

Classification
Physics Abstracts
07.80 — 42.30

Quantitative criteria for the matching of simulations with experimental HREM images

M.J. Hÿtch⁽¹⁾ and W.M. Stobbs⁽²⁾

⁽¹⁾ CECM–CNRS, 15 rue G. Urbain, 94407 Vitry cedex, France

⁽²⁾ Department of Materials Science and Metallurgy, University of Cambridge, Pembroke St., Cambridge CB2 3QZ, U.K.

(Received March 24; accepted June 8, 1994)

Résumé. — Nous proposons une méthode quantitative de distinguer entre des modèles structuraux possibles par la comparaison de simulations et d'images expérimentales à haute résolution. La comparaison se déroule en trois parties : la comparaison de l'intensité moyenne, du contraste et de la forme de l'image. Un coefficient de corrélation croisée est défini pour mesurer les différences de forme. L'effet du bruit est caractérisé et il est montré que le coefficient de corrélation répond linéairement aux erreurs dans les paramètres de simulation. Les résultats trouvent leur justification théorique dans l'analyse de Fourier d'image. L'analyse justifie aussi l'utilisation d'une méthode itérative de comparaison de simulation et image expérimentale pour trouver la valeur de plusieurs paramètres expérimentaux en combinaison.

Abstract. — We propose a quantitative way of distinguishing between proposed structural models by the comparison of simulations with experimental high resolution images. The comparison is divided into three parts : the comparison of the mean intensity, the contrast and the pattern. We define a cross-correlation coefficient for the assessment of the differences in pattern. The effect of noise is characterised and it is shown that the cross-correlation chosen responds linearly to errors in the simulation parameters. This is shown theoretically by Fourier analysis of the image and by using simulations. From this analysis it is shown that an iterative method can be used to determine the values of several of the parameters in combination.

1. Introduction.

The determination of structure from HRTEM images is usually based on a comparison with simulated images. The reason is that even assuming that the microscope is modelled perfectly, the dynamical scattering of the electrons, non-linear contributions to the image and the phase shifts introduced by the lens remove the direct link between the form of the image and the structure of

the material in question. Therefore, to determine structure quantitatively, we need objective criteria for judging how well a simulation fits the experimental data. This involves, in particular, the characterisation of the effect of noise on the matching process and the influence of uncertainties in the experimental parameters. It is our aim to establish such criteria, thus making it possible to judge quantitatively the structural models upon which the simulations were based. Typical applications would be for the determination of the structure and composition of a precipitate, the structure of a grain boundary or the variation in composition across an interface.

For specific materials and imaging conditions attempts have been made to extract quantitative data whilst at the same time avoiding the direct comparison of simulations with experimental images. For example, variations in image contrast across an interface between III-V semiconductors at [100] projection have been analysed quantitatively and related to the structure, i.e., compositional variations [1, 2]; the displacement of the positions of dots in the image have been measured and interpreted in terms of compositional variations across strained semi-conductor interfaces [3]. Justification for these methods however needed to be given in terms of image simulations where the structure was known.

An approach based on the comparison of simulations and experimental images is more general as no direct relation between the image contrast and the structure need be assumed. Suppose that we are trying to differentiate between two structural models, 'A' and 'B', one of which is correct. A set of simulations for both models, using best estimates for the imaging conditions, would be compared with the experimental images using a statistical measure. Neither set of simulations would resemble the experimental image exactly, even those for the correct structural model, because of the presence of noise and because of uncertainties over the experimental conditions (thickness, defocus, beam tilt, etc.). Study of the effect of experimental uncertainties on the simulations would then determine whether the level of mismatch was reasonable or not. If one of the simulation sets could be shown to be unreasonably dissimilar to the experimental images then the corresponding model for the structure could be eliminated.

Our investigation will have particular relevance to least squares fitting procedures used to determine structure (e.g., [4]). We will show, for example, that errors in the parameters used to model the microscope can be characterised in terms of linear changes in the matching criterion that we use. This result will provide the justification for iterative matching procedures, used previously for determining thickness and defocus [5], and here, extended to include microscope misalignments.

In this paper we test the basis of the approach we describe using simulated images of [001] $\text{Al}_3\text{Ga}_7\text{As}$ to represent the "experimental" image to see whether it is possible to determine the aluminium content. The aluminium content, in bulk, can be measured with an accuracy of 2% using dark field techniques [6]: it is not suggested that HREM might be better. The problem has been chosen for its simplicity. Having said that, analysis of experimental images of [001] $\text{Al}_3\text{Ga}_7\text{As}$, presented elsewhere [7], has shown that there can be serious difficulties with simulating the exit wavefunction for non-perfect crystals. There are many other experimental factors that are difficult to include in simulation procedures, namely, the effect of inelastic scattering, amorphous contamination, absorption and the effects of specimen damage. Because we shall be using standard simulations throughout these problems will not be dealt with explicitly. It is however essential to keep them in mind.

2. The basis of a quantitative matching process.

2.1 MEAN, CONTRAST AND PATTERN. — We base our comparison of the simulated and experimental images on the measurement of the root mean squared difference, δ , between the image intensities. To carry out the comparison, the experimental image would first be digitised and the intensities normalised with respect to the incident intensity, so as to be on the same intensity scale

as the simulated images. We shall assume that the electron intensities have been recorded linearly (for example, directly by a CCD camera or digitisation of a negative which was sufficiently underexposed for the response of the film to be linear). If x represents the intensity of each pixel in the experimental image, and y the intensity of corresponding pixels in the simulated image, then:

$$\delta \equiv \sqrt{\langle (x - y)^2 \rangle} \quad (1)$$

Unfortunately, there are problems with the use of δ as the sole matching criterion. Imagine, for example, that the mean intensity of the experimental image is less than the incident intensity, as could be due to electrons being scattered out of the objective aperture, and that the simulations did not take into account this effect. The value of δ would then have a value proportional to the differences in mean intensity even when the two images were otherwise identical. In addition, looking at the value of δ gives no indication as to where the problem might lie and hence how the simulations might be improved. However, we can rearrange equation (1) to give the following relation:

$$\delta^2 = [\langle x \rangle - \langle y \rangle]^2 + [\sigma(x) - \sigma(y)]^2 + 2(1 - \rho)\sigma(x)\sigma(y) \quad (2)$$

where $\langle x \rangle$ corresponds to the mean intensity, $\sigma(x)$ is the standard deviation of the intensity:

$$\sigma(x) = \sqrt{\langle x^2 \rangle - \langle x \rangle^2} \quad (3)$$

and ρ is defined according to the following definition:

$$\rho \equiv \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\sigma(x)\sigma(y)} \quad (4)$$

This cross correlation coefficient, ρ , has been used previously to compare quantitatively simulations and experimental images [5] whereas here we can see that it forms only a part of the matching process. It has the property that if the *form* of the intensity variations in the simulated and experimental image are identical, ρ takes the value of one. Adding a uniform background to the image or uniformly stretching the contrast levels does not alter the value of ρ obtained.

For HREM image analysis the patterns of the simulations will, in general, be similar to the experimental images, so that ρ will be close to one. In this case it is useful to define a correlation angle, θ [7], given by:

$$\theta \equiv \cos^{-1} \rho \quad (5)$$

For values of ρ close to one, the small angle approximation applies and

$$\rho \simeq 1 - \frac{1}{2}\theta^2 \quad (6)$$

Substituting this equation back into equation (2) gives us:

$$\delta^2 \simeq [\langle x \rangle - \langle y \rangle]^2 + [\sigma(x) - \sigma(y)]^2 + \theta^2 \sigma(x)\sigma(y) \quad (7)$$

It is this equation which suggests a way in which the matching process can be divided up: the equation consists of three positive terms. We shall refer to $\langle x \rangle$ as the mean intensity, $\sigma(x)$ as the image contrast (our definition) and the form of the variations in the intensities in an image as the pattern of the image. The first term therefore measures the difference in the mean intensity of the simulated and experimental images; the second term measures differences in the image contrast; and the last term measures differences in the pattern.

The main reason for splitting up the matching process is that differences in these three aspects of the experimental and simulated images result from a variety of effects. The mean intensity can be less than the incident intensity if there is an amorphous layer coating the specimen which scatters electrons to high angles, if there is significant inelastic scattering or if the atoms in the material are significantly shifted from their correct lattice sites (as can happen, for example, due to beam damage). The contrast can be reduced by such effects, as well as by, for example, crystal misorientation. The patterns can be different if there is noise present, if there are errors in the modelling parameters or if the structural model is incorrect. There are other possibilities and naturally there can be some correlation between differences in the three (for example, anomalous absorption affects all three aspects of the image). Despite the fact that the mean intensity, contrast and pattern are not completely independent, consideration of each aspect separately helps remove much of the ambiguities involved in the comparison of simulated and experimental intensities.

In practical terms, the comparison of the mean intensity or the contrast of the experimental images with the simulations is not difficult. In this paper we will therefore concentrate on the more complicated problem of comparing patterns. This is also the most critical part of any matching process as the pattern contains the majority of the structural information and is at the same time the most sensitive to misalignments, noise and the general imaging conditions.

3. The effects of noise on the matching process.

For the matching process to be useful we must understand how the criteria suggested will be affected not only by the values of the basic microscope and specimen parameters upon which the imaging depends but also by the various misalignments possible as well as by the presence of noise. We first examine how an image is affected by noise.

As has already been mentioned, the presence of noise will prevent a perfect match between an experimental image and its simulation even when using the correct structural model and imaging parameters. What then is the effect of random noise on θ ? As a preliminary first step it is sensible to consider the angle between two pictures which are identical except that one has some noise superimposed onto it. If n signifies the noise then

$$\rho = \frac{\langle x(x+n) \rangle - \langle x \rangle \langle x+n \rangle}{\sigma(x)\sigma(x+n)} \quad (8)$$

If we assume that there is no correlation between the signal, x , and the noise, n , then $\langle xn \rangle = 0$ and $\sigma(x+n) = \sqrt{\sigma^2(x) + \sigma^2(n)}$. The equation then simplifies to:

$$\rho = \frac{1}{\sqrt{1 + \frac{\sigma^2(n)}{\sigma^2(x)}}} \quad (9)$$

This is the same result as given by Frank [8] where the noise to signal ratio is defined as the ratio of $\sigma(n)$ to $\sigma(x)$.

The corresponding value for θ can be calculated from this result if the level of noise is much lower than the image contrast. Then, from equation (9), we find that

$$\rho \simeq 1 - \frac{1}{2} \frac{\sigma^2(n)}{\sigma^2(x)} \quad (10)$$

By comparison with equation (6) we can see that in the presence of noise:

$$\theta \simeq \frac{\sigma(n)}{\sigma(x)} \quad (11)$$

or the noise to signal ratio.

When calculating the effect on measuring the angle between two different pictures, one with noise (the experimental image) and one without (the simulation), it is tempting to extrapolate from the result above and imagine that the error in the measured θ would be:

$$\Delta\theta = \pm \frac{\sigma(n)}{\sigma(x)} \quad (12)$$

This is not, however, the case. We must go back to the definition of ρ and calculate the expectation value from there.

$$\rho = \frac{\langle(x+n)y\rangle - \langle x+n\rangle\langle y\rangle}{\sigma(x+n)\sigma(y)} \quad (13)$$

$$\simeq \rho_o \frac{1}{\sqrt{1 + \frac{\sigma^2(n)}{\sigma^2(x)}}} \quad (14)$$

where ρ_o is the value of ρ in the absence of noise. There will therefore be a systematic error in the measurement of ρ as it will be reduced by a factor depending on the level of the noise and only weakly on the contrast of the underlying signal.

If θ_0 is the angle corresponding to ρ_0 , the effect on θ can be calculated for small angles by expanding the multiplying factor:

$$\rho \simeq \left(1 - \frac{1}{2}\theta_0^2\right) \left(1 - \frac{1}{2}\frac{\sigma^2(n)}{\sigma^2(x)}\right) \quad (15)$$

$$\simeq 1 - \frac{1}{2}\left(\theta_0^2 + \frac{\sigma^2(n)}{\sigma^2(x)}\right) \quad (16)$$

To simplify, we will define θ_n as:

$$\theta_n \equiv \frac{\sigma(n)}{\sigma(x)} \quad (17)$$

By comparison with equation (6) we see that in the presence of noise :

$$\theta \simeq \sqrt{\theta_0^2 + \theta_n^2} \quad (18)$$

Experimentally it is easy to measure the standard deviation of the noise and therefore the expected mismatch due to noise is simple to calculate. The relationship also allows an assessment of the balance between the gains to be had on averaging and the consequent loss of spatial resolution in a given structural determination.

Computer simulations of the image of a convenient model structure can be used to test some of the results. Figure 1 shows a simulated image of [001] Al₃Ga₇As perfect crystal (unit cell size of 0.565 nm by 0.565 nm in projection), 14 nm thick, at a defocus of -128 nm, which has had some noise added to it. The multi-slice programme used was developed by G.J. Wood from the more basic calculations of Maclagan *et al.* [9] and the imaging of the exit wavefunction was simulated using routines within SEMPER [10]. The microscope was assumed to be perfectly aligned and the other imaging parameters were set to be those generally suitable for the Cambridge HREM operating at 500 kV, namely: a beam divergence of 0.5 mrad, a focal spread of 15 nm, a spherical aberration coefficient of 2.7 mm, a chromatic aberration coefficient of 2.3 mm and no vibration.

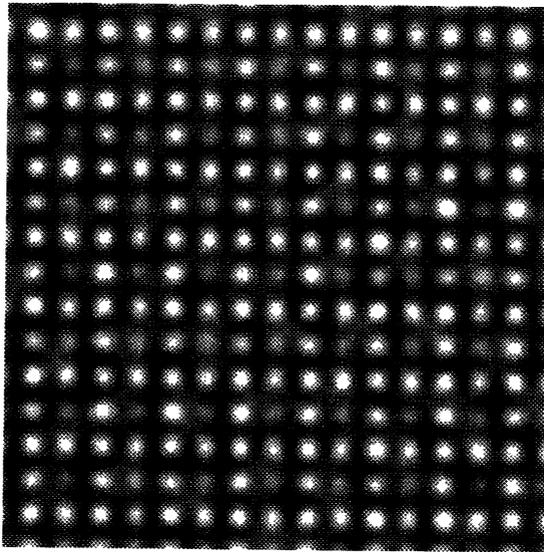


Fig. 1. — Simulated image of [001] $\text{Al}_3\text{Ga}_{7.7}\text{As}$. The crystal was 14 nm thick with 10 nm of amorphous material on the surface and the defocus was -128 nm. The image shows 4 by 4 unit cells, each of which are 0.565 nm square. The other parameters were those suitable to the Cambridge HREM operating at 500 kV namely: beam divergence 0.5 mrad, a focal spread 15 nm, a spherical aberration coefficient of 2.7 mm, a chromatic aberration coefficient of 2.3 mm and no vibration. Incident intensity = 1, black = 0.25, white = 2.5, original intensity range 0 to 2.5.

An image due to 10 nm of amorphous material having the same composition as the crystal was simulated using a recently developed procedure [11]. The mean intensity was then subtracted to produce “noise” having a mean of zero. This image was added incoherently to the image of the perfect crystal. A better way to simulate the effect of a contamination layer is to complete a full multi-slice calculation which propagates the electron wave through the amorphous layer and then through the crystal [11] but here the aim is to analyse the effect of noise which is completely uncorrelated with the image of the crystal.

The image of the perfect crystal without the contamination was compared with each of the unit cells in the image. The mean of θ was found to be 0.28 rad and the standard deviation 0.066 rad. Since the uncontaminated image had a contrast of 0.57 while that of the noise was 0.17, θ_n would be expected to be given by the ratio of these values as about 0.3 rad. With θ_0 (θ in the absence of noise) being zero we therefore have good agreement between the expected and observed values of θ as a function of the addition of noise.

To test the other prediction, an image with a little beam tilt was compared with each of the unit cells in turn. The difference between the misaligned image and the aligned image was 0.18 rad, i.e. θ_0 , and would thus predict an averaged value of $\sqrt{0.18^2 + 0.17^2} = 0.33$ rad for the comparison with the noisy images. The resulting value was in fact 0.34 rad confirming the suitability of the small angle approximations. The standard deviation was 0.057 rad which suggests that the spread in values for θ does not depend critically on the test image.

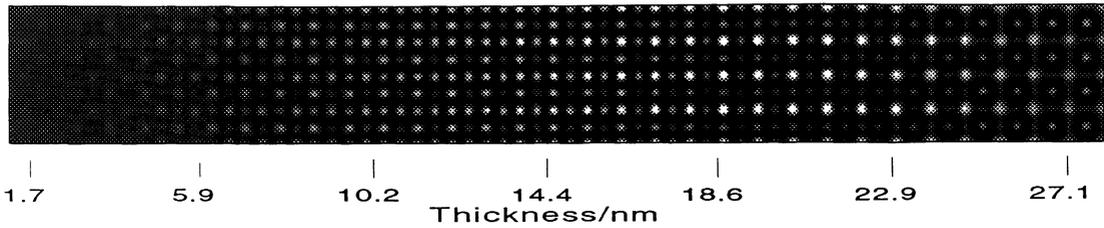


Fig. 2. — Simulations of [001] $\text{Al}_3\text{Ga}_{7.7}\text{As}$, defocus -128 nm, as a function of thickness. One unit cell horizontally and two unit cells vertically are shown for each thickness, the thicknesses being multiples of three unit cells (1.7 nm). Incident intensity = 1, black = 0.3, white = 3.

4. Effects of uncertainties in modelling parameters on the matching process.

Noise is relatively easy to deal with because it is, by its very nature, random. In contrast, changing an imaging parameter such as the beam tilt will produce a systematic change in the image. This change will also be correlated over defocus and thickness. Thus, the characterisation of the effect of misalignment on the matching process is likely to be much more difficult than that of noise. Similar concerns are associated with uncertainties in any of the other fitting parameters (thickness, defocus, astigmatism, crystal tilt, beam divergence, focal spread and specimen vibration). We will show, however, that the angle θ between two simulations is proportional to the error in a particular modelling parameter, for small errors. It will be possible to define the sensitivity of θ to errors in the parameters which can be calculated for all specimen thicknesses and defoci. The uncertainty in the value of a single imaging parameter can therefore be directly related to the uncertainty in θ .

4.1 CHARACTERISING ERRORS IN THE VALUE OF THE THICKNESS. — We shall begin by studying the effects that an error in thickness can have on the correlation between the simulations and the experimental image. A selection of image simulations, again for $\text{Al}_3\text{Ga}_{7.7}\text{As}$ and for the imaging parameters noted above, are shown for a range of thicknesses up to 27 nm in figure 2. Figure 3 shows the comparison of a simulation of a particular thickness representing the “experimental” image with all the others, beginning with the thinnest on the left and the thickest (27 nm) on the right. The five curves correspond to chosen thicknesses of 2.8 nm, 8.5 nm, 14.1 nm, 19.8 nm and 25.4 nm respectively.

It can be seen that the angle between the experimental image and the simulations form a well defined v -shape in the immediate vicinity of the correct thickness. It therefore seems reasonable to define a sensitivity of the images to the thickness, θ_{th} , at each point as the difference between a simulation having that thickness and one a unit cell thinner. The angle between two images of, say, three unit cells difference can be approximated as three times this angle of sensitivity. This approximation is in fact very good, as can be seen by looking at figure 4. The sensitivity has been measured by taking the angle between two simulations differing by only by one unit cell in thickness, differing by two unit cells and halving, and three unit cells and dividing by three. All three angles agree very well over the full thickness range and this remains true at other defoci as is demonstrated for the related graphs shown in figure 4b and figure 4c for defoci of -92 nm and -56 nm respectively.

At most places the change in sensitivity with thickness can be treated as a second order effect on the expected angle of mismatch. For the case examined here it seems that errors in thickness are well characterised and can be directly related to an uncertainty in the overall correlation. For

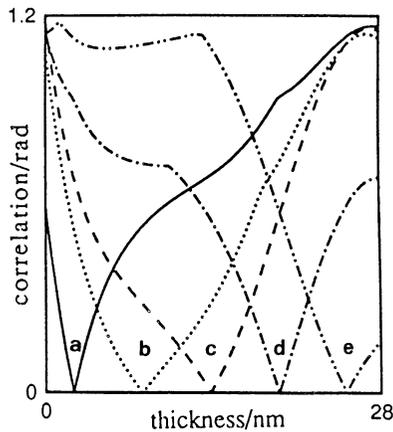


Fig. 3. — Investigation of the effect of errors in thickness. A simulation from figure 2 was chosen to represent the experimental image. This “experimental” image was then compared with each of the other simulations in turn, the horizontal axis corresponding to the thickness of the simulation, and the vertical axis corresponding to the value of θ obtained. The images are identical when $\theta=0$ and each curve corresponds to the use of a different “experimental” image. a) “experimental” image 2.8 nm thick b) “experimental” image 8.5 nm thick c) “experimental” image 14.1 nm thick d) “experimental” image 19.8 nm thick e) “experimental” image 25.4 nm thick.

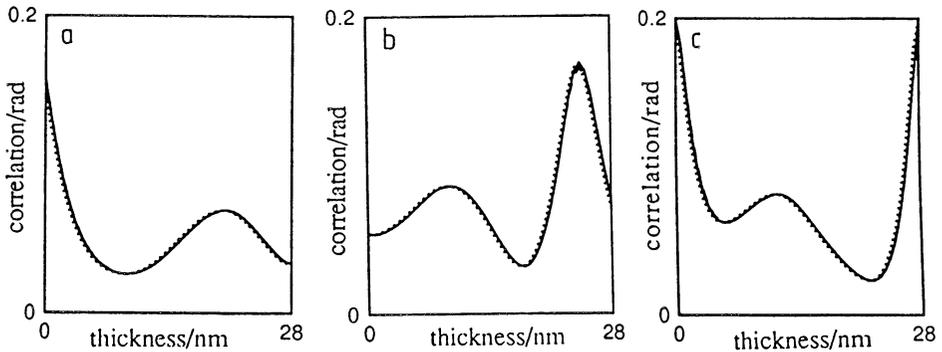


Fig. 4. — Sensitivity of simulations to thickness as measured by θ . If Δt is the difference in thickness (measured in unit cells) then $\theta_{th} = \theta / \Delta t$ with the full, dotted and dashed lines corresponding to a difference of one, two and three unit cells respectively. a) defocus -128 nm; b) defocus -96 nm; c) defocus -56 nm.

other circumstances it would be a relatively simple task to verify the behaviour, as has been done here.

It should incidentally be noted that the curves in figure 3 are not smooth: there are places, other than those at zero θ , where there is a sharp change in gradient. These kinks are due to the manner in which the images are compared. Because it would be impossible to know the correct relative displacement of a simulation and the experimental image (for example, in relation to which position in the image corresponds to a column of gallium atoms) the two images are compared over all possible relative displacements. The value for the mismatch is taken as the minimum possible value of the correlation angle for whatever relative displacement this happens to be. The kinks in the curves in figure 3 occur where the position of minimum mismatch switches abruptly from one

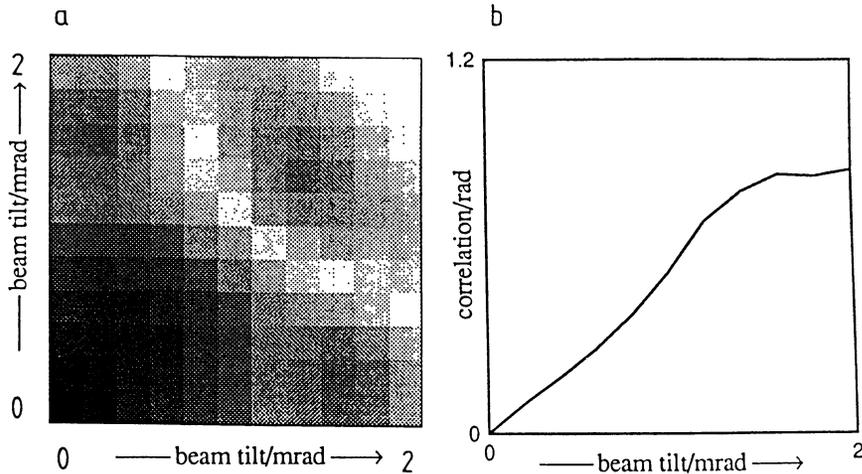


Fig. 5. — Effect of beam tilt on the pattern as measured by the correlation angle θ . A set of images of 14 nm thick [001] $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ were simulated at a defocus of -128 nm for various values of beam tilt. The aligned image was compared with each simulation. a) The results for θ plotted as a function of beam tilt. The beam tilt increases in steps of 0.4 mrad to a maximum of 2 mrad horizontally and vertically corresponding to beam tilts in the [100] and [010] directions respectively. Black=0, white=1.2 rad, b) average of θ over different directions, but equal magnitudes, of beam tilt.

position to another. It will be assumed throughout that the relative position of simulation and experiment is unknown and the minimum mismatch is the one that will be presented.

4.2 THE EFFECT OF BEAM TILT. — Figure 5 shows the effect of tilt on the image at a thickness of 25 unit cells (14 nm) and at a defocus of -128 nm. (The bottom right of the figure corresponds to 2 mrad of tilt in the horizontal, or x -direction, and top left to 2 mrad in the vertical, or y -direction.) For up to 1 mrad of beam tilt the mismatch angle increases uniformly and seems to be independent of the direction of the misalignment as is shown more clearly in figure 5b. We have investigated other defoci and found that this to be generally the case.

4.3 THE EFFECTS OF ASTIGMATISM. — Figure 6 shows the the correlation between the aligned image and the misaligned images for a set of simulations for a thickness of 14 nm and at a defocus of -128 nm. The behaviour of the correlation angle depends strongly on the direction of the astigmatism. Along one of the major axes the angle behaves linearly, as can be seen in figure 6b where θ is plotted along the horizontal axis. Away from the major axes, however, the angle rises quickly to a peak and then decays. This is shown in figure 6b where θ has also been plotted for the astigmatism in a direction $\pi/4$ rad to the major axes. It seems that in order to characterise the effect of astigmatism the sensitivity of θ needs to be defined along a principal axis and at 45° to it. It is anticipated that the presence and appropriate direction should be apparent from the amorphous carbon. If the level of astigmatism present is small (less than about five or six nanometre) and its direction is known, the astigmatism is well characterised by an angle of sensitivity.

4.4 THE EFFECTS OF CRYSTAL TILT. — Investigations into the effect of crystal tilt has shown that, similarly to beam tilt, θ changes linearly with the *magnitude* of the tilt. As a general comment, the relative sensitivity of θ to beam tilt and crystal tilt could be used to assess quantitatively the validity

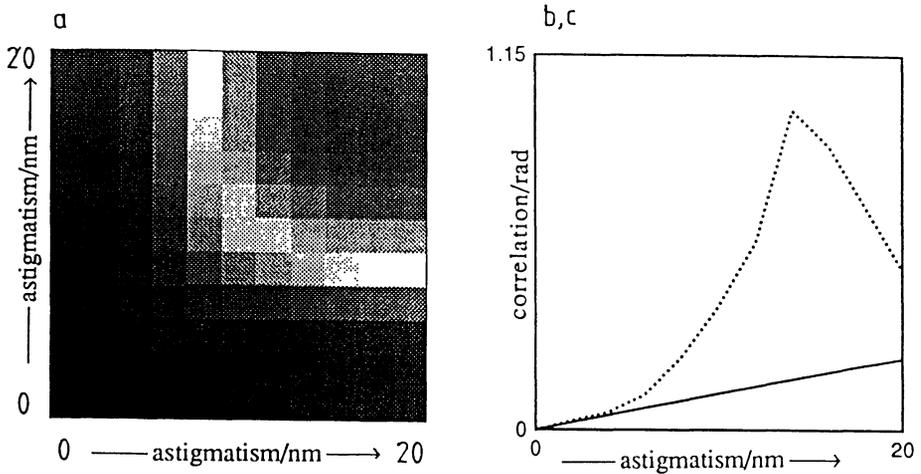


Fig. 6. — Effect of astigmatism on the pattern as measured by the correlation angle θ . A set of images of 14 nm thick [001] $\text{Al}_{1.3}\text{Ga}_{0.7}\text{As}$ were simulated at a defocus of -128 nm for various values of astigmatism. The aligned image was compared with each simulation. a) The results for θ plotted as a function of astigmatism. The astigmatism increases in steps of 4 nm to a maximum of 20 nm horizontally and vertically corresponding to astigmatism in the [100] and [010] directions respectively. Black=0, white=1.2 rad, b) θ for astigmatism along major axis (full line) and 45° to major axis (dotted line).

of the normal method of calculating the effect of beam divergence. What is usually assumed, as here, is that beam divergence calculations do not take into account the fact that the electrons impinging on the specimen have differing angles of incidence and thus scatter differently. The effect of having beams of electrons travelling at different angles to the optic axis usually has the most significant effect on the imaging behaviour.

4.5 THE EFFECTS OF OTHER PARAMETERS. — It turns out that the other imaging parameters — defocus, beam divergence, focal spread and vibration — all behave quite well, in terms of their effect on the pattern, as measured by θ . Figure 7 summarises this behaviour at the defocus of -128 nm, over the full thickness range as well as for the other parameters dealt with earlier. Each graph has three curves plotted on it, each showing the sensitivity of θ to the parameter in question but as determined using different degrees of parameter error so as to assess the linearity of the effect of the parameter in question. Figure 7a is, in fact, identical to figure 4a. In order to understand the approach we may note, as an example, that the first curve (full line) in figure 7c plots the angle θ between simulations having 0 and 0.4 mrad of beam tilt. The second curve (dotted line) plots half the angle between simulations having 0 and 0.8 mrad of tilt and the last curve (dashed line) plots a third of the angle between simulations having 0 and 1.2 mrad of tilt. If the relationship between errors in the parameter and θ is linear then all three curves should lie on top of each other. The exception is vibration where it seems that the square of the vibration parameter holds the linear relationship. All of the parameters behave quite well in this respect particularly thickness, defocus, astigmatism (along a major axis), beam divergence and vibration. Where the relation breaks down for the other parameters it is usually for the higher values of θ , which is to be expected. The first curve, if not accurate for higher errors in the parameter, usually provides an over-estimate which is good. Analysing this problem more generally is not feasible but in an actual experiment, curves such as those in figure 7 could be calculated for the specific conditions involved.

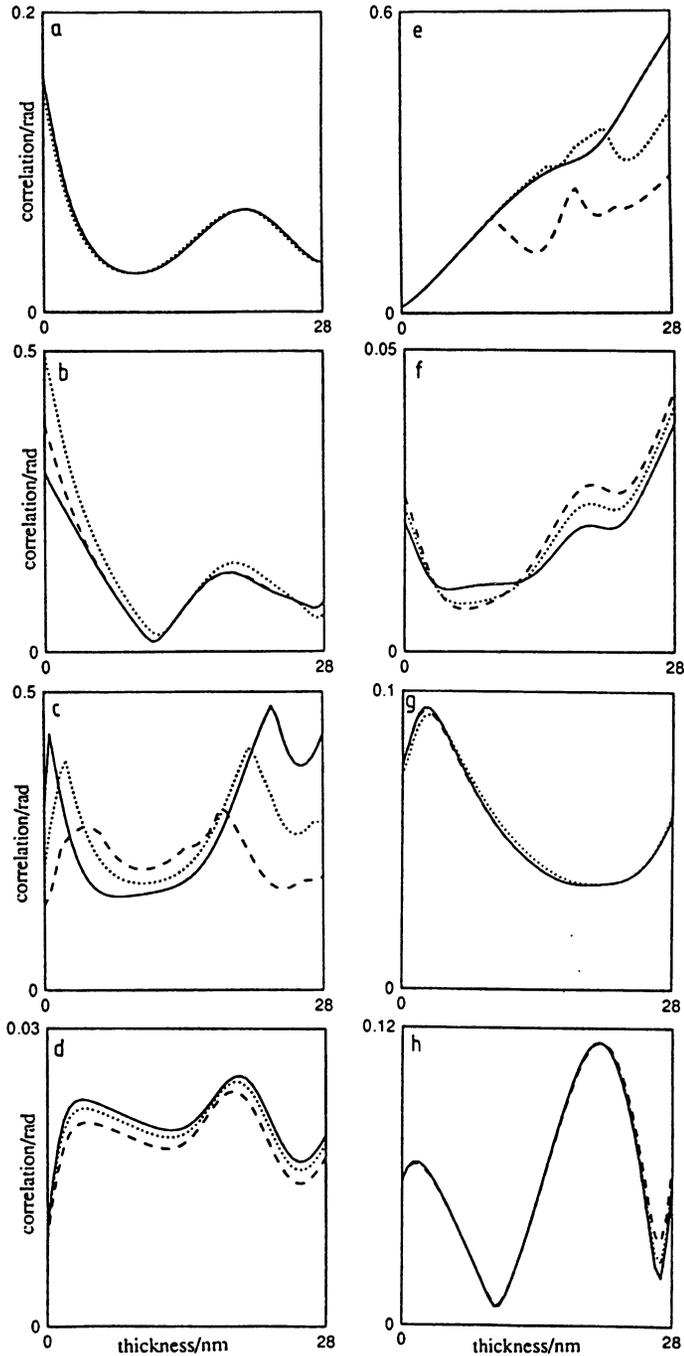


Fig. 7. — Sensitivity of pattern to the various parameters over the whole thickness range, at -128 nm defocus. For the misalignments the direction is along one of the major axes. The sensitivity (full line) is calculated by measuring θ between two simulations differing in one of the parameters (the degree of difference is quoted for each parameter below). The dotted line corresponds to taking a difference twice as big but dividing the result by two, and the dashed line corresponds to taking a difference three times as big but dividing the result by three. The similarity of the three curves is a measure of the linearity of the response of θ to errors in the parameters. a) thickness/unit cell; b) defocus/3 nm; c) beam tilt/0.4 mrad; d) astigmatism/4 nm; e) crystal tilt/mrad; f) beam divergence/0.1 mrad; g) focal spread/nm; h) vibration/ 10^{-4} nm².

5. Fourier analysis of the Image Contrast.

In the treatment we have described so far, we have not relied in any specific way on understanding how various specific parameters individually affect the image. If we are to gain any confidence in the general application of correlation coefficients we need to explain why θ should respond linearly with errors in the parameters. The way we can do this is by considering how the different Fourier components of the simulated image behave both in relation to the projected structure and as a function of the changes in the parameter.

Suppose, then, that an image x can be written in terms of its Fourier components, x_g , such that:

$$x(\mathbf{r}) = \sum_{\mathbf{g}} x_g \exp \{2\pi \mathbf{g} \cdot \mathbf{r}\} \quad (19)$$

It can be shown by substituting this equation, and the equivalent equation for y , into equation (4) that

$$\rho = \frac{\sum_{\mathbf{g} \neq 0} x_g y_g^*}{\sqrt{\sum_{\mathbf{g} \neq 0} |x_g|^2 \sum_{\mathbf{g} \neq 0} |y_g|^2}} \quad (20)$$

$$\sigma(x) = \sqrt{\sum_{\mathbf{g} \neq 0} |x_g|^2} \quad (21)$$

These expressions show that just as the images can be compared point to point so too can a Fourier component in one picture be compared with the corresponding component in another without any cross terms (see also [12]).

For small differences characterised by $y_g = x_g + \Delta x_g$ equation (20) reduces to :

$$\rho \simeq 1 - \frac{1}{2} \frac{\sum_{\mathbf{g} \neq 0} |\Delta x_g|^2}{\sum_{\mathbf{g} \neq 0} |x_g|^2} \quad (22)$$

where higher order terms in Δx_g have been ignored. Comparing this equation with equation (6) we see that

$$\theta \simeq \sqrt{\frac{\sum_{\mathbf{g} \neq 0} |\Delta x_g|^2}{\sum_{\mathbf{g} \neq 0} |x_g|^2}} \quad (23)$$

This expression can be used to explain why the correlation angle behaves in the way that it does. Suppose that a single parameter is changed and this in turn changes one of the Fourier components, then :

$$\theta \simeq \frac{|\Delta x_g|}{\sigma(x)} \quad (24)$$

so we see that changes in the angle are linearly proportional to changes in the Fourier component. In general, the Fourier components of the image change slowly with changes in the parameters. For example, changes in thickness will change the Fourier components of the image according to the gradient of the Pendellösung profile. Second order changes will occur at minima (or maxima) for the beams but rarely will all beams go through a minima at the same thickness. We would therefore expect changes in the Fourier components — and hence the angle θ — to be proportional to small changes in thickness. The same argument applies to the other parameters.

The only exception found was for vibration. The reason why this parameter behaves differently is that the changes in θ were with respect to zero vibration initially. Since we begin from a minimum we might expect increasing vibration to introduce a quadratic variation in the Fourier components — which we see in the variation of θ .

It is difficult to prove these statements analytically because of the complexity of the imaging, but for the case of beam tilt it is possible for linear interference. Let $\Phi(\mathbf{g})_{\text{linear}}$ be the Fourier component of the image with periodicity \mathbf{g} in the absence of non-linear interference. The linear interference between a beam \mathbf{g} and the central beam, in the presence of beam tilt, \mathbf{t} , but in the absence of beam divergence or focal spread, is given [13] by

$$\Phi(\mathbf{g})_{\text{linear}} \simeq 2\Re[\phi(\mathbf{g})\phi^*(\mathbf{0}) \exp\{-i\gamma(\mathbf{g})\}] \exp\{-2\pi i(g^2 + D)\mathbf{g}\cdot\mathbf{t}\} \quad (25)$$

where $\phi(\mathbf{g})$ is the complex amplitude of the beam \mathbf{g} , γ is the phase of the lens complex transfer function, and D the defocus (underfocus negative). The expression is in reduced units of Glazers and Scherzers. The effect of beam tilt is just to change the phase of the Fourier component, or in other words, to shift the position of the fringes. If

$$y_g = x_g \exp\{-i\varphi\} \quad (26)$$

$$|\Delta x_g| \simeq \varphi|x_g| \quad (27)$$

for small φ , thus explaining the earlier result.

The approximate independence of θ to the direction of the beam tilt can also be explained. For the special case of AlGaAs the x -axis and y -axis are equivalent, therefore, in the absence of beam tilt or crystal tilt, the Fourier components in the two directions will be identical. When beam tilt is introduced for any pair of beams, g_1 and g_2 at ninety degrees to each other, the total effect will be such that:

$$|\Delta x_1|^2 + |\Delta x_2|^2 = \varphi_1^2|x_1|^2 + \varphi_2^2|x_2|^2 \quad (28)$$

$$= |x_1|^2 \{(2\pi(g_1^2 + D)\mathbf{g}_1\cdot\mathbf{t})^2 + (2\pi(g_2^2 + D)\mathbf{g}_2\cdot\mathbf{t})^2\} \quad (29)$$

$$= |x_1|^2 4\pi^2 (g_1^2 + D)^2 \{(\mathbf{g}_1\cdot\mathbf{t})^2 + (\mathbf{g}_2\cdot\mathbf{t})^2\} \quad (30)$$

$$= |x_1|^2 4\pi^2 (g_1^2 + D)^2 g_1^2 t^2 \quad (31)$$

which is independent of tilt direction. When the effect on the envelope factors due to beam divergence and focal spread becomes important this result is likely not to hold, as is seen for high angles of tilt. A similar treatment can be carried out for crystal tilt. We expect that this would be the case for any material with four fold symmetry in projection.

6. Dealing with errors in more than one parameter.

6.1 HOW TO ESTIMATE THE TOTAL MISMATCH. — In all the preceding analysis we have implicitly assumed that there is an error in only one of the simulation parameters. To develop a realistic criterion for determining whether a simulation is too dissimilar to the experimental image for the structural model to be correct, we need to estimate the total mismatch expected due to errors in many parameters. We can gain insight into this problem using Fourier analysis. If there are errors

in just two parameters that affect different components x_1 and x_2 , then, according to equation (23):

$$\theta \simeq \frac{\sqrt{|\Delta x_1|^2 + |\Delta x_2|^2}}{\sigma(x)} \quad (32)$$

$$\simeq \sqrt{\theta_1^2 + \theta_2^2} \quad (33)$$

where θ_1 and θ_2 are the angles due to changes in the respective individual parameters. This would be the case for errors in thickness and beam tilt, for example. Thickness will only change the amplitude of the Fourier components of the image whereas beam tilt will only affect the phase, to a first approximation. The two effects are therefore independent and equation (33) will apply. We have found this to be the case for small beam tilts. In general, though, the effects are not independent so we propose that the total mismatch will be less than the sum of the angles for the individual parameter errors. Again, this cannot be proved analytically but seems reasonable given the form of equation (23).

6.2 THE ITERATIVE METHOD FOR THE ASSESSMENT OF THE VALUE OF SEVERAL PARAMETERS. — There is another approach to dealing with multiple parameter errors and this is to find the exact imaging parameters. In this context, we propose the use of an iterative method which can simultaneously minimise the uncertainty in six parameters these being the thickness, defocus, beam tilt (in the two directions) and astigmatism (in the two directions). This scheme has been used successfully for an experimental focal series of [001] $\text{Al}_{1.3}\text{Ga}_{0.7}\text{As}$ images [7] and a similar iterative scheme has been shown to work experimentally for determinations of thickness and defocus [5]. This approach to determining the experimental parameters and, more importantly, the exit wavefunction, is an alternative to the wavefunction reconstruction methods [14, 15] as it deals more directly with the experimental problems such as microscope misalignment, specimen misorientation and the presence of noise due to amorphous layers of contamination. Structural models are, however, necessary: the exit wavefunction cannot be determined independently of simulations of the electron-specimen interaction (via multislice calculations etc.).

We have investigated various iterative methods of determining the generating parameters using a simulation as the “experimental” picture. We shall describe one which works quite well. For AlGaAs and the microscope performance considered here we have found that sensible convergence of an iterative approach can be obtained when the “experimental” picture is less than 28 nm thick, the defocus is within a known range 50 nm wide, the beam tilt is less than 2 mrad, and the astigmatism is less than 20 nm while all the other parameters are taken to be known exactly. To begin with, simulations covering the possible values of thickness and defocus are generated and compared with the “experimental” picture. The values of thickness and defocus showing the best fit, as measured by θ , are then chosen as the first guess for these parameters. Then a set of simulations are produced ranging over the possible beam tilts, having set the thickness and defocus to the values which had already been guessed. The beam tilt in the simulations is now set to the value which fits the best and then astigmatism is finally analysed in the same fashion. Once approximate values for all four parameters (technically there are six since the beam tilt and astigmatism have two independent components) have been found it is time to go round the loop again.

The parameters are continuous variables (with the exception of the thickness) but the test simulations can only be of finite number. There is a balance between wanting to sample the parameter space finely and wanting to keep the time spent calculating the images to a minimum. From our experience it seems that it is adequate to sample the thickness every 1.7 nm (three unit cells), the defocus every 5 nm, the beam tilt every 0.4 mrad (in both directions) and the astigmatism every 4

Table I. — *Parameter sensitivity calculated at a thickness of 9 nm and a defocus of -120 nm. Sensitivity is defined as the angle θ between the original image and an image simulated after a small change in one of the parameters.*

| parameter | sensitivity |
|-------------|--|
| thickness | 0.16rad / 1.6nm |
| defocus | 0.18rad / 5nm |
| beam tilt | 0.25rad / 0.4mrad |
| astigmatism | 0.024rad / 4nm parallel to a principal axe (100) or (010) 0.10rad / 4nm along the (110) direction |

nm (again in both directions.) Once the parameters have stopped changing their values, on going around the iterative loop, it is possible to increase the sampling density and simultaneously to reduce the range of values analysed. This method is nearly always successful. For example, given a simulation of [001] $\text{Al}_{1.3}\text{Ga}_{0.7}\text{As}$ of thickness 9 nm, defocus -120 nm, beam tilt (0.7,0.85)mrad and astigmatism (1.5,6)nm as the “experimental” image, the iterative routine produces a value of $\theta=0.08$ rad for a simulation with thickness 8.5 nm, defocus -118 nm, beam tilt (0.6,0.8)mrad and astigmatism (1,7.5)nm after only four cycles. When it does fail, in that the match finds a local rather than a global optimum, the mismatch is then large enough to make it obvious ($\theta > 0.3$ rad). In order to do better it is usually sufficient to change the order in which the parameters are looked at and begin again. Sometimes also the beam tilt can get caught in the wrong quadrant (referring to its angular position in the diffraction pattern) in which case artificially moving to another one usually helps.

It should be noted that the sampling density used can be justified very simply. The sensitivity of θ to each of the parameters explored, for the example given, are shown in table I. It can be seen that the changes in θ from point to point in parameter space are roughly equal for all the different parameters. As a general comment, the method seems to work well because of the monotonic effect of the parameters on the correlation angle θ . Equally importantly, it is that effects of beam tilt and astigmatism are fairly independent of the thickness and defocus which allows them to be dealt with successively, the thickness and defocus having been optimised simultaneously because of their strong interaction.

7. The possibility of measuring the aluminium concentration in $\text{Al}_x\text{Ga}_{1-x}\text{As}$.

All the tools are now available which would allow the assessment of whether the concentration of aluminium in $\text{Al}_x\text{Ga}_{1-x}\text{As}$ could be measured by direct comparison of simulations with HREM images. A general theoretical investigation is beyond the scope of this paper but we can examine how to go about such an assessment in practice.

Simulations of [001] GaAs and [001] AlAs were performed for a range of thicknesses and defoci using parameters suitable to the Cambridge HREM operating at 500 kV. The images of GaAs were then compared with the images of AlAs having the same thickness and defocus. Figure 8 shows the root mean squared difference, δ , between an image of GaAs and as image of AlAs of the same thickness and defocus. In general terms, the differences increase with thickness and

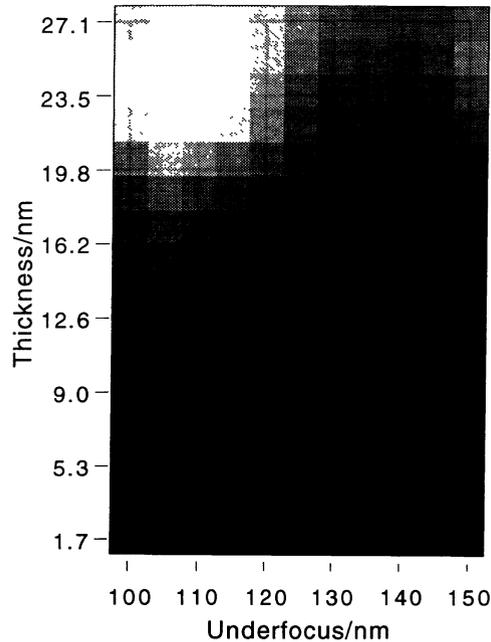


Fig. 8. — Sensitivity of the images to aluminium content. Simulations of [001] GaAs and [001] AlAs were compared for the same thickness and defocus. The results for the root mean squared difference, δ , are plotted as a function of thickness and defocus. Black=0, white=1.

become considerable (~ 1) for thicknesses of about 20 nm.

Because the effects of absorption were not included in the calculations, the root mean squared difference, δ , can be written, according to equation (7), as

$$\delta^2 \simeq |\sigma_{\text{AlAs}} - \sigma_{\text{GaAs}}|^2 + \theta^2 \sigma_{\text{AlAs}} \sigma_{\text{GaAs}} \quad (34)$$

The only contributions are due to differences in contrast and pattern. Figure 9a shows the values obtained for θ , which now represents the degree to which chemistry affects the pattern for each thickness and defocus, and figure 9b shows the modulus of the difference in contrast. In equation (34), however, the term concerning the pattern was multiplied by the product of the contrast of the two images. The differences in contrast shown in figure 9b were therefore normalised with respect to $\sqrt{\sigma_{\text{AlAs}} \sigma_{\text{GaAs}}}$ so that a fair comparison could be made with the values of θ .

It can be seen that for thicknesses lower than about 20 nm, differences in pattern contribute more to the root mean squared difference, δ , than differences in contrast. However, for thicknesses greater than 20 nm differences in contrast become much more important.

The sensitivity of the image to the aluminium content is not however the only factor: the image sensitivity must be compared with the uncertainties introduced by the incomplete knowledge of the remaining simulation parameters. If a region which is highly sensitive to the aluminium content is also highly sensitive to changes in thickness, the accuracy of any measurement could be severely compromised by any small error in the value of thickness used in the simulations. In the previous section the means of estimating the expected mismatch in the pattern due to parameter errors were presented. Once the errors in parameters have been estimated, it is therefore possible to plot the expected mismatch, as measured by θ , as a function of thickness and defocus.

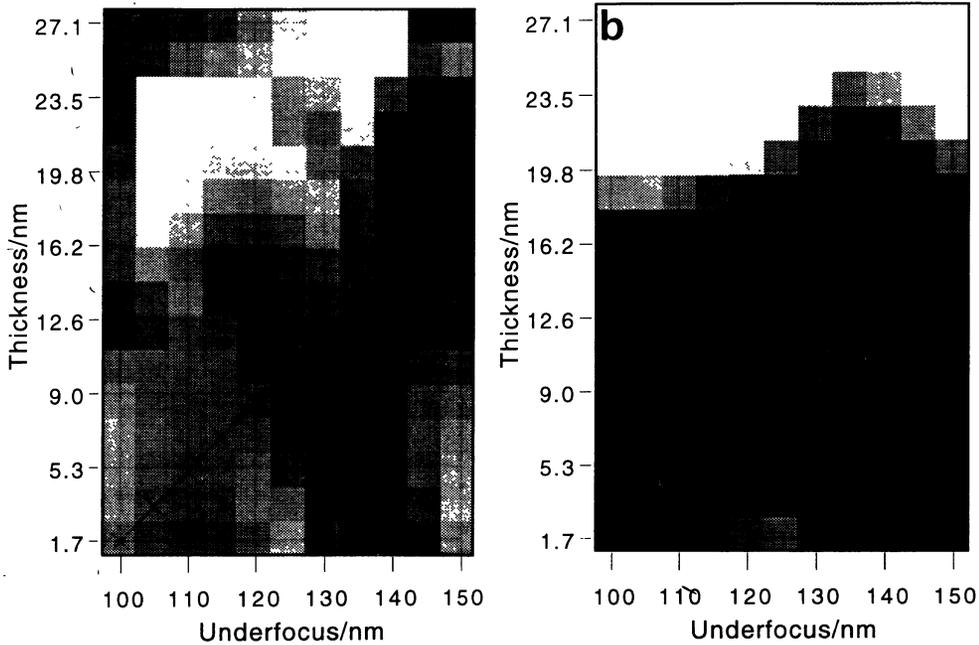


Fig. 9. — The contributions to the root mean squared difference, δ , between images of [001] GaAs and [001] AlAs due to differences in pattern and contrast. The values have been plotted on a scale which allows direct comparison of the relative importance of the two contributions. Black=0, white=1. a) differences in pattern measured by θ ; b) differences in contrast measured by $|\sigma_{\text{AlAs}} - \sigma_{\text{GaAs}}|/\sqrt{\sigma_{\text{AlAs}}\sigma_{\text{GaAs}}}$.

The values obtained could then be directly compared with the sensitivity of the image to the aluminium content. Where the expected mismatch due to parameter errors was as least as important as those introduced by changes in chemistry, it would not be possible to measure the aluminium content with any accuracy. As refinements in the modelling of the aluminium content become relatively more important, the accuracy of the measurement would increase correspondingly. (If an iterative method had been used to find the best matching simulation it would be the remaining degree of mismatch which should be compared with the sensitivity of the image to the aluminium content.)

For thicknesses less than 20 nm, the average value of θ is 0.4 rad. Comparing this value to the sensitivities in table I shows that even small errors in the parameters are going to produce large uncertainties in the measured values of x . It therefore seems from this brief analysis that the best way to differentiate between the two structures would be to measure the contrast of the image for thicknesses greater than 20 nm, the contrast of the AlAs image being as much as 16 times greater than the GaAs image. To determine the accuracy, however, the effect on the contrast of errors in the simulation parameters would have to be assessed.

8. Conclusion.

Here we have described how a general method of quantitative analysis of high resolution images can be based on the root mean squared difference between simulated and experimental image intensities. It was shown that the matching process can usefully be separated into the comparison of

mean intensity, contrast and pattern. We have shown that the pattern, as measured by θ , changes linearly with almost all of the modelling parameters. Furthermore the effect of noise can be well characterised and effects due to combinations of several imaging parameters can be satisfactorily managed. An iterative method can, for example, be used to follow the angle θ in order to determine the thickness, defocus, beam tilt and astigmatism of a simulated "experimental" picture. The simple relationship which we have demonstrated to exist for AlGaAs between θ and changes in the Fourier components which make up the image examined suggests that the correlation angle will behave well for other materials. While this cannot be proved generally, whether or not such a relationship exists for any other material and range of imaging conditions of interest could easily be tested. We should also like to add that θ could be quoted to describe objectively the success of a match between the *pattern* of experimental images and simulations.

Previously HREM image analysis has concentrated almost entirely on the pattern of the images. It has been shown that a complete analysis considers differences in the mean intensity and contrast; for the case studied it was even shown that the contrast could be the aspect of the image most responsive to changes in structure. There is, however, a more important reason for basing the analysis on the root mean squared difference. It is that, here, we have only considered simulations. The effects of surface layers of contamination or damaged material and inelastic scattering have not been included. These aspects can have important consequences when real experimental images are compared with simulations [7]. If large differences in the absolute values of the image intensities remain, there must be doubt in the validity of a comparison based only on the consideration of the patterns of the simulated and experimental images.

Acknowledgments.

The authors are grateful to Owen Saxton for help with the computing. Most of the work was completed whilst MJH was at the Materials Science department at Cambridge and the authors would like to thank the SERC for financial support. MJH is currently funded by the E.U. under the Human Capital and Mobility scheme. The authors would also like to thank Profs. Hull and Humphreys for the use of the laboratory facilities at Cambridge and Dr. Chevalier for useful discussion and for the use of the laboratory facilities at Vitry.

References

- [1] Ourmazd A., Taylor D.W., Bode M. and Kim Y., *Science* **246** (1989) 1571.
- [2] de Jong A.F. and van Dyck D., *Ultramicroscopy* **33** (1990) 269.
- [3] Bierwolf R., Hohenstein H., Philipp F., Brandt O., Crook G.E. and Ploog K., *Ultramicroscopy* **49** (1993) 273.
- [4] King W.E. and Campbell G.H., *Ultramicroscopy* **51** (1992) 128.
- [5] Thust A. and Urban K., *Ultramicroscopy* **45** (1992) 23.
- [6] Bithell E.G. and Stobbs W.M., *Philos. Mag. A* **60** (1989) 39.
- [7] Hýtch M.J. and Stobbs W.M., *Ultramicroscopy* **53** (1994) 191.
- [8] Frank J., *Computer Processing of Electron Microscope Images*, P.W. Hawkes Ed. (Springer, Berlin, 1980).
- [9] Maclagan D.S., Bursill L.A. and Spargo A.E.C., *Philos. Mag.* **35** (1977) 757.
- [10] Saxton W.O., Pitt T.J. and Horner M., *Ultramicroscopy* **4** (1979) 343.
- [11] Chevalier J.P. and Hýtch M.J., *Ultramicroscopy* **52** (1993) 415.

- [12] van Heel M., Schatz M; and Orlova e., *Ultramicroscopy* **46** (1992) 307.
- [13] Smith D.J., Saxton W.O., O'Keefe M.A., Woods G.J. and Stobbs W.M., *Ultramicroscopy* **11** (1983) 263.
- [14] Saxton W.O., *J. Microsc. Spectrosc. Electron.* **5** (1980) 661.
- [15] Coene W, Janssen G., de Beeck M.O. and van Dyck D., *Phys. Rev. Lett.* **69** (1992) 3743.