 μm $\times$ 3 mm slit was placed in front of the detector (an NaI scintillation counter). Resolution at the specular rod was $\Delta q_x/q = 2.1 \times 10^{-4}$ in the plane and $\Delta q_z = 1.8 \times 10^{-3}$ Å^{-1} out of the plane.

The superlattice sample, B, has 10 bilayers, each made of Si of thickness 50 Å and $Si_{0.9}Ge_{0.1}$ of thickness 50 Å, grown by solid-source MBE on a Si(111) substrate with its surface off-oriented by 0.63° towards a direction 20° from [11$\bar{2}$]. First a homoepitaxial layer of Si of approximately 800 Å was grown at 800 °C. Ten layers of SiGe and Si were deposited at 1.0 Å s^{-1}, with the substrate temperature 500 °C.

X-ray experiments on the superlattice were performed on a diffractometer using Cu Kα_1 radiation from an 18 kW rotating-anode source. An asymmetric channel-cut Si(111) crystal was used as a monochromator. With a miscut of 12° optimized for Cu Kα_1 radiation, this monochromator produces a beam 3.4 times more intense than a conventional symmetric channel-cut crystal. With a 50 μm slit in front of the sample and another before the detector, the resulting resolution of this configuration is $\Delta q_x/q = 8.7 \times 10^{-4}$ in the plane and $\Delta q_z = 1.0 \times 10^{-3}$ Å^{-1} out of the plane. Both of the diffractometers described here had relaxed collimation perpendicular to the plane of diffraction, i.e. they integrated over the q_y-direction in reciprocal space.

Intensity ranges of more than six decades were covered, using a varying number of aluminium attenuator foils at both the synchrotron and the rotating-anode x-ray source. Varying loads on the latter source caused significant changes in the optical alignment of the diffractometer, probably due to thermal changes of the rotating-target diameter. So, the generator was run at a constant output of 14.4 kW.

3. Data analysis

We use the distorted-wave Born approximation (DWBA) to calculate a scattering cross-section for a model layer system, and then perform a fit to experiment using Powell's method [16]. Sinha *et al* [3] applied the DWBA to the scattering from a single interface near its critical angle for total external x-ray reflection. Holy *et al* [4, 5] have recently extended the theory to the general case of many layers. In table 1 we summarize the DWBA cross-section for elastic scattering, and the associated terms, using the notation of Schlomka *et al* [6].

3.1. The modified correlation function

Most importantly, the diffuse cross-section is sensitive to the Fourier transform $\tilde{C}_{jk}(q_x)$ of the height–height correlation function $C_{jk}(X)$ between layer j and layer k for points separated by an in-plane separation X:

$$C_{jk}(X) = \left\langle z_j(\boldsymbol{r}) z_k(\boldsymbol{r} + X) \right\rangle_{\boldsymbol{r}} \tag{1}$$

Table 1. A summary of the factors entering the DWBA expression for the scattering cross-section. j and k index the interfaces, and n_j is the (complex) refractive index for x-rays of layer j. τ is the mean transverse replication cut-off length. A single interface is characterized by its r.m.s. roughness, σ_j, Hurst parameter, h_j, and cut-off length, ξ_j. $\mathcal{G}$ is a factor which accounts for the total illuminated sample area. G_j^m is shorthand for various combinations of the Fresnel specular reflection R_j- and transmission T_j-coefficients, indexed to either the incoming wave-vector, $\boldsymbol{k}_1$, or the scattered wave-vector $\boldsymbol{k}_2$.

DWBA

$$\left[\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega}\right]_{\mathrm{diff}} = \frac{\mathcal{G}k_1^2}{16\pi^2}\sum_{j,k=1}^{N} n_{jk}\tilde{A}_{jk}^2\tilde{C}_{jk}(q_x)\sum_{m,n=0}^{3}\tilde{G}_j^m\tilde{G}_k^n \exp\left\{-\frac{1}{2}\left[(q_{z,j}^m\sigma_j)^2 + (q_{z,k}^n\sigma_k)^2\right]\right\}$$

Refraction indices	$n_{jk} = (n_j^2 - n_{j+1}^2)(n_k^2 - n_{k+1}^2)^*$
Transverse replication	$\tilde{A}_{jk}^2(q_x) = (1-\delta_{jk})\exp\left[-2(\tau/2\pi)^2 q_x^2\right] + \delta_{jk}$
Correlations	$\tilde{C}_{jk}(q_x) = \int \mathrm{d}X\ \mathrm{e}^{-\mathrm{i}q_x X} C_{jk}(X)$
	$C_{jk}(X) = \frac{1}{2}\left[\frac{\sigma_k}{\sigma_j}C_j(X) + \frac{\sigma_j}{\sigma_k}C_k(X)\right]\exp(-\lvert z_j - z_k\rvert/\xi_\perp)$
	$C_j(X) = C_{jj}(X) = \sigma_j^2 \exp\left\{-(X/\xi_j)^{2h_j}\right\}$
Fresnel coefficients	$\tilde{G}_j^m = G_j^m \exp(-\mathrm{i}q_{z,j}^m z_j)$
	$G_j^0 = T_{i,j+1}T_{f,j+1};\ G_j^1 = T_{i,j+1}R_{f,j+1};\ G_j^2 = R_{i,j+1}T_{f,j+1};\ G_j^3 = R_{i,j+1}R_{f,j+1}$
Momentum transfers	$\boldsymbol{q}_j^0 = -\boldsymbol{q}_j^3 = \boldsymbol{k}_{i,j+1} + \boldsymbol{k}_{f,j+1};\ \boldsymbol{q}_j^1 = -\boldsymbol{q}_j^2 = \boldsymbol{k}_{i,j+1} - \boldsymbol{k}_{f,j+1}$

where $z_j(\boldsymbol{r})$ denotes the local height of interface j at an in-plane position $\boldsymbol{r}$ (i.e. $\boldsymbol{r}$ has no z-component), and according to the convolution theorem, equation (1) implies $\tilde{C}_{jk}(q_x) = \tilde{z}_j(q_x)\tilde{z}_k(q_x)$. Here we consider only correlations as a function of X which lies in the diffracting plane, since our diffractometers average over the q_y-direction in reciprocal space.

At very large length scales, the r.m.s. roughness approaches a limiting value σ_j. Schlomka *et al* [6] use for the correlations between interfaces

$$C_{jk}(\boldsymbol{X}) = \frac{1}{2}\left[\frac{\sigma_k}{\sigma_j}C_j(X) + \frac{\sigma_j}{\sigma_k}C_k(X)\right]\exp(-|z_j - z_k|/\xi_\perp). \tag{2}$$

Here, $\xi_\perp$ is the mean perpendicular correlation length connecting fluctuations at interface j with those at interface k according to the average distance $|z_j - z_k|$ between the two.

While equation (2) approaches an appropriate limit for completely correlated interfaces (as $\xi_\perp \to \infty$), it makes no distinction between large and small fluctuations. Large features should be reproduced from one interface to another, even when the small-wavelength fluctuations are completely uncorrelated. Stearns [17] writes the Fourier transform of the height profile of interface j as

$$\tilde{z}_j(q_x) = \tilde{h}_j(q_x) + \tilde{a}_j(q_x)\tilde{z}_{j-1}(q_x) \tag{3}$$

identifying two kinds of contribution to the roughness of interface j: $\tilde{h}_j(q_x)$ is the intrinsic roughness of interface j, and roughness from the previously grown interface $j-1$ propagates to interface j according to a function $\tilde{a}_j(q_x)$, which is probably a Gaussian function of q_x, according to suggestions of Edwards and Wilkinson [18].

Now, consider the extreme, yet not implausible situation, where the low-frequency components with wavelengths larger than τ are completely replicated from one interface to the next, and the intrinsic roughness of each layer $\tilde{h}_j$ has a frequency spectrum with vanishingly small amplitudes at fluctuation wavelengths above τ. The frequency spectrum of an arbitrary layer would be

$$\tilde{z}_j(q_x) = \tilde{h}_j(q_x) + \tilde{L}(q_x) \tag{4}$$

where the low-frequency part $\tilde{L}(q_x)$ with zero amplitude at spatial frequencies larger than $1/\tau$ is completely replicated from the substrate to all subsequent layers.

If there is a frequency gap between the spectra of the intrinsic roughnesses $\tilde{h}_j(q_x)$ and the low-frequency substrate roughness $\tilde{L}(q_x)$, and $1/\tau$ is a frequency somewhere in that gap, then choosing a Gaussian function for the replication function in equation (3):

$$\tilde{a}_j(q) = \tilde{A} = \exp[-(\tau/2\pi)^2 q^2] \tag{5}$$

would result in equation (4). Note that $\tilde{A} \to 1$ as $q \to 0$ for wavelengths long compared to τ, and $\tilde{A} \to 0$ for wavelengths small compared to τ.

Table 2. Parameters for multilayer sample A: layer thicknesses t_{dif} and t_{ref}, and root mean square interface roughnesses σ_{dif} and σ_{ref}, as determined respectively from the diffuse and reflectivity fits. Some of the thickness values for the diffuse fit were kept fixed to the reflectivity values (indicated by asterisks). Only the diffuse scattering is sensitive to the Hurst parameter h (the effective fractal dimension is $D = 3 - h$) and the in-plane cut-off length ξ. The model was particularly insensitive to ξ when the dimension of a layer approached 2, and thus no value is shown for these cases. The germanium fraction is found to be $x = 0.12$. The correlation parameters were $\xi_\perp = 461 \pm 26$ Å, and $\tau = 14.4$ μm. The variances shown for the diffuse scattering fits were those found to increase the χ^2-fitting criterion by 2%.

	t_{dif} (t_{ref}) (Å)	σ_{dif} (σ_{ref}) (Å)	h (D)	ξ (Å)
Vacuum				
		4.55 ± 0.02 (4.54)	0.181 ± 0.002 (2.82)	$20\,000 \pm 1800$
Si oxide	19.96 ± 0.19 (16.39)			
		2.04 ± 0.04 (4.83)	0.86 ± 0.18 (2.14)	$15\,000 \pm 900$
Si	230.81 ± 0.33 (212.3)			
		11.1 ± 2.4 (10.29)	0.57 ± 0.27 (2.43)	2600 ± 600
$Si_{1-x}Ge_x$	65.5*			
		11.1 ± 1.6 (10.59)	$\to$1.0 (2.0)	—
Si	229.3*			
		3.2 ± 3.0 (8.66)	$\to$1.0 (2.0)	—
$Si_{1-x}Ge_x$	66.7*			
		10.7 ± 1.2 (10.51)	0.95 ± 0.4 (2.05)	2600 ± 1000
Si substrate				

The intrinsic height variations h_j are statistically independent of each other. If we also assume that the high-frequency fluctuations are unrelated to the low-frequency ones, it may be seen that using equation (4) results in the correlation function $\tilde{C}_{jk} = \tilde{L}\tilde{L}$. Because of our assumptions for the frequency spectrum of $\tilde{L}$, it makes no difference whether we use the function $\tilde{L}$ or $\tilde{A}\tilde{L}$, and thus we could just as easily write this as

$$\tilde{C}_{jk}(\boldsymbol{q}) = \exp\left[-2(\tau/2\pi)^2 q^2\right] \tilde{L}\tilde{L}. \tag{6}$$

Therefore, we can impart a physically desirable property—low-frequency fluctuations should be highly correlated from one interface to another—to the otherwise appealing correlation function given in equation (2) by multiplying it by the exponential given in

equation (6). Fitting for the value τ should give us an approximate idea of the length scales at which fluctuations become well correlated between interfaces.

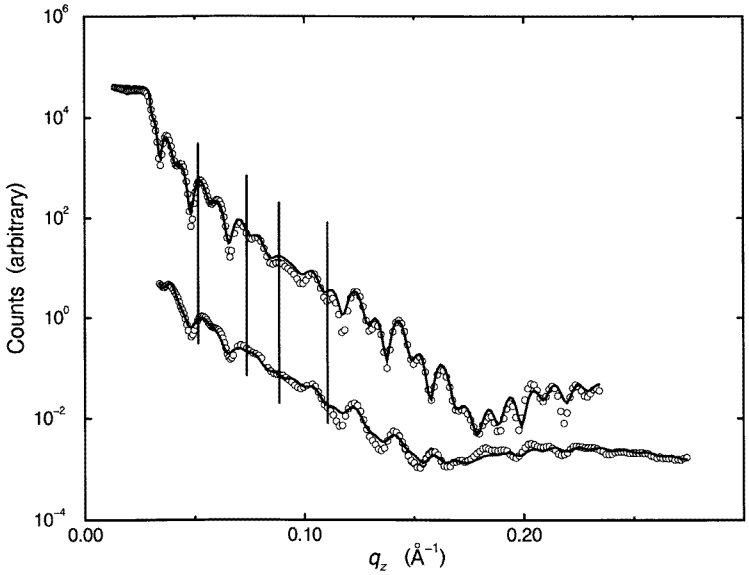

Figure 2. Reflectivity data (upper curve, circles) and an offset scan (lower curve, circles) with $\delta = -0.02^\circ$ for sample A. The reflectivity was fitted to a roughness-modified Fresnel function (upper line). The offset data were fitted (bottom line) simultaneously with the rocking curve data shown in figure 3. The q_z-positions of the rocking curves in figure 3 are indicated with line segments.

3.2. Geometry and resolution

To apply the theory to the experimentally observed quantities, we have to take account of several other quantities. The total sample area illuminated by the beam changes with θ_1. The multilayer sample A was a large (three-inch) wafer. For the rocking scans shown, the beam was always completely intercepted by the sample. Calculated cross-sections were then multiplied by a 'footprint correction' factor, $1/\sin\theta_1$.

During a scan the diffractometer will integrate over a varying resolution volume in reciprocal space. Like Gibaud *et al* [19], we assume small angles, and that the wave-vector resolution $\Delta k_0/k_0$ is small compared to the range of angles $\Delta\theta_1$ and $\Delta\theta_2$ which the diffractometer accepts. It may then be shown that the resolution perpendicular to the specular rod has the form

$$\frac{\Delta q_x}{|\boldsymbol{q}|} \propto \sqrt{C\sin^2\theta_1 + \sin^2\theta_2} \tag{7}$$

where C is a constant which may be found from the data. ($C = 1.29$ for this experiment.) The out-of-plane resolution depends only very weakly on θ_1 and θ_2, so we take it as constant. Thus, we multiplied the calculated diffuse cross-sections by a resolution factor $\sqrt{C\sin^2\theta_1 + \sin^2\theta_2}$.

Though the cross-section is defined per unit solid angle $\mathrm{d}\Omega$, it may be shown that $\mathrm{d}\Omega$ is simply proportional to an area $\mathrm{d}A_q$ in q-space, so no further factors are needed to account

for resolution for the diffuse measurements. However, the cross-section for reflectivity involves delta functions: $\delta(q_x)\delta(q_y)$. Under these circumstances the orientation of the area dA_q relative to the q_z-axis becomes important, resulting in a factor $1/\sin\theta_2$ [3].

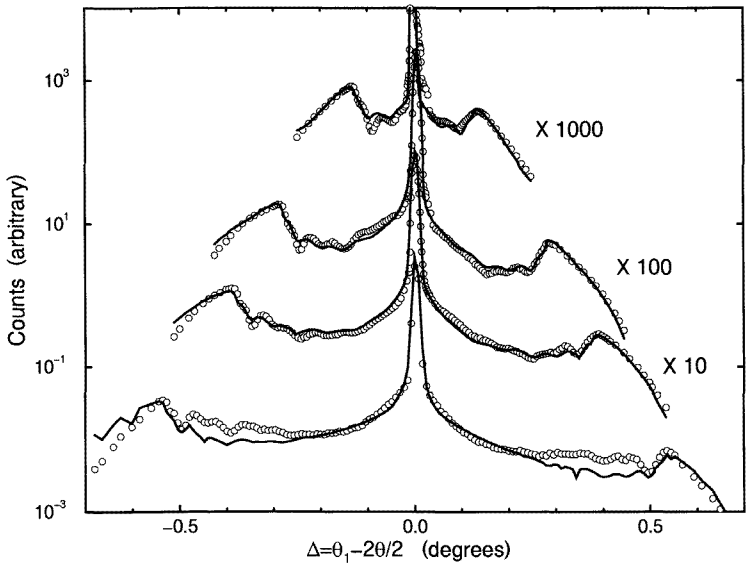

Figure 3. Rocking curves (circles) at $2\theta = 0.7^\circ$ (top, scaled $\times 1000$), 1.0°, 1.2°, and 1.5° (bottom) together with the DWBA fit to the data (solid lines). All four curves as well as the offset scan of figure 2 were fitted simultaneously with one set of layer/interface parameters. For clarity of display, the intensities were plotted against the diffractometer coordinate Δ. For comparison to other figures: $\Delta = 0$ corresponds to $q_x = 0$, the top scan ranges from $q_x = -0.32\times10^{-3}$ Å^{-1} to $+0.32 \times 10^{-3}$ Å^{-1}, and the bottom rocking curve from $q_x = -1.45 \times 10^{-3}$ Å^{-1} to $+1.45 \times 10^{-3}$ Å^{-1}.

4. Results

The top curve in figure 2 shows the specular scattering from sample A, along with a fit from a Fresnel theory [20] modified to take into account interface roughness from many layers [21]. The reflectivity data in figure 2 were acquired by rocking the sample across the specular rod at each successive 2θ-position. As seen from the rocking curves of figure 3, the specular peak appears as a narrow peak at $\theta_1 = (2\theta)/2$ atop the slowly varying diffuse scattering. Each reflectivity point in figure 2 shows the integrated intensity from the specular component of the rocking curve alone. The resulting reflectivity layer thicknesses, $t_{\rm ref}$, and roughnesses, $\sigma_{\rm ref}$, are tabulated in table 2.

To probe the non-specular intensity, we made an offset scan, also shown in figure 2, and wide rocking scans shown in figure 3. The offset scan shows peaks and valleys which are very similar to those in the specular scan. This indicates that the roughness structures of the individual interfaces are strongly correlated in this sample. The scans of figure 3 show a wealth of features. We have already pointed out the specular rod. The leftmost and rightmost peaks in each scan are the so-called Yoneda peaks [22], which occur whenever the incident angle θ_1 or the exit angle θ_2 is just equal to θ_c, the critical angle for total external reflection. The oscillations close to either Yoneda peak are due to dynamical scattering

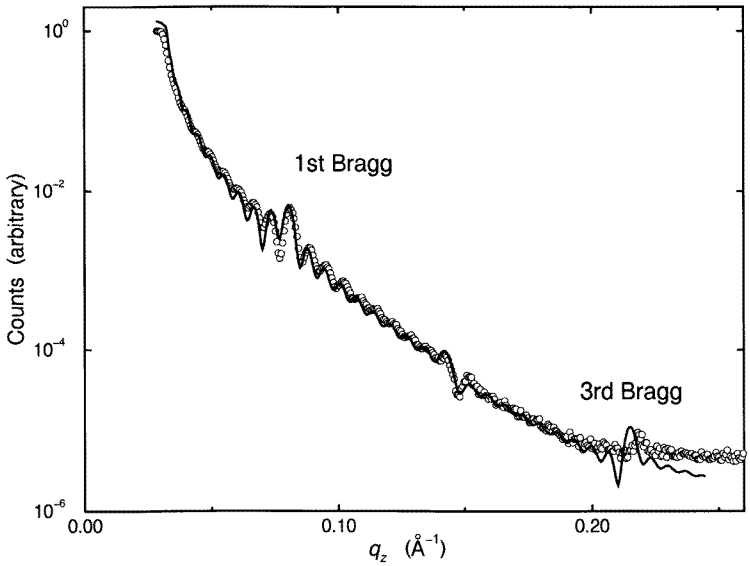

Figure 4. The reflectivity (circles) and fit (line) for the Si/SiGe superlattice, sample B. The intense peaks (labelled 1st Bragg and 3rd Bragg) correspond to scattering from structure with the superlattice periodicity of 100 Å.

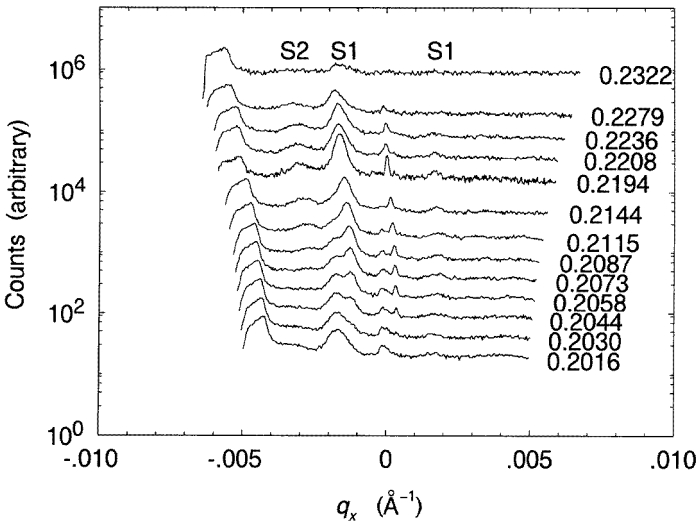

Figure 5. The scattered intensity near the third-order Bragg peak for sample B. The number on the right is the q_z-value (in Å^{-1}) for each rocking curve.

effects, but resemble the kinematical reflectivity oscillations. The similarity occurs in spite of the different scattering mechanisms, because in both cases the scattering is modified by interference from the same layered structure as a function of the incident and exit angles.

We fitted the offset scan and the rocking curves simultaneously to the DWBA model above with one set of layer/interface parameters [6]. Hereafter this is referred to as the *diffuse* fit, and the fitting parameters are tabulated in table 2. With the exception of the

top two layers, where there was reason to believe that the reflectivity and the diffuse scattering would yield noticeably different results, the layer thicknesses for the diffuse fit were constrained to the values found from the reflectivity fit. Though reflectivity and diffuse scattering measure slightly different types of roughness, there is generally good agreement between the values from the reflectivity fit, σ_{ref}, and the diffuse fit, σ_{dif}.

Figure 4 shows the experimental reflectivity of sample B (solid line) from a θ–2θ scan along the specular rod. Because the Si and SiGe layer thicknesses are very nearly equal, the odd-order Bragg peaks are strong, but the even-order peaks nearly vanish.

A preliminary fit (dotted line) to this curve indicates no significant difference between the roughnesses of the Si on SiGe interfaces and the SiGe on Si interfaces.

Figure 5 shows rocking scans made with the miscut direction in the plane of scattering (that is, perpendicular to the surface steps). 13 scans were made at different q_z-positions around the third-order Bragg peak. The one at $q_z = 0.2194$ Å^{-1} is the scan through the centre of the third Bragg peak. The peak appearing at $q_x = 0$ is the specular peak, and the leftmost one is a Yoneda peak. We will discuss the peaks S1 and S2 in the next section.

5. Discussion

We discuss first the results of the simultaneous diffuse fits for sample A summarized in table 2. The near-surface layers are much more precisely characterized than those deeper in the sample. Since the scattering depends on the electron charge-density difference across each interface, a large portion of the scattering will come from the top vacuum/oxide interface. Also, absorption effects will reduce the scattering from deeper interfaces.

Calculations of the diffuse scattering as a function of q_x from a single surface show that as it becomes fractally flat ($D \to 2$) the scattered intensity becomes concentrated in a narrow range ever closer to $q_x = 0$, where it becomes experimentally indistinguishable from the specular scattering. Thus, the diffuse fit is relatively insensitive to the parameters of those interfaces for which D approaches 2.

Clear evidence of the distinction between the sensitivity of diffuse scattering and reflectivity is seen at the silicon oxide (SiO_x)/silicon interface. An interface can be *both* graded *and* have vertical fluctuations in the local average height, and the reflectivity measures the mean width of the transition region due to the two effects. However, the diffuse scattering is sensitive only to the r.m.s. vertical height fluctuations—the roughness measured by diffuse scattering should always be less than that measured by reflectivity. Our reflectivity-measured roughness value of almost twice that measured by diffuse scattering indicates that the (SiO_x)/silicon interface is significantly graded, a plausible result considering that this interface arises from a process of atomic diffusion.

For Si/SiGe heterostructures grown on Si(100), it has been found that the tops of the SiGe layers are often rougher than the tops of the Si layers. This may be understood as follows: the average lattice parameter of a Si–Ge alloy is larger than that of Si. Therefore, in a heterostructure of two kinds of layer with epitaxially matching interfaces, the SiGe layers would be in compression and the Si layers under tension. Xie *et al* [23] have argued that the energy of a free surface can be lowered by either island formation (roughening) for layers under compression (such as SiGe) or smoothing out for layers under tension (Si)—this last perhaps analogous to stretching taut the wrinkled skin of a balloon.

The data from sample A do not show a large asymmetry between the properties of the Si and SiGe layers. The roughnesses of the first (counting from the vacuum) Si-on-SiGe and SiGe-on-Si interfaces are indistinguishable. As far as the fractal dimension is concerned, the first Si-on-SiGe layer has a surface of dimension $D = 2.43$; however, the fit is rather

insensitive to the next few layers below it, though the Si substrate would appear to have a somewhat lower dimension.

The perpendicular correlation length was found to be $\xi_\perp = 460$ Å, slightly less than the total sample thickness. This means that adjacent layers are well correlated, and that there is some correlation between layers across the whole sample. This experiment sets an upper bound on the replication factor τ, since the data set includes just one, well-correlated offset scan: the fitted value was $\tau = 14$ μm, implying that fluctuations with wavelengths longer than 14 μm are quite well replicated from one interface to another. But it is possible that fluctuations with a somewhat shorter wavelength are also well correlated.

Turning now to sample B, we note several interesting features of the rocking scans of figure 5. First, in addition to the specular and the Yoneda peaks, there are additional distinct side peaks on each scan (labelled S1, S2). Second, both the specular peak and the side peak S1 are split in some scans. Such a splitting evolves smoothly when changing q_z. Third, the spacing between the specular peak and S1, and that between S1 and S2 are not the same, indicating that S2 is not a higher-order resonance of S1. Fourth, the side peak S1 does not change its q_x-position when one scans through different q_z-positions. This holds true even further away from the Bragg peak (additional scans, not shown). On the other hand, the side peak S2 does change its q_x-position slightly as q_z changes. The appearance of the side peaks is an indication of strong lateral correlations of the roughness. The separation q_x between the side peak and the specular peak corresponds to a mean lateral undulation wavelength of $2\pi/q_x$. S1, the stronger side peak, indicates that the dominant roughness component has a mean lateral wavelength, $L1$, of about 3700 Å.

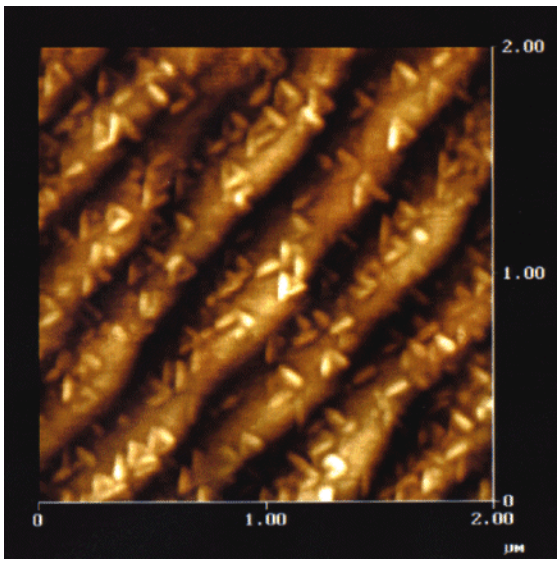

Figure 6. An atomic force microscope image of the surface of sample B. Terraces with approximate length 3700 Å correspond to the scattering component labelled S1 in figure 5.

(This figure can be viewed in colour in the electronic version of the article; see `http://www.iop.org/EJ/welcome`)

We examined the top surface of sample B with an atomic force microscope (AFM). The surface, shown in figure 6, does indeed show a wavy structure with a period of approximately

3700 Å. The mean height difference between the peaks and valleys of these waves, estimated from the AFM picture, is 35 ± 8 Å. The wavy surface structure is due to a replication of the terraced structure of the substrate, or a manifestation of 'ripples' which spontaneously form in lattice-mismatched systems [24].

A vicinal surface will consist of regularly spaced terraces separated by steps or step bunches. If we assume that the 3700 Å period is due to replicated terraces, then taking into account the 0.63° miscut angle of our substrate, the steps between terraces would have a mean height of about 41 Å. This implies a step bunch consisting of about 13 single steps (step height = 3.18 Å) separating successive terraces. Swartzentruber *et al* [25] observed a step bunch size of approximately 10 steps over a wide range of miscut angles for Si(111). A mechanism for this effect has recently been suggested [26]. Thus, the step bunch size implied by our measurements is compatible with the expected substrate terrace structure.

According to a theory elaborated, and measurements made on Si/SiGe systems, by Pidduck *et al* [27], ripples of wavelength λ can form in an overlayer of thickness t if the inequality

$$\frac{t}{\lambda} < \frac{Y g_0^2}{4\gamma\pi^2} \tag{8}$$

is satisfied. The right-hand side, containing the Young's modulus Y for the overlayer, the lattice mismatch strain g_0, and the surface free energy per unit area γ, is proportional to the germanium fraction x in our $Si_{1-x}Ge_x$-on-Si system. This inequality should be satisfied in our system if the behaviour observed is ripples. However, using a thickness $t = 50$ Å, $\lambda = 3700$ Å, $x = 0.10$, and the proportionality constant for $Si_{1-x}Ge_x$ measured by Pidduck *et al*, we find that inequality (8) is violated, the left-hand side exceeding the right-hand side by about 50%. It appears more likely that what we are observing is the result of the replication of the substrate surface structure.

We note that terrace-related peaks were not observed for sample A in figure 3, even though the sample had a miscut of 0.44°. Also, it is puzzling at first glance that the intensity of S1 is greater than the specular intensity. These observations may be explained in the following way. It is straightforward to calculate the scattering from an ideal vicinal surface miscut at angle α, with terraces all of length L and steps all of height h, where $\tan\alpha = h/L$. The results, shown in cartoon form in figure 7, have the following general features. (1) The scattering is concentrated into rods parallel to q_z. (2) Scattering maxima occur at $q_z = (2\pi/h)n$, and $q_x = (2\pi/L)m$, with the highest intensity for rod n when $m = n$ (or, as in figure 7, $m = -n$). (3) The peak position of each rod lies along a line perpendicular to the surface of the terraces. Disorder will broaden the rods out [28], and reduce the intensity of the rods with increasing n.

For sample B, assuming a mean step height $\overline{h}$ of 41 Å, the scans of figure 5 were made for $q_z = 1.31 \times 2\pi/\overline{h}$ to $1.50 \times 2\pi/\overline{h}$. In figure 7, this is close to the peak ($n = 1$) of the first rod ($m = -1$), but far from the peak of the $m = 0$ rod (the specular rod) at the origin. The rocking scans of figure 3 for sample A were made much closer to the q_z-origin, ranging from $q_z = 0.33 \times 2\pi/\overline{h}$ to $0.78 \times 2\pi/\overline{h}$.

Rocking scans at other azimuths (not shown) show that peak S2 rotates about the sample normal with the same phase as S1, indicating that this is scattering due to some structure tilted in the same direction as the miscut. Since its q_x-position is changing with q_z, we cannot associate a unique length scale with this peak, but the peak always lies on a line oriented at 0.84° to the specular rod. No exact structure has been identified for S2. The peak splitting within S1 and the specular rod implies a regularly spaced structure in the in-plane direction. At $q_z = 0.2058$ Å^{-1} the measured peak splitting in S1 is $\Delta q_x = 4.9 \times 10^{-4}$ Å^{-1},

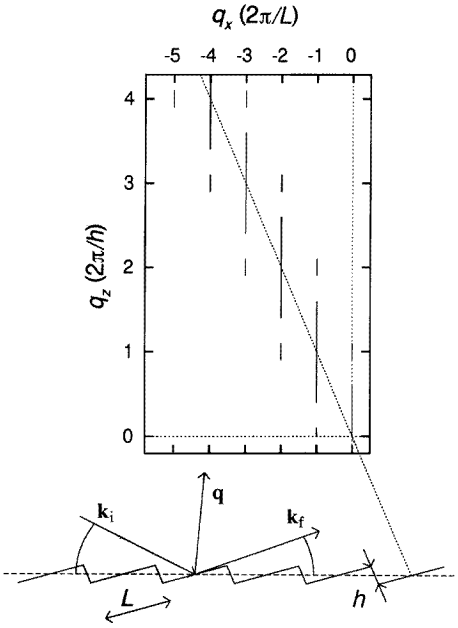

Figure 7. Scattering from a regular array of terraces of length L separated by steps of height h is concentrated along rods in reciprocal space. The figure shows that terrace scattering from the first rod is small compared to the reflectivity (the zero-order rod) near the origin, but becomes more important with increasing q_z. In terms of a step height of 41 Å, the rocking curves shown in figure 5 for sample B were made for $q_z = 1.31 \times (2\pi/h)$ and $1.50 \times (2\pi/h)$, and the rocking scans of figure 3 for sample A ranged from $q_z = 0.33 \times (2\pi/h)$ to $0.78 \times (2\pi/h)$.

which corresponds to an in-plane undulation distance of 1.28 μm—much larger than the scale of the AFM picture of figure 6, and several times the typically 2000 Å corrugation length of self-organized islands formed on SiGe interfaces [29]. Both the peak in S1 and the specular rod appear symmetric at $q_z = 0.2144$, and most split at 0.2058 Å^{-1}, from which we obtain $\Delta q_z \approx 0.017$ Å^{-1}. This in turn implies a tilt of $\tan^{-1}(\Delta q_x/\Delta q_z) \approx 1.6^\circ$.

Perhaps more likely than an island structure is the scenario in which the splitting in S1 and the specular rod arises from the undulations of each interface being not exactly vertically correlated but correlated at a slight angle. Then the splitting arises from interference effects between multiply stepped interfaces. Successive interfaces would be correlated at a tilt angle θ_t of about 1.6°, tilted in the opposite direction to the miscut. Headrick *et al* [14] found evidence using somewhat different measurements for a misorientation of $\theta_t = 45^\circ$ for heterostructures grown on Si(001) miscut by 4°. The distance implied by Δq_z is $2\pi/\Delta q_z = 370$ Å, which is of the same order as the vertical correlation length for large-scale fluctuations found for sample A of about 460 Å.

6. Conclusions

We have introduced and justified a modified correlation function useful in analysing low-angle diffuse x-ray scattering. It allows us to find a mean vertical correlation length connecting height fluctuations of different interfaces in a heterostructure consisting of a few layers. Applying this to a $Si_{0.9}Ge_{0.1}$/Si heterostructure, we found a mean correlation length

of 460 Å. With our model, we found no significant differences between the roughness of the Si-on-SiGe interface, and the SiGe-on-Si interface (both roughnesses were about 11.1 Å); however, we found modest evidence that the fractal dimension of the SiGe-on-Si interface is somewhat lower than that of the Si-on-SiGe interface.

We also examined a $Si_{0.9}Ge_{0.1}$/Si superlattice with both x-rays and an AFM, and found a strongly periodic in-plane structure with wavelength 3700 Å. This structure is probably due to replication of the miscut substrate structure all the way up to the surface.

Acknowledgments

P M Reimer and J-H Li gratefully acknowledge support from the Japan Society for the Promotion of Science. We thank Garry Foran and David Cookson for experimental assistance at the Photon Factory. The synchrotron experiment was supported by the Photon Factory under proposal 94G320. This work was also supported by Monbusho Grants-in-Aid, No 95169 and No 06555092.

References

[1] Sunamura H, Fukatsu S and Shiraki Y 1995 *J. Cryst. Growth* **150** 1038
[2] Tersoff J, Teichert C and Lagally M G 1996 *Phys. Rev. Lett.* **76** 1675
[3] Sinha S K, Sirota E B, Garoff S and Stanley H B 1988 *Phys. Rev.* B **38** 2297
[4] Holy V, Kubena J, Ohlidal I, Lischka K and Plotz W 1993 *Phys. Rev.* B **47** 15 896
[5] Holy V and Baumbach T 1994 *Phys. Rev.* B **49** 10 668
[6] Schlomka J-P, Tolan M, Schwalowsky L, Seeck O H, Stettner J and Press W 1995 *Phys. Rev.* B **51** 2311
[7] Schlomka J-P, Fitzsimmons M R, Pynn R, Stettner J, Seeck O H, Tolan M and Press W 1996 *Physica* B **221** 44
[8] Bauer G, Li J and Koppensteiner E 1995 *J. Cryst. Growth* **157** 61
[9] Headrick R L and Baribeau J-M 1995 *J. Vac. Sci. Technol.* A **13** 782
[10] Ming Z H, Huang S, Soo Y L, Kao Y H, Carns T and Wang K L 1995 *Appl. Phys. Lett.* **67** 629
[11] Phang Y H, Savage D E, Kuech T F, Lagally M G, Park J S and Wang K L 1992 *Appl. Phys. Lett.* **60** 2986
[12] Headrick R L and Baribeau J-M 1993 *Phys. Rev.* B **48** 9174
[13] Phang Y, Teichert C, Lagally M G, Peticolos L J, Bean J C and Kasper E 1994 *Phys. Rev.* B **50** 14 435
[14] Headrick R L, Baribeau J-M and Strausser Y E 1995 *Appl. Phys. Lett.* **66** 96
[15] Barnea Z, Creagh D C, Davis T J, Garrett R F, Janky S, Stevenson S and Wilkins S W 1992 *Rev. Sci. Instrum.* **63** 1068
[16] Brent R P 1973 *Algorithms for Minimization without Derivatives* (Englewood Cliffs, NJ: Prentice-Hall)
[17] Stearns D G 1992 *J. Appl. Phys.* **71** 4286
[18] Edwards S F and Wilkinson D R 1981 *Proc. R. Soc.* A **381** 17
[19] Gibaud A, Vignaud G and Sinha S K 1993 *Acta Crystallogr.* A **49** 642
[20] Parratt L G 1954 *Phys. Rev.* **95** 359
[21] Cowley R A and Ryan T W 1987 *J. Phys. D: Appl. Phys.* **20** 61
[22] Yoneda Y 1963 *Phys. Rev.* **131** 2010
[23] Xie Y H, Gilmer G H, Roland C, Silverman P J, Buratto S K, Cheng J Y, Fitzgerald E A, Kortan A R, Schuppler S, Marcus M A and Citrin P H 1994 *Phys. Rev. Lett.* **73** 3006
[24] Cullis A G 1996 *MRS Bull.* **21** (4) 21
[25] Swartzentruber B S, Mo Y-W, Webb M B and Lagally M G 1989 *J. Vac. Sci. Technol.* A **7** 2901
[26] Tersoff J 1995 *Phys. Rev. Lett.* **75** 2730
[27] Pidduck A J, Robbins D J and Cullis A G 1993 *Microscopy of Semiconducting Materials* ed A G Cullis, J L Hutchison and A E Staton-Bevan (Bristol: Institute of Physics Publishing)
[28] Pukite P R, Lent C S and Cohen P I 1985 *Surf. Sci.* **161** 39
[29] Usami N, Sunamura H, Mine T, Fukatsu S and Shiraki Y 1995 *J. Cryst. Growth* **150** 1065