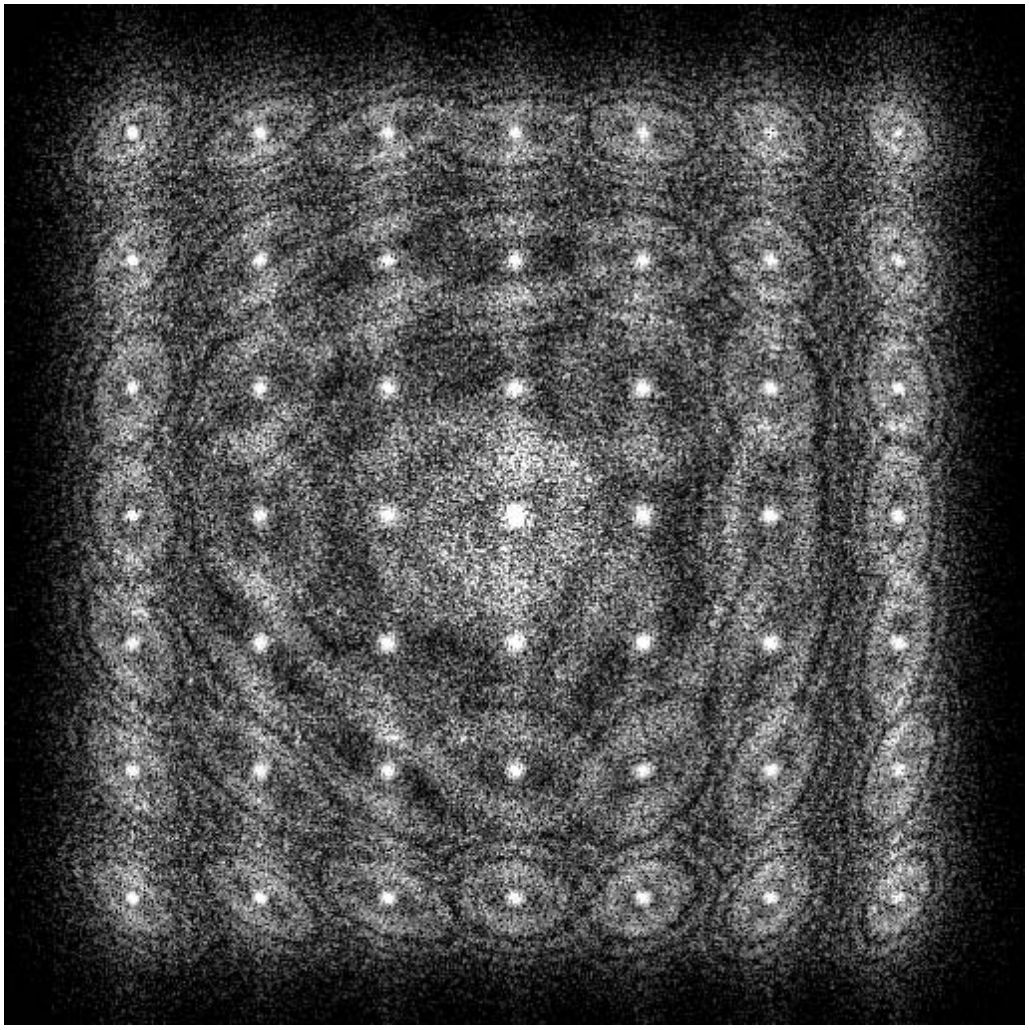
Aberration Determination and Compensation in High Resolution Transmission Electron Microscopy

Gopal Chand
Trinity College
Cambridge, U.K.

A Dissertation submitted to the University of Cambridge
for the degree of Doctor of Philosophy

July 1995



A diffractogram tableau from the JEM 4000 EXII electron microscope clearly showing the presence of *three-fold* astigmatism. For more details see section 4.9

Abstract

The main area addressed by this dissertation is the accurate measurement of the Transmission Electron Microscope (TEM) aberrations present in the objective lens. These measurements allow correct adjustment of the TEM which should improve the throughput of the microscope without so much dependence on highly skilled operators. They also provide precise information about imaging conditions essential for resolution improvement through image reconstruction.

Various techniques are described which can be used to determine TEM aberrations. Out of all these, the measurement of image shifts and diffractograms (power spectra) for various predetermined injected beam tilts hold most promise. It has proved possible to automate the measurement of image shifts; this however needs special filtering techniques at high resolution which are explored at length.

A number of large image data sets were collected to test the current imaging models critically. Serious discrepancies between predicted and actual data observations were found to be well accounted for by a hitherto neglected aberration, three-fold astigmatism. This aberration was found to be present at levels that cause serious image degradation and its value has been determined for one particular microscope.

Reconstruction methods are reviewed. The aberration determination methods developed have made viable for the first time a reconstruction method based on multiple beam tilt images which achieves close to a doubling of resolution. Some experimental results from the use of this approach are presented.

Preface

This dissertation describes a course of research carried out at the High Resolution Electron Microscopy Unit of the Department of Materials Science and Metallurgy and at the Department of Engineering in Cambridge University, England. Except where specifically acknowledged, it is entirely the result of my own work and includes nothing which is the result of work done in collaboration. This dissertation has not been submitted in whole or in part as consideration for any other degrees of qualification at this University or any other Institute of Learning. The dissertation is less than 65,000 words in length.

I would like to take this opportunity to thank my supervisor, Dr. D.M. Holburn of the Department of Engineering, for his supervision and support in this project. I am deeply indebted to Dr. W.O. Saxton of the Department of Materials Science and Metallurgy for his ideas and encouragement throughout the project and also his help in explaining electron microscopy concepts. My thanks are also due to Dr. A.I. Kirkland of the Chemistry Department who operated the microscope during the project and provided useful discussions, to Dr. K. Tsuno and Mr. M. Kawasaki of JEOL who were involved in the capture of micrographs in Japan, and to Professor C.J. Humphreys of the Department of Materials Science and Metallurgy for allowing me to have access to the JEOL 4000 EXII electron microscope. Particular thanks go to Dr. R.C. Thomas of the Department of Engineering who allowed me to use her powerful HP workstations for image processing. Help at the initial stages of the project from Dr. K.C.A. Smith and Dr. K.-L. Chau both formerly of the Department of Engineering is gratefully acknowledged. Finally I would like to thank my parents and my sister for their support and encouragement.

This work was supported by the Science and Engineering Research Council and the Royal Society Paul Instrument Fund to whom I am sincerely grateful.

Contents

List of Figures

List of Tables

List of Symbols and Abbreviations

A_0	Image position
A_1	Two-fold astigmatism (overfocus positive)
C_1	Defocus (overfocus positive)
A_2	Three-fold astigmatism
B_2	Axial coma
A_3	Four-fold astigmatism
B_3	Axial star
C_3	Third order spherical aberration
C	Cross-power spectral density
E	Even part of wave aberration function
F	Specimen object function
G	Coherence envelope
I	Image contrast
N	Noise
O	Odd part of wave aberration function
P	Weak phase object transfer function
T	Wavefunction transfer function
W	Wave aberration function (displacements)
$\mathbb{C}$	Sum of cross-terms in restoring filter
$\mathbb{S}$	Sum of squared terms in restoring filter
d	R.m.s. focus spread
$\mathbf{x}$	Space variable
$\mathbf{k}$	Fourier variable
$\mathbf{k}_0$	Beam tilt
$\mathfrak{T}$	Effective contrast transfer function

Ψ	Specimen exit plane wavefunction
β	Beam divergence
γ	Wave aberration function (phase shifts)
ϵ	Signal to noise weighting
λ	Electron wavelength
μ	Complex tilt from coma-free axis
ν	Image noise-to-signal ratio
τ	Complex tilt
τ_c	Coma-free axis position
ω	Complex diffraction angle

ACF	Auto-correlation function
CCD	Charge coupled device
CTF	Contrast transfer function
Gl	Glaser units : $\sqrt[4]{C_3\lambda^3}$
MTF	Camera modulation transfer function
PC-XCF	Phase compensated cross-correlation function
PHAT	Phase transform
PSF	Point spread function
Sch	Scherzer units : $\sqrt{C_3\lambda}$
SCOT	Smoothed coherence transform
SNR	Signal to noise ratio
ST-XCF	Squared transfer cross-correlation function
TCC	Tilt coil calibration
TD-XCF	Transfer divided cross-correlation function
TEM	Transmission Electron Microscope
XCF	Cross-correlation function
XPSD	Cross-power spectral density

Chapter 1

Introduction

1.1 The Purpose of Aberration Determination

It is a constant aim of High Resolution Transmission Electron Microscopy (HRTEM) to achieve higher resolution. This is necessary in order to interpret high-resolution TEM images with more confidence. Although computer aided TEM aberration determination is important as a step towards fully automated microscope adjustment and a consequent large improvement in the productivity of individual instruments, it has an even greater value in characterising the conditions in which micrographs are captured as a vital step in resolution improvement techniques.

1.2 Outline of the Dissertation

It is the purpose of this dissertation to describe aberration determination techniques as aids for microscope adjustment and image reconstruction. The rest of this chapter reviews the techniques normally used by microscope operators to aid manual adjustment of the TEM and the various techniques that have previously been explored for automated adjustment. Chapter 2 reviews previous configurations for computer controlled microscopy and describes the configuration used for the current project. Chapter 3 provides a detailed analysis of objective lens aberrations and their effect and shows

how linear and non-linear parameter fitting techniques can be used to determine these aberrations given observations of either image displacements or diffractograms (power spectra). The chapter also covers various techniques for image displacement and diffractogram parameter estimation. Three-fold astigmatism is included in the analysis. Chapter 4 presents extensive results from the fitting of parameters using displacement and diffractogram measurements. The practical aspects of using phase compensating filters in displacement estimation, given limited *a priori* information about the imaging conditions, are also investigated. The chapter ends with the presentation of evidence for a significant level of three-fold astigmatism in the TEM available to this project. Chapter 5 reviews image reconstruction methods and examines the potential of tilt series reconstruction compared with conventional focus series reconstruction techniques. Experimental results from the use of the tilt series technique are presented. One of these shows a significant improvement in resolution. Finally, chapter 6 draws conclusions and suggests future work.

1.3 Uncontrollable Microscope Parameters

There are three microscope aberrations or parameters which are not usually under the direct control of the microscopist but whose value it is helpful to determine: accelerating voltage, spherical aberration and focus spread.

Accelerating voltage is likely to be within 0.5% of its stated value which is approximately 2kV for the JEOL 4000 EXII microscope operating at an accelerating voltage of 400 kV. Barry (Barry, J.C., 1992) notes that this level of error is not significant and that it is possible to measure accelerating voltage very accurately if required.

The manufacturer usually supplies a value for the spherical aberration for the microscope which can be used if standard objective lens currents are applied. The value for the JEOL 4000 EXII microscope is 0.9mm. There are, however, two techniques available for the measurement of spherical aberration: the diffractogram method and the displacement method. Krivanek's diffractogram method (Krivanek, O.L., 1976) involves measuring the position of rings in the diffractogram of an amorphous specimen with no injected beam tilt and using graphical techniques to determine an estimate for the aberration. Budinger and Glaeser (Budinger, T.F. and Glaeser, R.M., 1976)

suggested that spherical aberration can be measured by examining the specimen image displacement which results when a large beam tilt is injected. This method is prone to errors because it ignores the effect of possible initial misalignment. Furthermore, the distortion suffered by images when a large tilt has been injected makes the displacement estimate itself unreliable in practice. The diffractogram method is more tolerant of initial misalignment although the noisy nature of diffractograms makes the measurement of ring positions difficult.

Focus spread embodies lens current, voltage fluctuations and emission energy spread. It increases with beam current so that it is not invariant like accelerating voltage and spherical aberration. It is defined by:

$$d = C_c\sqrt{(\tfrac{\Delta V}{V})^2 + 4(\tfrac{\Delta I}{I})^2 + (\tfrac{\Delta E}{V})^2} \tag{1.1}$$

where C_c is the coefficient of chromatic aberration, $\frac{\Delta V}{V}$ is the fractional lens voltage fluctuation, $\frac{\Delta I}{I}$ is the fractional lens current fluctuation and $\frac{\Delta E}{V}$ is the fractional emission energy spread. Based on typical values for the JEOL 4000 EXII microscope, the value of d derived from equation (1.1) is 6.0nm. The most reliable method for estimating focus spread is the measurement of the microscope information limit (described in section 5.1) using diffractograms at low values of beam divergence. Although the information limit is theoretically defined in terms of the damping produced by a negative exponential, its experimental definition is the point on a diffractogram where the transfer envelope is reduced to a set value which is normally 5% or 10% in practice. The focus spread, which is proportional to the square of the information limit, can thus be derived (Frank, J., 1974), though not with high accuracy.

1.4 Adjustable Microscope Parameters

Defocus, two-fold astigmatism, beam-tilt and beam divergence are the important microscope parameters which are under operator control. Defocus is measured relative to Gaussian defocus which is defined as 0 Sch [1]. The two-

[1] Generalised units such as Sch and Gl/Sch are defined on page xvii and described in (Hawkes, P.W., 1980)

fold astigmatism in the objective lens should be effectively zero which usually means less than 6.0 nm (0.16 Sch) which corresponds to a 45° maximum phase shift for a 2 Å beam. Beam tilt, like two-fold astigmatism, should be effectively zero which usually means less than 0.4 mrad (0.06 Gl/Sch) which also corresponds to a 45° maximum phase shift for a 2 Å beam. Beam divergence is set by the condenser aperture radius and condenser lens strength. For images from the JEOL 4000 EXII, two values are typically used for the diameter of this aperture: CA3 (50 μm) and CA4 (20 μm) which correspond to about 0.9 mrad and 0.4 mrad of divergence respectively.

Defocus can be measured from the diffractograms or texture of amorphous support films or contamination in the vicinity of the object being examined. Operators can determine Gaussian defocus by finding the defocus value at which image contrast is at a minimum. This can then be used as the reference value for subsequent microscope work. Two-fold astigmatism is determined by examining fringe isotropy around the edge of a hole or by assessing the texture of amorphous support films or contamination visually at high resolution. Use of beam tilt will be described in section 1.5.2. Barry (Barry, J.C., 1992) describes crystal tilt, intermediate lens astigmatism and objective aperture radius in addition to the parameters mentioned above.

1.5 Manual Adjustment Techniques

1.5.1 Current and voltage centre alignment

It should be noted that alignment generally refers to beam tilt and is therefore distinct from adjustment described in the previous section. The earliest adjustment technique for the TEM was LePoole's 'wobbler' (LePoole, J.B., 1947) in which the parallax caused by tilting the illuminating beam produced an image shift which could be detected by the operator. This allowed it to be used as a focussing aid. Since then, current centre and voltage centre alignment procedures (Hines, R.L., 1984) have been considered adequate for low and medium resolution work (4-100 Å).

Current centre alignment involves varying the current to the objective lens. Image rotation and displacement occur as a result and beam tilt is adjusted until the stationary point in the image is brought to the centre of the viewing screen.

Voltage centre or dispersion-free alignment consists of superposing a small alternating voltage on the high tension voltage supply. This changes the energy of the electrons and is therefore equivalent to changing all the lens currents. The resulting point of expansion and contraction is brought to the centre of the viewing screen.

The voltage and current centres are usually distinct; apart from anything else, high tension voltage changes additionally affect all other lenses which are not usually perfectly coaxial with the objective. The voltage centre is usually preferred because of hysteresis problems and objective lens instability associated with current centre alignment. Furthermore, in most microscopes, fluctuations in the high tension supply lead to larger displacements in microscope images than instabilities in the objective lens current.

1.5.2 Coma-free alignment

In 1978, Zemlin et al. (Zemlin, F. et al., 1978) suggested that axial coma was important for high resolution work (<4 Å). They were able to determine values of two-fold and three-fold astigmatism and axial coma from measurements of diffractograms in a tableau. It will be shown in section 3.5 that axial coma corresponds to a beam tilt or misalignment from the so called coma-free axis. The procedure of adjusting beam tilt so that coma is eliminated is known as coma-free alignment. Saxton and O'Keefe (Saxton, W.O. and O'Keefe, M.A., 1981) investigated the implication of misalignment and over the next decade, the manual location of the coma-free axis usually employed the $\pm x$ and $\pm y$ tilt alternator method of Smith et al. (Smith, D.J. et al., 1983). This involved injecting equal but diametrically opposite beam tilts in both the x- and y-directions and adjusting the beam tilt until the amorphous textures in the tilted images were roughly identical for both directions. It should be noted that coma-free alignment methods usually make use of the same beam deflector coils above the objective as the voltage centre method does. As a result coma-free alignment is obtained at the expense of voltage centre alignment unless beam deflector coils below the objective are used (Ishizuka, K. and Shirota, K., 1994).

1.6 Automated Adjustment Techniques

Erasmus and Smith (Erasmus, S.J. and Smith, K.C.A., 1982) and Saxton et al. (Saxton, W.O. et al., 1983) provide an overview of early alignment methods of which suitable candidates for automated adjustment such as the minimum variance method, diffractogram methods and the auto-correlation function method will be described. The displacement methods developed later by Van der Mast (van der Mast, K.D., 1984) and Koster et al. (Koster, A.J. et al., 1987) will also be described.

1.6.1 The minimum variance method

The minimum variance method makes use of the fact that TEM image variance is at a global minimum when the image is focused, stigmated and aligned (Saxton, W.O. et al., 1983). The method involves varying the alignment, two-fold astigmatism and defocus in turn and fixing the value of these parameters at a level for which the variance of the image is at a minimum. The five-dimensional minimum search can be effectively reduced to a series of five one-dimensional searches. The method requires a disordered specimen which satisfies the weak phase object criterion in order to work effectively. Another major drawback is the high level of irradiation suffered by the specimen because the number of images which need to be captured is in the region of 100 for each parameter. However, the method is conceptually and experimentally simple, self calibrating and has been successfully implemented.

1.6.2 Diffractogram methods

Methods which make use of diffractograms (power spectra) require a disordered, weak phase object just like the minimum variance method. The diffractogram of such a specimen exhibits a characteristic ring pattern, dependent on microscope imaging conditions such as defocus, two-fold astigmatism and misalignment. It has been the most widely used diagnostic tool since Thon (Thon, F., 1966) used it *a posteriori*. He showed how to calculate the defocus by measuring the positions of maxima and minima in the diffractogram from a disordered specimen. As mentioned in section 1.3, Krivanek (Krivanek, O.L., 1976) describes an extension of the method to measure the coefficient of spherical aberration as well. Zemlin et al. (Zemlin, F. et al.,

1978) further extended diffractogram measurement to elucidate all the first and second order aberrations as mentioned in section 1.5.2.

Erasmus (Erasmus, S.J., 1982) concluded that it is generally quite difficult to measure the characteristics of diffractograms automatically for focus and two-fold astigmatism correction because they suffer from various sources of noise. The object structure factor attenuates the higher frequencies and this, in conjunction with the noise and coherence envelopes, makes the judgement of ring positions in the diffractogram less reliable at these frequencies. Erasmus also pointed out that the large number of calculations required for the automatic procedure rendered the method infeasible with the computing power available at the time.

Automation of diffractogram measurement was introduced by Baba et al. (Baba, N. et al., 1987). However, their technique is incapable of working for a large set of microscope starting conditions, typically failing for diffractograms obtained at overfocus with large levels of two-fold astigmatism. Krivanek and Fan (Krivanek, O.L. and Fan, G.Y., 1992) showed how diffractograms with at least two rings can be fitted automatically and introduced a method based on the diffractograms of four tilted images. The procedure is normally performed at high magnification (above 100 KX) after an initial automated adjustment at low magnification using the beam-tilt-induced image displacement (BID) method described below. Three-fold astigmatism is omitted from their analysis of diffractograms and displacements so that their estimates of misalignment are likely to be in error if this aberration is significant. Krivanek and Leber (Krivanek, O.L. and Leber, M.L., 1994) include three-fold astigmatism in the estimation of misalignment using the four tilt diffractogram method thus correcting the deficiency in the Krivanek and Fan paper (Krivanek, O.L. and Fan, G.Y., 1992). However, they use measurements of effective two-fold astigmatism only and discard measurements of effective defocus. An analysis of the four tilt diffractogram method which takes effective defocus into account is presented in section 3.6.4.1.

In general, diffractogram methods require only 5-10 images in order to elucidate important aberrations making them attractive for low radiation dose specimens. The methods are not adversely affected by specimen drift at high magnification as are the displacement methods described below. Consequently, a manual diffractogram measurement procedure was implemented in the project to elucidate aberration values off-line (described in section 3.6.3). Results of off-line diffractogram fitting are presented in chapter 4.

1.6.3 The auto-correlation function method

Frank and Al-Ali (Frank, J., 1975), (Al-Ali, L. and Frank, J., 1980) cover the use of the cross-correlation function of two images from the same area of specimen as estimates for the auto-correlation function of the point spread function. Like the minimum variance and diffractogram methods, a disordered and weak object is required. Auto-correlation functions (ACFs) estimated by this method do not suffer from the effects of noise as do the diffractogram methods and make use of the same low number of images. The ACF contains exactly the same information as the diffractogram albeit expressed differently. This is its virtue and its weakness: the signal is concentrated and is clear even with low exposure but no way has yet been found of deducing defocus let alone two-fold astigmatism directly from the ACF.

1.6.4 Displacement methods

The displacement methods hold greatest promise as candidates for automated adjustment. They can cope with both strong and weak objects and are therefore less restrictive than the above methods. Van der Mast (van der Mast, K.D., 1984) describes how injecting beam tilt allows the determination of defocus after examining the resulting image displacement. Koster et al. (Koster, A.J. et al., 1987) extend the technique to determine two-fold astigmatism by injecting tilts in two directions. Theoretical limitations on the beam-tilt-induced image displacement (BID) method as they describe it are derived although the frequency-dependent image shifts due to beam tilt are ignored.

Chang (Chang, M.M.Y., 1988) shows how injecting a defocus step leads to a convenient estimation of misalignment which does not depend upon an estimate of two-fold astigmatism. However, injecting a defocus step involves changing the objective lens current which implies that the current centre (described in section 1.5.1) is found by this method. This is unlikely to coincide with the coma-free axis so a residual misalignment will still be present.

Koster et al. (Koster, A.J. et al., 1988) simulate the compensation of frequency-dependent image shifts by a phase correcting filter for small tilts and show how the cross-power spectrum can be used to estimate image displacement. More sophisticated phase compensation techniques are described in section 3.6.1.3 and the failure of the cross-power spectrum method due to

coherence envelopes is described in section 4.4.

Koster et al. (Koster, A.J. et al., 1989) further extend the BID and cross-power spectrum methods to measure misalignment based on three images: one formed without beam tilt and two with equal but opposite beam tilts. This technique is attractive because the frequency-dependent image shift can be cancelled out. However, three-fold astigmatism is missing from the analysis (covered in sections 3.6.1.4 and 3.6.2.1) and this is likely to undermine subsequent estimates of defocus and two-fold astigmatism.

Koster and de Ruijter (Koster, A.J. and de Ruijter, W.J., 1992) describe the results of applying the BID method to 100 kV and 400 kV microscopes. Automated calibration procedures are also introduced. Although they note that the reproducibility of the method is dependent upon specimen and microscope instrumentation, it is likely that some of the lack of reproducibility is due to the presence of three-fold astigmatism (described in section 3.6.2.1).

De Jong and Koster (de Jong, A.F. and Koster, A.J., 1992) review automated adjustment techniques. They also describe how the third-order polynomial fitting of the displacements which result from a series of injected tilts can allow for the estimation of spherical aberration, defocus, two-fold astigmatism and misalignment.

As mentioned above, Krivanek and Fan (Krivanek, O.L. and Fan, G.Y., 1992) use the BID method as a low magnification procedure and the automated fitting of diffractograms as the high magnification procedure for automated adjustment. This is because the displacement method suffers from specimen drift far more than the diffractogram method at high magnification.

Chau (Chau, K.-L., 1993) introduces a least squares technique for displacement fitting which is also described by Saxton (Saxton, W.O., 1994c). In this, images are collected for a number of predetermined injected tilts and aberrations are fitted to the over-determined set using a linearised least squares procedure; this allows a flexible compromise between speed and reliability. Chau performed simulations to investigate this procedure for various injected tilt configurations, and also cross-power spectral techniques.

As with diffractogram and auto-correlation function methods, only a small number of images need to be captured for displacement methods. Given the applicability of the latter to both strong and weak objects, they may be the ideal automated adjustment candidate. However, the crucial omission of three-fold astigmatism in all the previously described displacement methods combined with their susceptibility to specimen drift has meant that the

full potential of such methods has not been realised at high magnification. Consequently, detailed examination of various displacement techniques was carried out in the project. Results of the displacement method are presented in chapter 4.

Chapter 2

System Hardware and Software

2.1 Introduction

The fundamental configuration of a system for computer controlled microscope experiments has remained unchanged since the early 1980s. It consists of a microscope with an image pick up device connected to a framestore controlled by a computer. The computer analyses images captured by the framestore and injects compensating microscope lens currents by means of an interface device.

Improvements to this basic configuration have occurred in each of the components. One of the first configurations was that developed by Erasmus (Erasmus, S.J., 1982). He used a phosphor screen as the image pick up which is described by Catto et al. (Catto, C.J.D. et al., 1981). The framestore called GEMS was developed in a joint venture with the Computer Aided Design Centre of Cambridge and it was capable of storing 4 images of 512 x 1024 x 8 bits each. A 16-bit mini-computer, the Computer Automation LSI-4/90, was used to control this framestore. A custom interface device called the General Purpose Interface (GPI) was also controlled by the computer and was capable of controlling defocus and two-fold astigmatism in the Cambridge 600 kV microscope described by Cosslett et al. (Cosslett, V.W. et al., 1979).

Chang (Chang, M.M.Y., 1988) investigated the alternative use of a yttrium aluminium garnet (YAG) pick up via Monte Carlo simulations. He continued to use the GEMS framestore and the GPI but changed the computer to an Intel 8086 based Olivetti M24 Personal Computer (PC).

Chau (Chau, K.-L., 1993) used a Synoptics Synapse framestore capable of storing one 512 x 512 x 8 bit image. The card was inserted in the expansion bus of an IBM compatible 486 DX 33MHz PC. The other components in his configuration are identical to those used in the project and are described below. A useful overview of TEM interfacing is given by Downing et al. (Downing, K.H. et al., 1992).

The present project was carried out using the JEOL JEM-4000 EX-II electron microscope at the High Resolution Electron Microscope unit in the Department of Materials Science and Metallurgy at Cambridge University. A Synoptics Sprynt system was used as the integrated computation, image capture and display device. The system configuration is shown in figure 2.1, and is described more fully in the next section.

2.2 System Components

2.2.1 The microscope

The JEOL JEM-4000 EX-II is a 400kV microscope equipped with a LaB_6 filament. It has a top entry goniometer and a low coefficient of spherical aberration (0.9mm) for high resolution work. An on-board computer controls the lens currents and there is an external serial port for connection to a remote computer. In addition to entering commands from the remote computer via the serial interface, it is possible to enter commands using a terminal which is part of the microscope.

2.2.2 The camera interface

The purpose of the camera unit is to convert the electron intensity distribution in the TEM into a video signal which can be digitised. A Gatan 622 Mk-II video camera with an electron photomultiplier is directly coupled to a single crystal (YAG) screen with optical fibre. The camera output is connected to a filter unit. This allows averaging of the images as well as contrast stretching.

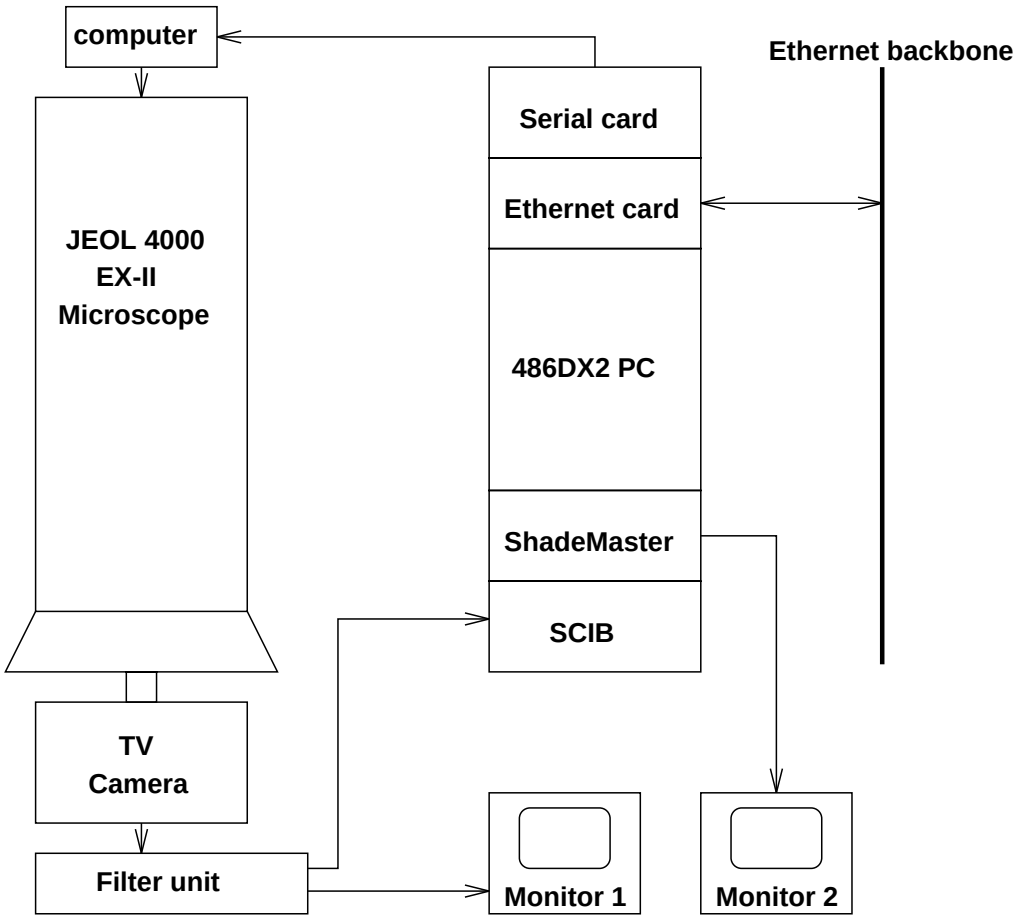


Figure 2.1: System configuration

2.2.3 The computer

The computer is an IBM compatible 486 DX2 66MHz PC running MSDOS. The detailed internal configuration of the PC with the Sprynt system is shown in figure 2.2. The following cards are installed in the PC:

- VGA card
- Floppy drive and 400Mb hard drive controller
- 16-bit ethernet card

- Multiple serial port card
- The Sprynt system

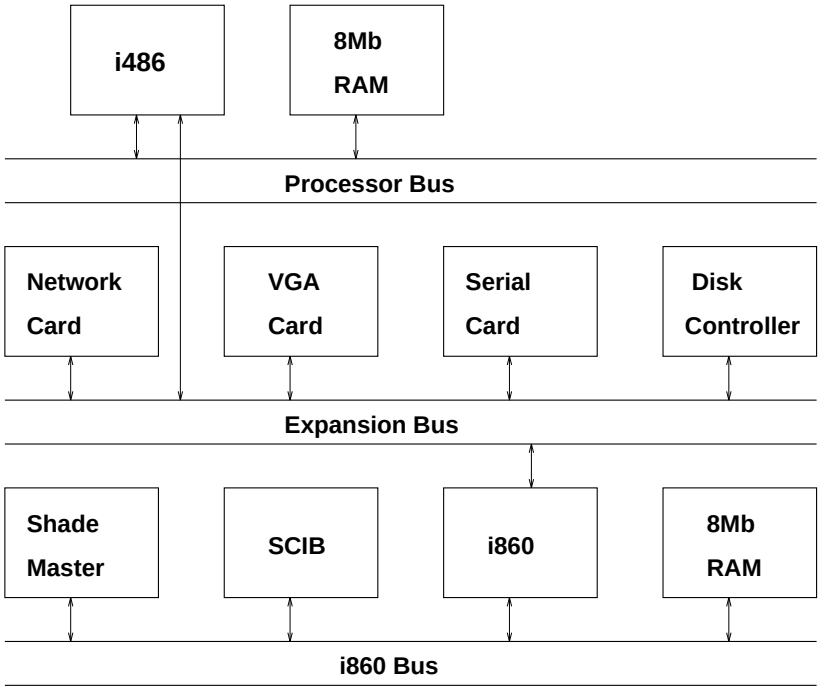


Figure 2.2: PC configuration

The multiple serial port card allows enables access to up to four serial ports. The Sprynt system hardware consists of a 24 bit true-colour framegrabber, true-colour framestore and an Intel *i*860 microprocessor board. These three cards are connected to each other by the high speed (200Mbps) *i*860 bus.

2.2.3.1 The framegrabber

The framegrabber is a 2Mb Synoptics Colour Input Board (SCIB) capable of capturing both analogue and digital images. Analogue colour images up to PAL resolution (768 x 576) can be captured at a digitisation depth of 24 bits provided they have been split into red, green and blue components. Digital images can be captured at pixel rates of up to 20MHz. For this project, this

board was used to grab analogue grey scale images (8 bits) of 512 x 512 from the PAL source of the image filter.

2.2.3.2 The framestore

The framestore is a Myriad ShadeMaster capable of displaying 8-bit images at resolutions of up to 1280 x 1024 or 24-bit true-colour images at resolutions of up to 800 x 600.

2.2.3.3 The *i*860 microprocessor board

The *i*860 is a high performance RISC microprocessor which is capable of working up to 80 Million Floating Point operations a second (MFLOPs). It has access to 8Mb of memory on its high speed data bus which is independent of the host PC. The framegrabber and framestore have direct access to the memory via the bus. Table 2.1 shows the speed of the *i*860 in performing 2 dimensional fast Fourier transforms (FFTs) and cross-correlation functions (XCFs) compared with other RISC microprocessor based computers available.

	FFT		XCF	
	256^2	512^2	256^2	512^2
*i*860	0.49	1.81	2.30	7.69
SG Indigo	0.33	1.69	1.11	5.53
HP 735/125	0.15	0.85	0.60	3.05

Table 2.1: Comparison of computation speeds in seconds

2.3 Software

2.3.1 Low level software

The *i*860 runs a Unix-like operating system called Xnix. This is a Terminate-and-Stay-Resident utility which allows many standard Unix applications to be executed on the *i*860 whilst the PC is running MSDOS. Xnix sleeps until

a Unix development tool or the *i*860 requires servicing, whereupon it wakes up and performs the required service.

2.3.2 Image processing software

Semper (Saxton, W.O. et al., 1979), an image processing package supplied as part of the Sprynt system, runs on the *i*860 and utilises both the framegrabber and the framestore. It has an easy-to-use command line interface and allows rapid prototyping of new image processing algorithms. The algorithms are either built up from existing Semper commands or added as faster Fortran or C code.

2.4 System Performance

2.4.1 Noise reduction

The filter unit is capable of reducing noise present in the TEM by continuous updating of the stored frame, S_n, according to:

$$S_{n+1} = \alpha S_n + (1 - \alpha)V_n \qquad 0 < \alpha < 1 \tag{2.1}$$

where V_n is the nth incoming video frame ('recursive' filtering with respect to time). A value of α closer to unity effectively averages over a larger number of input video frames. The filter unit is marked by the manufacturer with a number of frames averaged without any precise definition. The signal-to-noise ratio (SNR) should increase with this number of frames, N, which is fixed by the filter unit at 1,2,4,8,16,32,64 or 128. This improvement can be investigated by finding the ratio of the SNR with and without filtering.

The SNR improvement is in practice limited by specimen and focal drift, which cause image blurring for large values of N. If it is assumed that specimen drift is linear, its effect can be modeled by convolving the image, $f(\mathbf{x})$, with a rectangular function, $h(\mathbf{x})$:

$$g(\mathbf{x}) = f(\mathbf{x}) * h(\mathbf{x}) \tag{2.2}$$

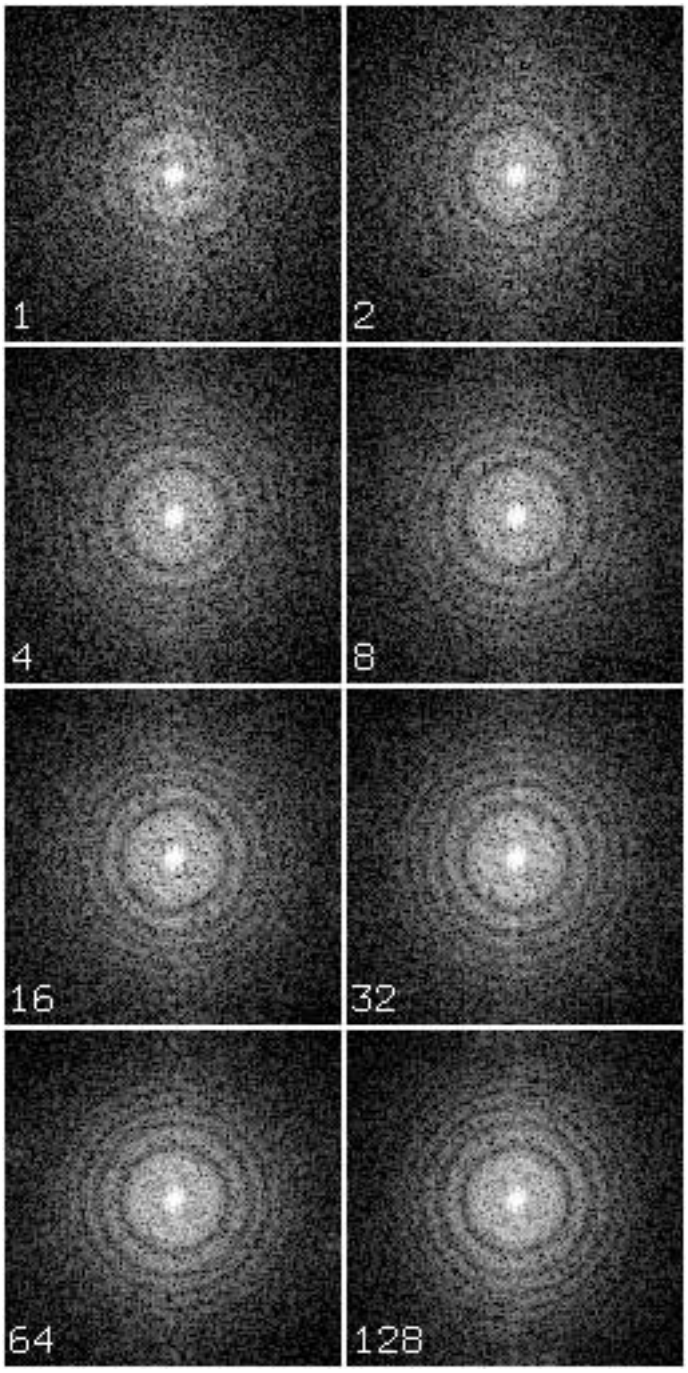


Figure 2.3: Diffractograms for different values of the number of averaged frames, N. The value chosen for subsequent experiments is $N = 8$ as this is the smallest value which leads to the presence of all the rings in the diffractogram. The images from which these diffractograms are obtained are those of an amorphous germanium film at an arbitrarily large underfocus value and no injected tilt. The remaining microscope parameters are identical to those shown in figure A.1

where the $*$ denotes convolution and $h(\mathbf{x})$ is a 1-D rectangle of width $\mathbf{r}$ and unit height. The Fourier Transform of $h(\mathbf{x})$ is:

$$H(\mathbf{k}) = \frac{\sin \pi \mathbf{k}.\mathbf{r}}{\pi \mathbf{k}} \tag{2.3}$$

This has the effect of attenuating high frequencies and thus imposes an upper bound on N to which $\mathbf{r}$ is proportional.

In order to find an optimum N, images were collected with different values of N at high defocus to ensure a rapidly oscillating contrast transfer function i.e. many rings. The resulting diffractograms of these images are shown, labeled with N, in figure 2.3. For these images, the specimen drift is very small and is in a direction approximately $\frac{\pi}{4}$ to the horizontal.

The SNR can be estimated by examining the peak height in the cross-correlation of two image captured with the same value of N (Saxton, W.O., 1978):

$$\text{SNR} = \sqrt{\frac{\rho}{1-\rho}} \tag{2.4}$$

where ρ is the cross-correlation peak height.

Figure figure 2.4 plots SNR, which is estimated from cross-correlation peak heights using equation (2.4), against N. As expected, the SNR increases with N and a value of 8 for N was chosen as it is the smallest value which leads to the presence of all the rings in the diffractogram.

2.4.2 Computation

The *i*860 compares favourably with current RISC based workstations (table 2.1). The Silicon Graphics Indigo is a MIPS R4000 machine and the Hewlett Packard 735/125 is a PA-RISC machine. The times for the XCFs are greater than those for three FFTs because of the overhead due to the complex multiplication.

The speed of the XCF calculation is the most important factor in determining how fast displacement estimation methods (described in section 3.6.1) can provide aberration estimates for possible real-time control. In order to decrease the time taken to calculate XCFs, the use of a number of powerful

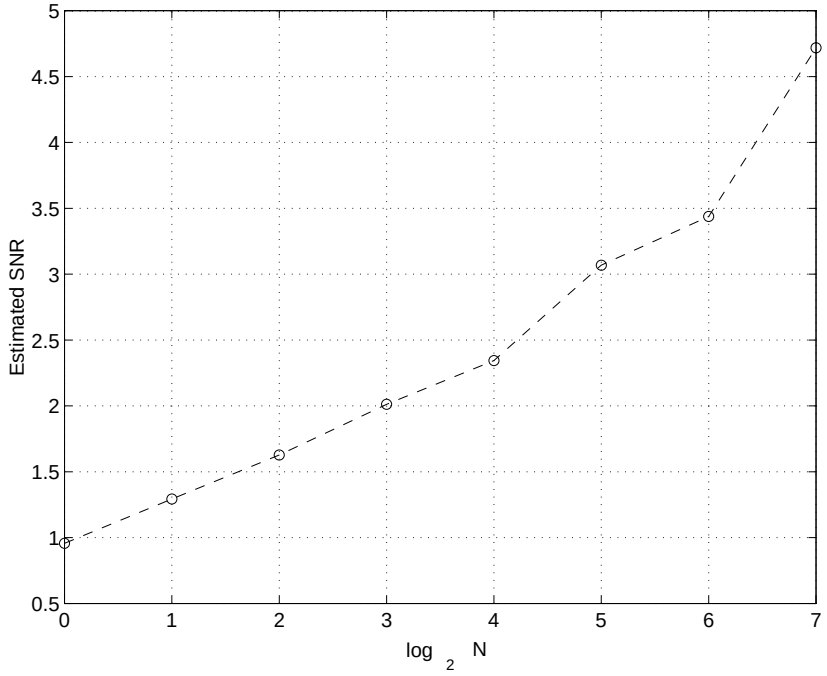


Figure 2.4: A graph of estimated SNR for different values of the number of averaged frames, N.

workstations in parallel over a network (Chand, G. and Holburn, D.M., 1994) was investigated as an alternative to the *i*860 based system. The results are presented in appendix C.

Chapter 3

Theory of Image Formation and Analysis

3.1 Introduction

This chapter provides a detailed analysis of objective lens aberrations and their effect on imaging at high resolution. The analysis includes the effect of a tilt coil calibration matrix and the hitherto neglected aberrations: three-fold and four-fold astigmatism and axial star aberration. Linear and non-linear parameter fitting techniques are derived which can be used to determine the aberrations given observations of either image displacements or diffractograms (power spectra). Techniques for improving the accuracy of displacement estimation are also described.

3.2 The Wave Aberration Function

3.2.1 Ideal Gaussian imaging

Consider the imaging of a one-dimensional lattice grating with period d as shown in figure 3.1. The layout is similar to that encountered in the imaging system of a Transmission Electron Microscope (TEM) in which case the object is the specimen, the converging lens is the objective and the image plane is the fluorescent screen (Misell, D.L. and Brown, E.B., 1987). The optic axis of the lens is CC' and the lens focal length is f. The four parallel

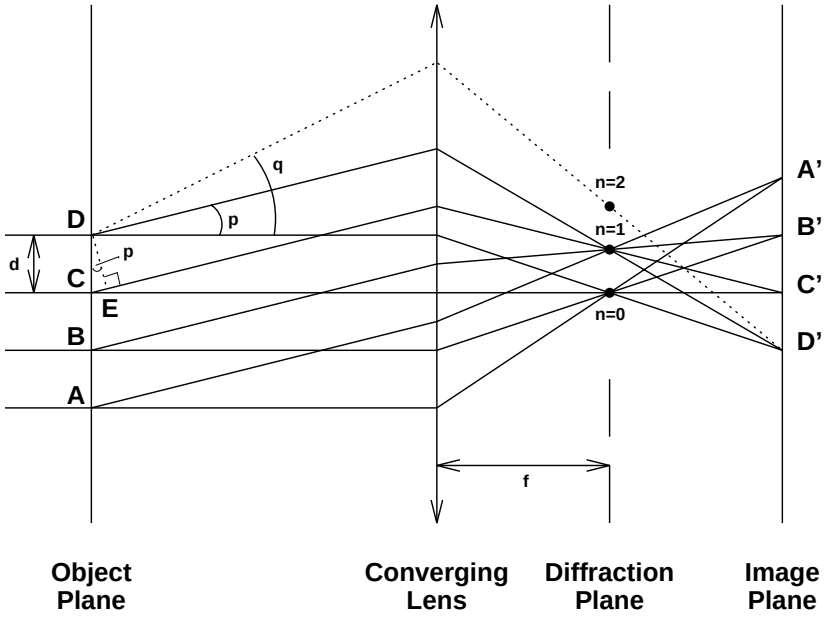


Figure 3.1: Lens Diffraction

rays leaving the object at angle p to the optic axis are combined by the lens into a single point in the diffraction plane. In the TEM, these rays would be normal to the electron wavefronts. If CE $= \lambda$ where λ is the electron wavelength, then these beams will combine constructively in the diffraction plane. In general, constructive interference will occur when CE $= n\lambda$ where n is an integer. The figure shows the diffraction spots for $n = 0, 1$ and 2 which correspond to ray angles $\theta = 0, p$ and q. When $\theta = p$ and p is small then, because the angle $\angle$CDE is also p, we derive $p = \frac{\lambda}{d} = \lambda k_p$ where k_p is a frequency which corresponds to the *diffraction angle*, p, and the lattice grating period, d.

Figure 3.2 is a three dimensional extension of part of the system in figure 3.1. The lens has been omitted for clarity. The diffraction angles, x and y, define a vector, $\mathbf{k}$ in the diffraction plane for a *general* two-dimensional lattice grating in the object plane. If the components of the lattice grating are d_x and d_y in the x- and y-direction respectively, then $d_x = \frac{\lambda}{x}$ and $d_y = \frac{\lambda}{x}$. If OA = L and both x and y are small then AB $= L.x = L.\frac{\lambda}{d_x} = L.\lambda k_x$ and BC $= L.y = L.\frac{\lambda}{d_y} = L.\lambda k_y$ where k_x and k_y are components of $\mathbf{k}$.

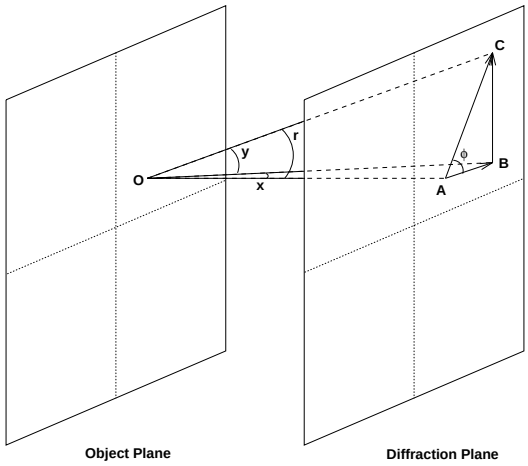


Figure 3.2: The diffraction vector

3.2.2 The effect of aberrations

All real lenses suffer from aberrations which cause the rays illustrated in figure 3.2 to deviate from their ideal directions. A point object at (x_0, y_0) in the object plane ideally leads to a spherical wavefront in the diffraction plane contracting to a conjugate image point in the image plane. The *wave aberration function*, W, measures the distance between the ideal spherical surface and the actual diffractogram plane wavefront. In general, W can be expanded in powers of x_0, y_0, x and y where the latter are the diffraction angles as illustrated in figure 3.2. At sufficiently high magnification, the object region being examined is so small that terms.m.s. in x_0 and y_0 can be neglected. This implies that W can be expressed in terms of aperture aberrations only:

$$\begin{aligned} W(x,y) =& c_1 x + c_2 y \\ &+ c_3 x^2 + c_4 xy + c_5 y^2 \\ &+ c_6 x^3 + c_7 x^2 y + c_8 x y^2 + c_9 y^3 \\ &+ c_{10} x^4 + c_{11} x^3 y + c_{12} x^2 y^2 + c_{13} x y^3 + c_{14} y^4 \end{aligned} \tag{3.1}$$

A complex position variable, $(\omega = x + iy)$, can be introduced to simplify subsequent algebra as there is no vector equivalent of complex conjuga-

tion. This allows W to be re-expressed in a form where important terms are grouped together (Typke, D. and Dierksen, K., 1992):

$$\begin{aligned} W(\omega) = \Re\big(&A_0\omega^* + \tfrac{1}{2}A_1\omega^{*2} + \tfrac{1}{2}C_1\omega^*\omega \\ &+ \tfrac{1}{3}A_2\omega^{*3} + \tfrac{1}{3}B_2\omega^{*2}\omega \\ &+ \tfrac{1}{4}A_3\omega^{*4} + \tfrac{1}{4}B_3\omega^{*3}\omega + \tfrac{1}{4}C_3\omega^{*2}\omega^2\big) \end{aligned} \tag{3.2}$$

The * denotes the complex conjugation. This equation can be written out in Cartesian form where $A_0 = A_{0r} + iA_{0i}$ etc.:

$$\begin{aligned} W(x,y) =&A_{0r}x + A_{0i}y \\ &+ \tfrac{1}{2}A_{1r}(x^2-y^2) + \tfrac{1}{2}A_{1i}2xy + \tfrac{1}{2}C_1(x^2+y^2) \\ &+ \tfrac{1}{3}A_{2r}(x^3-3xy^2) + \tfrac{1}{3}A_{2i}(3x^2y-y^3) \\ &+ \tfrac{1}{3}B_{2r}x(x^2+y^2) + \tfrac{1}{3}B_{2i}y(x^2+y^2) \\ &+ \tfrac{1}{4}A_{3r}[(x^2-y^2)^2 - 4x^2y^2] + \tfrac{1}{4}A_{3i}4xy(x^2-y^2) \\ &+ \tfrac{1}{4}B_{3r}(x^4-y^4) + \tfrac{1}{4}B_{3i}2xy(x^2+y^2) + \tfrac{1}{4}C_3(x^2+y^2)^2 \end{aligned} \tag{3.3}$$

and in polar form where $A_0 = |A_0|e^{i\phi_{11}}$, $A_1 = |A_1|e^{2i\phi_{22}}$ etc.:

$$\begin{aligned} W(r,\phi) =&|A_0|r\cos(\phi-\phi_{11}) \\ &+ \tfrac{1}{2}|A_1|r^2\cos 2(\phi-\phi_{22}) + \tfrac{1}{2}C_1r^2 \\ &+ \tfrac{1}{3}|A_2|r^3\cos 3(\phi-\phi_{33}) + \tfrac{1}{3}|B_2|r^3\cos(\phi-\phi_{31}) \\ &+ \tfrac{1}{4}|A_3|r^4\cos 4(\phi-\phi_{44}) + \tfrac{1}{4}|B_3|r^4\cos 2(\phi-\phi_{42}) + \tfrac{1}{4}C_3r^4 \end{aligned} \tag{3.4}$$

A_0 simply determines image position as it describes a linear phase variation in the Fourier plane. C_1 is the *defocus* where overfocus is defined as positive. C_3 is the third order *spherical aberration*.

Looking at the polar form, it is possible to recognise A_1 as *two-fold astigmatism* which is a two-fold azimuthal variation of defocus. This way of measuring astigmatism is convenient because changes in the real and imaginary parts are proportional to changes in the two stigmator currents apart from an overall rotation. A_2 has three maxima w.r.t. azimuth and is known as

three-fold astigmatism. B_2 is *axial coma*, having only a single maximum. A_3 is *four-fold astigmatism* with four maxima w.r.t. azimuth and B_3 is *axial star aberration* with two such maxima. Figures 3.3 to 3.10 are three-dimensional plots of W against x and y showing the form that each aberration takes. The accelerating voltage is 400 kV and the spherical aberration is 0.9 mm giving Sch $\approx$ 38 nm and Gl/Sch $\approx$ 6.5 mrad. Figure 3.11 shows the form of a typical wave aberration function, combining the effects of two-fold astigmatism, defocus and spherical aberration.

Until the late 70s, the only aberrations considered to be significant in electron microscope image formation were two-fold astigmatism (A_1), defocus (C_1) and third order spherical aberration (C_3). In 1978, Zemlin et al. (Zemlin, F. et al., 1978) deduced that axial coma (B_2) was predominantly responsible for anomalies in the diffractogram tableaus from a stigmated and focused TEM. It will be shown later that axial coma corresponds to a beam misalignment from the so called coma-free axis.

It has recently been reported that unless the effect of three-fold astigmatism (A_2) is corrected, 1 Å resolution, although achievable, will lead to images which are difficult to interpret as a consequence of phase shifts introduced by the aberration (Krivanek, O.L. and Leber, M.L., 1993; Krivanek, O.L. and Leber, M.L., 1994; Saxton, W.O. et al., 1994; Saxton, W.O., 1994b).

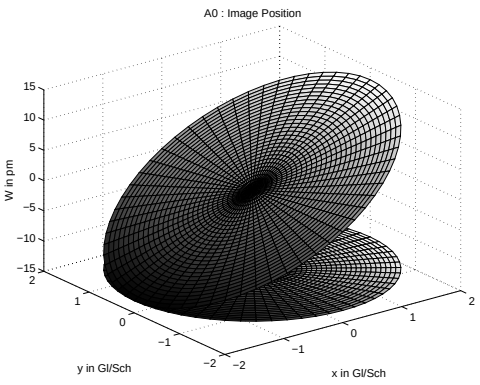


Figure 3.3: A_0=1 nm

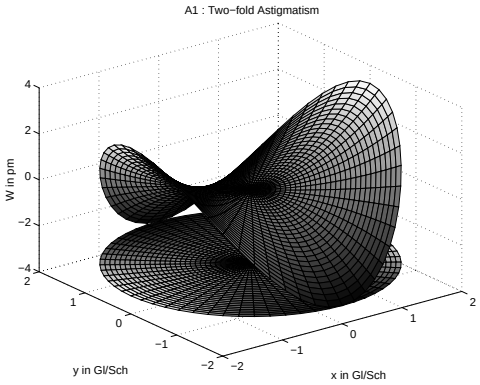


Figure 3.4: A_1=1 Sch

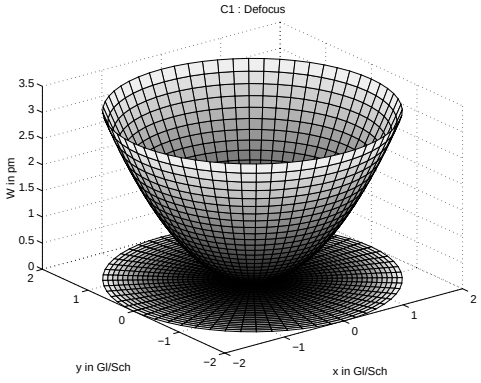


Figure 3.5: C_1=1 Sch

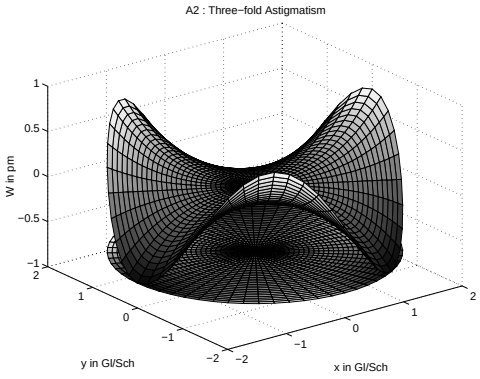


Figure 3.6: A_2=1 μm

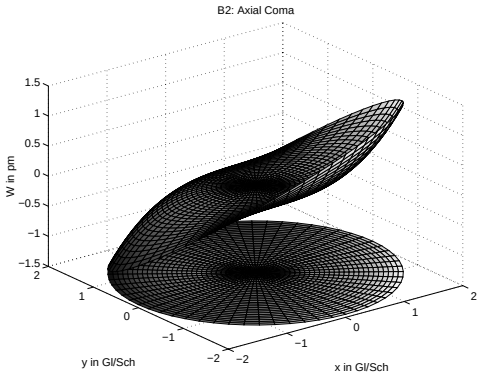


Figure 3.7: B_2=2 μm

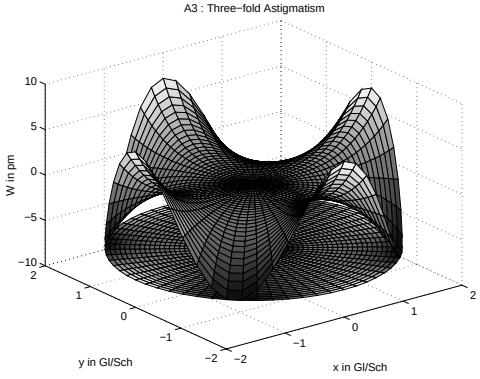


Figure 3.8: A_3=1 mm

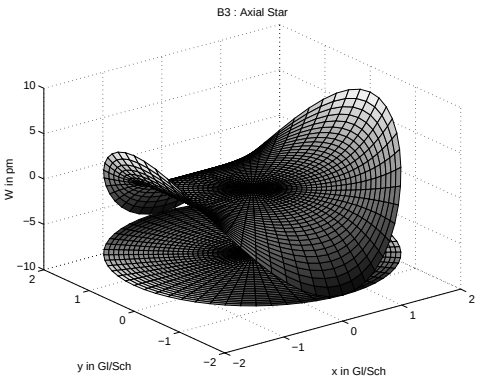


Figure 3.9: B_3=1 mm

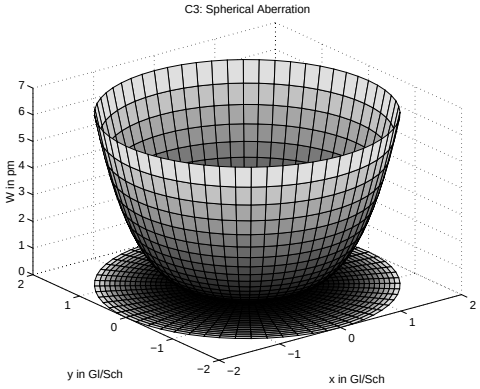


Figure 3.10: C_3=0.9 mm

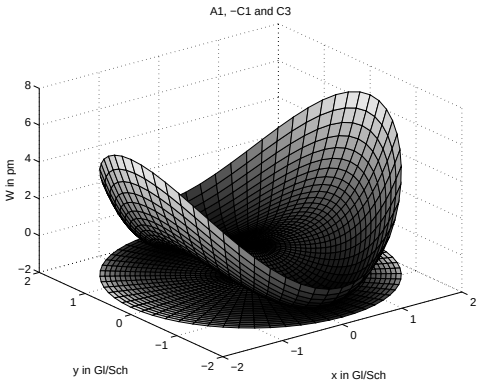


Figure 3.11: Combined effect of $A_1, -C_1$ and C_3. Figures 3.3 to 3.11 illustrate the effect of various aberrations on the wave aberration function.

3.3 The Contrast Transfer Function

Contrast transfer theory has been developed by Thon (Thon, F., 1966) to assess the effect of the aberrations described above upon the final image contrast. Hanszen (Hanszen, K.-J., 1971) reviews image transfer theory including strong and weak object approximations. Hanszen and Trepte (Hanszen,

K.-J. and Trepte, L., 1971) and Frank (Frank, J., 1974) cover the effect of coherence envelopes in axial imaging. Tilted imaging is covered by MacFarlane (McFarlane, S.C., 1975), Jenkins and Wade (Jenkins, W.K. and Wade, R.H., 1977; Wade, R.H. and Jenkins, W.K., 1978) and Saxton (Saxton, W.O., 1978).

In the weak scattering approximation (Saxton, W.O., 1978), the scattered wave is weak compared with the transmitted wave everywhere across the specimen. This implies that any terms in image intensity involving the square of the scattered wave can be ignored. Under these circumstances, the image contrast, i.e. the fractional intensity deviation, is related to the scattered wave by a linear mechanism.

Using vector notation, a perfectly monochromatic plane wave incident normally on a weak specimen produces an exit wave of the form:

$$\psi_e(\mathbf{x}) = 1 + \psi(\mathbf{x}) \tag{3.5}$$

The two terms on the RHS represent the transmitted primary beam and the scattered wave respectively ($|\psi| \ll 1$ everywhere for the weak scattering approximation).

For *tilted* illumination, the Fourier components of ψ are related to those of micrograph contrast, i, as follows (Saxton, W.O., 1978):

$$I(\mathbf{k}, \mathbf{k}_0) = \Psi(\mathbf{k})T(\mathbf{k}, \mathbf{k}_0) + \Psi^*(-\mathbf{k})T^*(-\mathbf{k}, \mathbf{k}_0) \tag{3.6}$$

where second order terms in ψ have been neglected (the *weak object approximation*) and $T(\mathbf{k})$ is the wavefunction transfer function or contrast transfer function (CTF).

$$T(\mathbf{k}) = G(\mathbf{k}_0 + \mathbf{k}, \mathbf{k}_0) \exp[-i(\gamma(\mathbf{k}_0 + \mathbf{k}) - \gamma(\mathbf{k}_0))] \tag{3.7}$$

$$\gamma(\mathbf{k}) = \frac{2\pi W(\mathbf{k})}{\lambda} \tag{3.8}$$

$$G(\mathbf{k}_1, \mathbf{k}_2) = \exp[-\tfrac{\beta^2}{4}|\nabla\gamma(\mathbf{k}_1) - \nabla\gamma(\mathbf{k}_2)|^2 \exp[-\tfrac{1}{2}\pi^2 d^2(\mathbf{k}_1^2 - \mathbf{k}_2^2)^2] \tag{3.9}$$

$G(\mathbf{k}_0 + \mathbf{k}, \mathbf{k}_0)$ is the coherence envelope term (Saxton, W.O., 1988) and it models the effect of source spatial and temporal coherence. $\mathbf{k}_0$ measures the

illumination tilt relative to the optical axis and λ is the electron wavelength. β is the beam divergence and d is the r.m.s. focus spread.

If the object is *weak phase* then the $\psi(\mathbf{x})$ is entirely imaginary:

$$\psi(\mathbf{x}) = e^{i\phi(\mathbf{x})} \approx i\phi(\mathbf{x}) \qquad |\phi(\mathbf{x})| \ll 1 \tag{3.10}$$

where $\phi(\mathbf{x})$ is the object phase shift. This means that $\Psi(-\mathbf{k}) = -\Psi^*(\mathbf{k})$ so that equation (3.6) becomes:

$$I(\mathbf{k}, \mathbf{k}_0) = \Psi(\mathbf{k})(T(\mathbf{k}, \mathbf{k}_0) - T^*(-\mathbf{k}, \mathbf{k}_0)) \tag{3.11}$$

ignoring the coherence envelope term (a reasonably close approximation for the lower spatial frequencies).

$W(\mathbf{k})$ and hence $\gamma(\mathbf{k})$ can be derived from equation (3.4) by using the substitutions:

$$\begin{aligned} r &= \lambda.|\mathbf{k}| \\ \phi &= \angle\mathbf{k} \end{aligned} \tag{3.12}$$

as in figure 3.2 so that from equation (3.8):

$$\begin{aligned} \gamma(\mathbf{k}) = 2\pi(&A_0\tfrac{\mathbf{k}}{\lambda} + \tfrac{1}{2}A_1\lambda|\mathbf{k}|^2\cos 2(\phi - \phi_{22}) + \tfrac{1}{2}C_1\lambda|\mathbf{k}|^2 \\ &+ \tfrac{1}{3}A_2\lambda^2|\mathbf{k}|^3\cos 3(\phi - \phi_{33}) + \tfrac{1}{3}B_2\lambda^2|\mathbf{k}|^3\cos(\phi - \phi_{31}) \\ &+ \tfrac{1}{4}A_3\lambda^3|\mathbf{k}|^4\cos 4(\phi - \phi_{44}) + \tfrac{1}{4}B_3\lambda^3|\mathbf{k}|^4\cos 2(\phi - \phi_{42}) + \tfrac{1}{4}C_3\lambda^3|\mathbf{k}|^4) \end{aligned} \tag{3.13}$$

Under tilted illumination, the variation of $\gamma(\mathbf{k})$ around the primary beam position can be separated into even and odd parts:

$$\gamma(\mathbf{k}_0 + \mathbf{k}) = \gamma(\mathbf{k}_0) + E(\mathbf{k}, \mathbf{k}_0) + O(\mathbf{k}, \mathbf{k}_0) \tag{3.14}$$

Therefore, from equation (3.11):

$$I(\mathbf{k}, \mathbf{k}_0) = \Psi(\mathbf{k})(e^{-i[E(\mathbf{k},\mathbf{k}_0)+O(\mathbf{k},\mathbf{k}_0)]} - e^{i[E(\mathbf{k},\mathbf{k}_0)-O(\mathbf{k},\mathbf{k}_0)]}) = \Phi(\mathbf{k}).P(\mathbf{k}, \mathbf{k}_0) \tag{3.15}$$

where $P(\mathbf{k}, \mathbf{k}_0)$ is the weak phase object transfer function defined by

$$\boxed{P(\mathbf{k}, \mathbf{k}_0) = e^{-iO(\mathbf{k},\mathbf{k}_0)}.2i\sin(E(\mathbf{k}, \mathbf{k}_0))} \tag{3.16}$$

3.4 The Effects of Changing the Tilt Origin

All the aberration coefficients depend on the tilt origin chosen. It is helpful to examine how the coefficients are affected when the origin is changed (Typke, D. and Dierksen, K., 1992). Reverting to complex notation and defining the tilt as $\tau = t_x + it_y$,

$$\begin{aligned} W(\omega+\tau) - W(\tau) = \Re\big(&A_0'\omega^* + \tfrac{1}{2}A_1'\omega^{*2} + \tfrac{1}{2}C_1'\omega^*\omega \\ &+ \tfrac{1}{3}A_2'\omega^{*3} + \tfrac{1}{3}B_2'\omega^{*2}\omega \\ &+ \tfrac{1}{4}A_3'\omega^{*4} + \tfrac{1}{4}B_3'\omega^{*3}\omega + \tfrac{1}{4}C_3'\omega^{*2}\omega^2\big) \end{aligned} \tag{3.17}$$

where $W(\tau)$ is a trivial change of phase origin and

$$\begin{aligned} A_0' &= A_0 + A_1\tau^* + C_1\tau + A_2\tau^{*2} + \tfrac{1}{3}B_2^*\tau^2 + \tfrac{2}{3}B_2\tau^*\tau \\ &\quad + A_3\tau^{*3} + \tfrac{3}{4}B_3\tau^{*2}\tau + \tfrac{1}{4}B_3^*\tau^3 + C_3\tau^*\tau^2 \\ A_1' &= A_1 + 2A_2\tau^* + \tfrac{2}{3}B_2\tau + 3A_3\tau^{*2} + \tfrac{3}{2}B_3\tau^*\tau + C_3\tau^2 \\ C_1' &= C_1 + \Re\big(\tfrac{4}{3}B_2\tau^* + \tfrac{3}{2}B_3\tau^{*2}\big) + 2C_3\tau^*\tau \\ A_2' &= A_2 + 3A_3\tau^* + \tfrac{3}{4}B_3\tau \\ B_2' &= B_2 + \tfrac{9}{4}B_3\tau^* + 3C_3\tau \\ A_3' &= A_3; B_3' = B_3; C_3' = C_3 \end{aligned} \tag{3.18}$$

It is possible to simplify these expressions by assuming that A_3 and B_3 are negligible:

$$\begin{aligned} W(\omega+\tau) - W(\tau) = \Re\big(&A_0'\omega^* + \tfrac{1}{2}A_1'\omega^{*2} + \tfrac{1}{2}C_1'\omega^*\omega \\ &+ \tfrac{1}{3}A_2'\omega^{*3} + \tfrac{1}{3}B_2'\omega^{*2}\omega + \tfrac{1}{4}C_3'\omega^{*2}\omega^2\big) \end{aligned} \tag{3.19}$$

$$\begin{aligned}
A_0' &= A_0 + A_1\tau^* + C_1\tau + A_2\tau^{*2} + \tfrac{1}{3}B_2^*\tau^2 + \tfrac{2}{3}B_2\tau^*\tau + C_3\tau^*\tau^2 \\
A_1' &= A_1 + 2A_2\tau^* + \tfrac{2}{3}B_2\tau + C_3\tau^2 \\
C_1' &= C_1 + \Re(\tfrac{4}{3}B_2\tau^*) + 2C_3\tau^*\tau \\
A_2' &= A_2 \\
B_2' &= B_2 + 3C_3\tau \\
C_3' &= C_3
\end{aligned} \tag{3.20}$$

Using these terms, it is possible to obtain simplified versions of $E(\mathbf{k}, \mathbf{k}_0)$ and $O(\mathbf{k}, \mathbf{k}_0)$. If W is re-expressed:

$$W(\omega + \tau) = W(\tau) + E_W(\omega, \tau) + O_W(\omega, \tau) \tag{3.21}$$

then

$$\begin{aligned}
E_W(\omega, \tau) &= \Re(\tfrac{1}{2}A_1'\omega^{*2} + \tfrac{1}{2}C_1'\omega^*\omega + \tfrac{1}{4}C_3'\omega^{*2}\omega^2) \\
O_W(\omega, \tau) &= \Re(A_0'\omega^* + \tfrac{1}{3}A_2'\omega^{*3} + \tfrac{1}{3}B_2'\omega^{*2}\omega)
\end{aligned} \tag{3.22}$$

$$\begin{aligned}
E_W(\omega, \tau) = \Re(&\tfrac{1}{2}A_1\omega^{*2} + A_2\tau^*\omega^{*2} + \tfrac{1}{3}B_2\tau\omega^{*2} + \tfrac{1}{2}C_3\tau^2\omega^{*2} + \tfrac{1}{2}C_1\omega^*\omega \\
&+ \Re(\tfrac{2}{3}B_2\tau^*)\omega^*\omega + C_3\tau^*\tau\omega^*\omega + \tfrac{1}{4}C_3\omega^{*2}\omega^2) \\
O_W(\omega, \tau) = \Re(&A_0'\omega^* + \tfrac{1}{3}A_2\omega^{*3} + \tfrac{1}{3}B_2\omega^{*2}\omega + C_3\tau\omega^{*2}\omega)
\end{aligned} \tag{3.23}$$

so that

$$\begin{aligned}
E_W(\omega, \tau) = \Re(&\tfrac{1}{2}A_1\omega^{*2} + A_2\tau^*\omega^{*2} + \tfrac{1}{2}C_1\omega^*\omega + \tfrac{1}{4}C_3\omega^{*2}\omega^2 \\
&+ \tfrac{1}{3}B_2\tau\omega^{*2} + \Re(\tfrac{2}{3}B_2\tau^*)\omega^*\omega + C_3\tau\omega^*.\Re(\tau\omega^*) + \tfrac{1}{2}C_3|\tau|^2|\omega|^2) \\
O_W(\omega, \tau) = \Re(&A_0'\omega^* + \tfrac{1}{3}A_2\omega^{*3} + \tfrac{1}{3}B_2\omega^{*2}\omega + C_3\tau\omega^*|\omega|^2)
\end{aligned} \tag{3.24}$$

3.5 The Coma-Free Axis

In the above derivations, the origin which the tilts are measured against has been intentionally vague. It is convenient to define an origin where B_2' or

axial coma disappears. This corresponds to a primary beam position, τ_c, such that:

$$\tau_c = -\frac{B_2}{3C_3} \tag{3.25}$$

where A_3 and B_3 have again been ignored. This choice of tilt origin defines the *coma-free axis* of the lens because placing the primary beam there provides coma-free imaging (Zemlin, F. et al., 1978),(Typke, D. and Dierksen, K., 1992). The new effective aberrations when the beam is tilted by μ from this axis are:

$$\begin{aligned}
A_0'' &= A_0' + A_1'\mu^* + C_1'\mu + A_2'\mu^{*2} + C_3'\mu^*\mu^2 \\
A_1'' &= A_1' + 2A_2'\mu^* + C_3'\mu^2 \\
C_1'' &= C_1' + 2C_3'\mu^*\mu \\
A_2'' &= A_2' \\
B_2'' &= 3C_3'\mu \\
C_3'' &= C_3'
\end{aligned} \tag{3.26}$$

If tilted imaging is assumed relative to coma-free axis ($B_2 = 0$) then we can derive $E_{W_c}(\omega,\mu)$ and $O_{W_c}(\omega,\mu)$ from equation (3.24):

$$\begin{aligned}
E_{W_c}(\omega,\mu) = \Re\big(&\tfrac{1}{2}A_1\omega^{*2} + A_2\mu^*\omega^{*2} + \tfrac{1}{2}C_1\omega^*\omega + \tfrac{1}{4}C_3\omega^{*2}\omega^2 \\
&+ C_3\mu\omega^*.\Re(\mu\omega^*) + \tfrac{1}{2}C_3|\mu|^2|\omega|^2\big) \\
O_{W_c}(\omega,\mu) = \Re\big(&A_0'\omega^* + \tfrac{1}{3}A_2\omega^{*3} + C_3\mu\omega^*|\omega|^2\big)
\end{aligned} \tag{3.27}$$

If A_2 is ignored:

$$\begin{aligned}
E_{W_c}(\omega,\mu) &= \Re\big(W(\omega) + C_3\mu\omega^*.\Re(\mu\omega^*) + \tfrac{1}{2}C_3|\mu|^2|\omega|^2\big) \\
O_{W_c}(\omega,\mu) &= \Re\big(A_0'\omega^* + C_3\mu\omega^*|\omega|^2\big)
\end{aligned} \tag{3.28}$$

Hence, $E(\mathbf{k},\mathbf{k}_0)$ and $O(\mathbf{k},\mathbf{k}_0)$ can be easily derived:

$$\begin{aligned}
E(\mathbf{k},\mathbf{k}_0) &= \gamma(\mathbf{k}) + 2\pi C_3\lambda^3[(\mathbf{k}_0.\mathbf{k})^2 + \tfrac{1}{2}|\mathbf{k}_0|^2|\mathbf{k}|^2] \\
O(\mathbf{k},\mathbf{k}_0) &= \nabla\gamma(\mathbf{k}_0).\mathbf{k} + 2\pi C_3\lambda^3\mathbf{k}_0.\mathbf{k}|\mathbf{k}|^2
\end{aligned} \tag{3.29}$$

This is the result quoted by Koster et al. (Koster, A.J. et al., 1989). The $(\mathbf{k}.\mathbf{k}_0)^2$ term in $E(\mathbf{k},\mathbf{k}_0)$ corresponds to a tilt-dependent astigmatism and the $|\mathbf{k}|^2|\mathbf{k}_0|^2$ term to an over-focus offset (Wade, R.H., 1992). The $|\mathbf{k}|^2\mathbf{k}_0.\mathbf{k}$ term in $O(\mathbf{k},\mathbf{k}_0)$ is responsible for a frequency-dependent image shift i.e. image distortion.

3.6 Estimation of Aberration Coefficients

Coarse TEM alignment consists of minimising two-fold astigmatism, setting the defocus to a predetermined value and making sure that the electron beam is aligned to the coma-free axis (or even the voltage centre in practice). If the shape of W and the position of its axis (i.e. the coefficients A_1,C_1,B_2 and C_3) could be easily determined then a general purpose auto-alignment procedure would have been developed a long time ago. However, the phase shift of a single beam given by γ, cannot be measured directly because the phase origin is arbitrary. Only the phase difference between two or more beams can be experimentally determined. This can be achieved by tilting the electron beam and observing how the image differs from that obtained when the beam is not tilted.

Looking at equation (3.18), it is noticed that the terms are linear in the coefficients so that all the coefficients can be found uniquely via a least squares fit of the observations of either A_0' or A_1' and C_1' for several different known τ.

The coma-free axis position can be obtained by solving

$$\hat{B_2} = -\tfrac{9}{4}\hat{B_3}\tau_c^* - 3\hat{C_3}\tau_c \tag{3.30}$$

for τ_c, where $\hat{B_2}$, $\hat{B_3}$ and $\hat{C_3}$ are the values for these coefficients returned by the least squares fit.

From equation (3.16) and equation (3.22), the simplified (A_3, B_3 ignored) weak phase object transfer function can be written:

$$P(\omega,\tau) = e^{-\frac{2\pi}{\lambda}i(A_0'\omega^* + \frac{1}{3}A_2'\omega^{*3} + \frac{1}{3}B_2'\omega^{*2}\omega)}.2i\sin\tfrac{2\pi}{\lambda}(\tfrac{1}{2}A_1'\omega^{*2} + \tfrac{1}{2}C_1'\omega^*\omega + \tfrac{1}{4}C_3'\omega^{*2}\omega^2) \tag{3.31}$$

This implies that the value of A_0' can be seen in image displacements and those of A_1' and C_1' can be seen in image power spectra or diffractograms.

3.6.1 Displacement estimation

There are two main methods available to estimate image displacement when a tilt is injected. These are the cross-correlation function (XCF) and the determination of cross-power spectrum phase angles.

3.6.1.1 The cross-correlation function

Electron microscopy applications of the cross-correlation functions are covered by Frank (Frank, J., 1980). In general, the cross-correlation (XCF) between two images at arbitrary tilts, $\mathbf{k_1}$ and $\mathbf{k_2}$ is given by

$$c(\mathbf{x}) = i(\mathbf{x}, \mathbf{k_1}) * i(\mathbf{x}, \mathbf{k_2}) = \iint i(\mathbf{x}' - \mathbf{x}, \mathbf{k}_1) i(\mathbf{x}', \mathbf{k}_2)\, d^2\mathbf{x}' \tag{3.32}$$

where the $*$ denotes cross-correlation. Defining $C(\mathbf{k})$ as the Fourier Transform of $c(\mathbf{x})$,

$$C(\mathbf{k}) = I^*(\mathbf{k}, \mathbf{k_1}) I(\mathbf{k}, \mathbf{k_2}) = I_1^*(\mathbf{k}) I_2(\mathbf{k}) \tag{3.33}$$

From equation (3.15):

$$I_i(\mathbf{k}) = \Psi(\mathbf{k}).P_i(\mathbf{k}) + N_i(\mathbf{k}) \tag{3.34}$$

$N_i(\mathbf{k})$ is the additive noise present in each image. Using equation (3.29), this can be re-written as

$$I_i(\mathbf{k}) = F(\mathbf{k}) P_i(\mathbf{k}) e^{-i\nabla\gamma(\mathbf{k_i}).\mathbf{k}} + N_i(\mathbf{k}) \tag{3.35}$$

where $F(\mathbf{k})$ is the specimen object function and $P_i(\mathbf{k})$ is the transfer function for the imaging conditions *centred on the origin*, which makes the exponential term explicit. The cross-power spectral density (XPSD), which is the Fourier transform of the XCF, is thus given by

$$C(\mathbf{k}) = |F(\mathbf{k})|^2 P_1^*(\mathbf{k}) P_2(\mathbf{k}).e^{-i\mathbf{k}.\mathbf{D}} \tag{3.36}$$

$$C(\mathbf{k}) = A_F(\mathbf{k}).C_P(\mathbf{k}).e^{-i\mathbf{k}.\mathbf{D}} \tag{3.37}$$

where

$$A_F(\mathbf{k}) = |F(\mathbf{k})|^2 \tag{3.38}$$

$$C_P(\mathbf{k}) = P_1^*(\mathbf{k}) P_2(\mathbf{k}) \tag{3.39}$$

$$\mathbf{D} = \nabla\gamma(\mathbf{k_2}) - \nabla\gamma(\mathbf{k_1}) \tag{3.40}$$

It is assumed that $N_1(\mathbf{k})$ and $N_2(\mathbf{k})$ are uncorrelated with $P_i(\mathbf{k})$, $F(\mathbf{k})$ and with each other. $A_F(\mathbf{k})$ is the power spectrum of the specimen object function and will usually be a good approximation to a delta function. $c_p(\mathbf{x})$ is the XCF of the two point spread functions (PSFs) so that the shape of $c(\mathbf{x})$ is essentially that of $c_p(\mathbf{x})$ smoothed by the finite width of the object auto-correlation function, $a_f(\mathbf{x})$. Ideally we would like $c(\mathbf{x})$ to be a delta function displaced by $\frac{\mathbf{D}}{2\pi}$.

3.6.1.2 Generalised correlation functions

Knapp and Carter (Knapp, C.H. and Carter, G.C., 1976),(Carter, G.C., 1987) describe general correlation methods for time delay estimation. It is possible to regard the two cross-correlated images as two-dimensional extensions of one-dimensional signals emanating from a remote source being monitored in the presence of noise by two spatially separated sensors and explore how image pre-filters can be used to improve the estimate of the time delay between the two signals. Using the notation in (Knapp, C.H. and Carter, G.C., 1976), the two images can be re-expressed:

$$i_1(\mathbf{x}) = p_1(\mathbf{x}) * f(\mathbf{x}) + n_1(\mathbf{x}) \tag{3.41}$$

$$i_2(\mathbf{x}) = p_2(\mathbf{x} - \mathbf{D}) * f(\mathbf{x}) + n_2(\mathbf{x}) \tag{3.42}$$

where $f(\mathbf{x})$, $p_1(\mathbf{x})$, $p_2(\mathbf{x})$, $n_1(\mathbf{x})$ and $n_2(\mathbf{x})$ are real, jointly stationary random processes and $f(\mathbf{x})$, $p_1(\mathbf{x})$ and $p_2(\mathbf{x})$ are assumed to be uncorrelated with $n_1(\mathbf{x})$ and $n_2(\mathbf{x})$. The form of these two equations differs from that in (Koster, A.J. et al., 1987) because the effect of different PSFs has been taken into account. The cross-power spectral density is defined by

$$\hat{G}_{i_1 i_2}(\mathbf{k}) = I_1^*(\mathbf{k}) I_2(\mathbf{k}) \tag{3.43}$$

so that

$$\hat{G}_{i_1 i_2}(\mathbf{k}) = \hat{G}_{ff}(\mathbf{k}) \hat{G}_{p_1 p_2}(\mathbf{k}) e^{-i\mathbf{k}.\mathbf{D}} \tag{3.44}$$

$G_{n_1 n_2}(\mathbf{k})$ is zero if $n_1(\mathbf{x})$ and $n_2(\mathbf{x})$ are uncorrelated. Equation (3.44) is identical to equation (3.37) : $\hat{G}_{ff}(\mathbf{k}) = A_F(\mathbf{k})$ and $\hat{G}_{p_1 p_2}(\mathbf{k}) = C_P(\mathbf{k})$. It is important to note that $\hat{G}_{i_1 i_2}$ is an estimator of the real XPSD, $G_{i_1 i_2}$ which is based on an ensemble average (Papoulis, A., 1989).

If two filters, $H_1(\mathbf{k})$ and $H_2(\mathbf{k})$ are introduced which filter $i_1(\mathbf{x})$ and $i_2(\mathbf{x})$ respectively, then the generalised cross power spectrum is given by

$$G_{j_1 j_2}(\mathbf{k}) = H_1^*(\mathbf{k}) H_2(\mathbf{k}) \hat{G}_{i_1 i_2}(\mathbf{k}) = \psi_g(\mathbf{k}) \hat{G}_{i_1 i_2}(\mathbf{k}) \tag{3.45}$$

$\psi_g(\mathbf{k})$ should be chosen to ensure a large sharp peak in $g_{j_1 j_2}(\mathbf{x})$, rather than a broad one to ensure resolution. However sharp peaks are more sensitive to errors due to noise so that the choice of $\psi_g(\mathbf{k})$ is a compromise between good resolution and stability.

3.6.1.2.1 The Roth processor

The weighting proposed by Roth is:

$$\psi_R(\mathbf{k}) = \frac{1}{G_{i_1 i_1}(\mathbf{k})} \tag{3.46}$$

so that

$$G_{j_1 j_2}(\mathbf{k}) = \frac{\hat{G}_{ff} \hat{G}_{p_1 p_2}}{G_{ff}(\mathbf{k}) G_{p_1 p_1}(\mathbf{k}) + G_{n_1 n_1}(\mathbf{k})} . e^{-i\mathbf{k}.\mathbf{D}} \tag{3.47}$$

$$= \frac{|\hat{F}|^2 \hat{P}_1^* \hat{P}_2}{(|F|^2 |P_1|^2 + |N_1|^2)} . e^{-i\mathbf{k}.\mathbf{D}} \tag{3.48}$$

The Roth Processor thus has the desirable effect of suppressing those frequencies where $G_{n_1 n_1}(\mathbf{k})$ is large and $\hat{G}_{i_1 i_2}(\mathbf{k})$ is more likely to be in error.

3.6.1.2.2 The smoothed coherence transform (SCOT)

The errors in $\hat{G}_{i_1 i_2}(\mathbf{k})$ may be due to frequencies where $G_{n_1 n_1}(\mathbf{k})$ is large or $G_{n_2 n_2}(\mathbf{k})$ is large. It is therefore uncertain whether to use $G_{i_1 i_1}(\mathbf{k})$ or $G_{i_2 i_2}(\mathbf{k})$ in the denominator of $\psi_R(\mathbf{k})$. The Smoothed Coherence Transform (SCOT) therefore selects:

$$\psi_S(\mathbf{k}) = \frac{1}{\sqrt{G_{i_1 i_1}(\mathbf{k}) G_{i_2 i_2}(\mathbf{k})}} \tag{3.49}$$

$$G_{j_1 j_2}(\mathbf{k}) = \frac{\hat{G}_{ff}(\mathbf{k})\hat{G}_{p_1 p_2}(\mathbf{k})}{\sqrt{[G_{ff}(\mathbf{k})G_{p_1 p_1}(\mathbf{k}) + G_{n_1 n_1}(\mathbf{k})][G_{ff}(\mathbf{k})G_{p_2 p_2}(\mathbf{k}) + G_{n_2 n_2}(\mathbf{k})]}} e^{-i\mathbf{k}.\mathbf{D}} \tag{3.50}$$

$$= \frac{|\hat{F}|^2 \hat{P}_1^* \hat{P}_2}{\sqrt{(|F|^2|P_1|^2 + |N_1|^2)(|F|^2|P_2|^2 + |N_2|^2)}} . e^{-i\mathbf{k}.\mathbf{D}} \tag{3.51}$$

3.6.1.2.3 The phase transform (PHAT)

Both the Roth processor and the SCOT spread out the peak of the correlation function relative to the ideal delta-function even though it is narrowed relative to an un-filtered XCF. The Phase Transform (PHAT) avoids this:

$$\psi_P(\mathbf{k}) = \frac{1}{|G_{i_1 i_2}(\mathbf{k})|} \tag{3.52}$$

$$G_{j_1 j_2}(\mathbf{k}) = \frac{\hat{G}_{ff}(\mathbf{k})\hat{G}_{p_1 p 2}(\mathbf{k})}{|G_{ff}(\mathbf{k})G_{p_1 p_2}(\mathbf{k})|} . e^{-i\mathbf{k}.\mathbf{D}} \tag{3.53}$$

$$= \frac{P_1^* P_2}{(|P_1^* P_2| + \epsilon)} . e^{-i\mathbf{k}.\mathbf{D}} \tag{3.54}$$

If $P_1 = P_2 = 1$ The result will ideally be a delta function displaced by $\frac{\mathbf{D}}{2\pi}$. If the denominator is small or zero, the phase is likely to be in error or undefined so that some additional S/N ratio weighting is required to make $\psi_P(\mathbf{k})$ feasible (ϵ in equation (3.54)).

3.6.1.3 Taking the transfer functions into account

As the displacement in (3.37) manifests itself in the exponential term, $c_p(\mathbf{x})$ needs to have its highest value at the origin. This is not always the case and using the position of the highest value of $c(\mathbf{x})$ is an unreliable estimate of $\frac{\mathbf{D}}{2\pi}$. This problem has been investigated by Saxton (Saxton, W.O., 1994a).

$c_p(\mathbf{x})$ can be written:

$$c_p(\mathbf{x}) = \iint P_1^*(\mathbf{k}) P_2(\mathbf{k}) e^{2\pi i \mathbf{k}.\mathbf{x}} \, d^2\mathbf{k} \tag{3.55}$$

Its value at the origin is thus given by

$$c_p(\mathbf{0}) = \iint P_1^*(\mathbf{k}) P_2(\mathbf{k}) \, d^2\mathbf{k} \tag{3.56}$$

The phase differences between the two transfer functions mean that the regions where the integrand is positive may be offset by regions where it is negative, resulting in a lower XCF value at the origin than outside it.

For the XCF value at the origin to have the largest value,

$$c_p(\mathbf{0}) \geq |c_p(\mathbf{x})| \qquad \forall \mathbf{x} \neq \mathbf{0} \tag{3.57}$$

Let

$$P_i(\mathbf{k}) = |P_i(\mathbf{k})| e^{i\theta_i(\mathbf{k})} \tag{3.58}$$

so that $\forall \mathbf{x} > \mathbf{0}$,

$$\iint |P_1(\mathbf{k})||P_2(\mathbf{k})| e^{i(\theta_2(\mathbf{k}) - \theta_1(\mathbf{k}))} \, d^2\mathbf{k} \geq \left| \iint P_1(\mathbf{k})||P_2(\mathbf{k})| e^{i(\theta_2(\mathbf{k}) - \theta_1(\mathbf{k}))} e^{2\pi i \mathbf{k}.\mathbf{x}} \, d^2\mathbf{k} \right| \tag{3.59}$$

If $\theta_2(\mathbf{k}) - \theta_1(\mathbf{k})$ is zero,

$$\iint |P_1(\mathbf{k})||P_2(\mathbf{k})| \, d^2\mathbf{k} \geq \left| \iint |P_1(\mathbf{k})||P_2(\mathbf{k})| e^{2\pi i \mathbf{k}.\mathbf{x}} \, d^2\mathbf{k} \right| \qquad \forall \mathbf{x} \neq \mathbf{0} \tag{3.60}$$

which is always true. Therefore, for the inequality in equation (3.57) to hold, $P_1(\mathbf{k})$ and $P_2(\mathbf{k})$ need to have the same phase everywhere.

Adjustments to the phase of $C_P(\mathbf{k})$ which compensate for the phase differences between $P_1(\mathbf{k})$ and $P_2(\mathbf{k})$ will thus improve the central XCF peak value relative to the value elsewhere. An additional benefit is the reduction of the peak width which leads to increased positional accuracy as well as detectability. This is because the integrated square of the XCF peak is fixed by Parseval's theorem:

$$\iint |c_p(\mathbf{x})|^2 \, d^2\mathbf{x} = \iint |C_P(\mathbf{k})|^2 \, d^2\mathbf{k} \tag{3.61}$$

so that a compensated XCF with a higher central peak value would reduce the extent of the peak region.

It is possible for the peak shape in the XCF to be centro-symmetric even if $\theta_1(\mathbf{k}) \neq \theta_2(\mathbf{k})$ in equation (3.59). This is the case for axial illumination where the two transfer functions, although different, are centro-symmetric The Fourier transform of a centro-symmetric function is real and symmetric so that it has a phase of 0 or π. Doubling the phase would give a function with zero phase throughout which is sufficient to provide a peak at the origin. This technique does not require knowledge of imaging conditions.

All the processors described previously do not acknowledge the existence of the different transfer functions in the two images and consequently they are incapable of improving on the XCF displacement estimate. Phase compensation filters can be used if *a priori* information is available regarding $P_1(\mathbf{k})$ and $P_2(\mathbf{k})$.

3.6.1.3.1 The phase-compensated XCF (PC-XCF)

This is defined by

$$\psi_{PC}(\mathbf{k}) = \frac{G^*_{p_1p_2}}{(\sqrt{G_{p_1p_1}(\mathbf{k})G_{p_2p_2}(\mathbf{k})} + \epsilon)} \tag{3.62}$$

so that

$$G_{j_1j_2}(\mathbf{k}) = \frac{G_{ff}(\mathbf{k})|G_{p_1p_2}(\mathbf{k})|^2}{(\sqrt{G_{p_1p_1}(\mathbf{k})G_{p_2p_2}(\mathbf{k})} + \epsilon)} . e^{-i\mathbf{k}.\mathbf{D}} \tag{3.63}$$

$$= \frac{|F|^2|P_1|^2|P_2|^2}{(|P_1||P_2| + \epsilon)} . e^{-i\mathbf{k}.\mathbf{D}} \tag{3.64}$$

The PC-XCF effectively compensates the phase differences between $P_1(\mathbf{k})$ and $P_2(\mathbf{k})$ and consequently $c_p(\mathbf{x})$ has its highest value at the origin. The ϵ is a S/N ratio weighting to avoid singularities. There is a variant of the PC-XCF which does not need ϵ:

$$\psi_{PCN}(\mathbf{k}) = \frac{G^*_{p_1p_2}}{\sqrt{G_{i_1i_1}(\mathbf{k})G_{i_2i_2}(\mathbf{k})}} \tag{3.65}$$

so that

$$G_{j_1j_2}(\mathbf{k}) = \frac{|G_{p_1p_2}(\mathbf{k})|^2}{\sqrt{(G_{p_1p_1}(\mathbf{k}) + G_{n_1n_1(\mathbf{k})})(G_{p_2p_2}(\mathbf{k}) + G_{n_2n_2})}}.e^{-i\mathbf{k}.\mathbf{D}} \tag{3.66}$$

$$= \frac{|P_1|^2|P_2|^2}{\sqrt{(|P_1|^2 + |N_1|^2)(|P_2|^2 + |N_2|^2)}}.e^{-i\mathbf{k}.\mathbf{D}} \tag{3.67}$$

3.6.1.3.2 The squared-transfer XCF (ST-XCF)

This is defined by

$$\psi_{ST}(\mathbf{k}) = G^*_{p_1p_2} \tag{3.68}$$

so that

$$G_{j_1j_2}(\mathbf{k}) = G_{ff}(\mathbf{k})|G_{p_1p_2}(\mathbf{k})|^2.e^{-i\mathbf{k}.\mathbf{D}} \tag{3.69}$$

$$= |F|^2|P_1|^2|P_2|^2.e^{-i\mathbf{k}.\mathbf{D}} \tag{3.70}$$

The ST-XCF is more pessimistic than the PC-XCF since it smoothes the XCF peak whilst reducing noise levels. It may be considered as a phase compensating XCF with different modulus weighting and has a variant for axial illumination which does not require knowledge of imaging conditions : the original XCF is simply squared.

3.6.1.3.3 The transfer-divided XCF (TD-XCF)

This is defined by

$$\psi_{TD}(\mathbf{k}) = \frac{G^*_{p_1p_2}}{(G_{p_1p_1}(\mathbf{k})G_{p_2p_2}(\mathbf{k}) + \epsilon)} \tag{3.71}$$

so that

$$G_{j_1 j_2}(\mathbf{k}) = \frac{G_{ff}(\mathbf{k})|G_{p_1p_2}(\mathbf{k})|^2}{(G_{p_1p_1}(\mathbf{k})G_{p_2p_2}(\mathbf{k}) + \epsilon)}.e^{-i\mathbf{k}.\mathbf{D}} \tag{3.72}$$

$$= \frac{|F|^2|P_1|^2|P_2|^2}{(|P_1|^2|P_2|^2 + \epsilon)}.e^{-i\mathbf{k}.\mathbf{D}} \tag{3.73}$$

Like the ST-XCF, it may be considered as a phase compensating XCF with different modulus weighting. It is more optimistic than the PC-XCF because it improves the XCF peak whilst increasing noise levels. The ϵ is a S/N ratio weighting to avoid singularities. There is a variant of the TD-XCF which does not need ϵ:

$$\psi_{TDN}(\mathbf{k}) = \frac{G^*_{p_1p_2}}{G_{i_1i_1}(\mathbf{k})G_{i_2i_2}(\mathbf{k})} \tag{3.74}$$

$$G_{j_1 j_2}(\mathbf{k}) = \frac{|G_{p_1p_2}(\mathbf{k})|^2}{(G_{p_1p_1}(\mathbf{k}) + G_{n_1n_1(\mathbf{k})})(G_{p_2p_2}(\mathbf{k}) + G_{n_2n_2})}.e^{-i\mathbf{k}.\mathbf{D}} \tag{3.75}$$

$$= \frac{|P_1|^2|P_2|^2}{|F|^2.(|P_1|^2 + |N_1|^2)(|P_2|^2 + |N_2|^2)}.e^{-i\mathbf{k}.\mathbf{D}} \tag{3.76}$$

3.6.1.4 Cross-power spectrum phase angles

Rather than using the position of the peak in the XCF, it should be possible to determine image displacement by fitting a linear ramp to the phase in the cross-power spectrum. Consider two images: one without tilt and the other with an arbitrary injected tilt, τ. Using an obvious notation, the XPSD will contain the term:

$$P(0)^*P(\tau) = P_0^*P_+ = e^{-i(O_+ - O_0)} \sin E_+ \sin E_0 \tag{3.77}$$

This is introduced by Koster et al. (Koster, A.J. et al., 1988; Koster, A.J. et al., 1989) although their analysis omits A_2. From equation (3.22), the exponent term is:

$$\frac{2\pi}{\lambda}\Re\big((A_{0+}-A_0)\omega^* + \tfrac{1}{3}(A'_{2+}-A_2)\omega^{*3} + \tfrac{1}{3}(B'_{2+}-B_2)\omega^{*2}\omega\big) \quad (3.78)$$

Now, from equation (3.20),

$$\begin{aligned} A'_{0+} &= A_0 + A_1\tau^* + C_1\tau + A_2\tau^{*2} + \tfrac{1}{3}B_2^*\tau^2 + \tfrac{2}{3}B_2\tau^*\tau + C_3\tau^*\tau^2 \\ A'_{2+} &= A_2 \\ B'_{2+} &= B_2 + 3C_3\tau \end{aligned} \quad (3.79)$$

so that the exponent term becomes:

$$\frac{2\pi}{\lambda}\Re\big((A_1\tau + C_1\tau + A_2\tau^{*2} + \tfrac{1}{3}B_2^*\tau^2 + \tfrac{2}{3}B_2\tau^*\tau + C_3\tau^*\tau^2)\omega^* + C_3\tau\omega^{*2}\omega\big) \quad (3.80)$$

This shows that the phase shift is not linear because of the $\omega^{*2}\omega$ term. Even if this term is ignored for small ω, the problem still remains that there will be $\pm\pi$ phase jumps due to the sine terms and also a modulo 2π ambiguity in the phase measurement itself. Furthermore, the different coherence envelope terms for the two images will restrict the size of the linear region in the XPSD. The first two of these problems can be compensated for using phase doubling. However, coherence envelopes are difficult to compensate without *a priori* knowledge of the imaging conditions.

3.6.2 Analysing displacement observations

3.6.2.1 The three-tilt and five-tilt displacement methods

According to Koster et al. (Koster, A.J. et al., 1989), the initial misalignment can be estimated using the three tilt method. In this, three images are recorded : one without injected tilt and two with equal and opposite tilts. If these tilts are 0, τ and $-\tau$ then the three images can be cross-correlated in such a way that the phase ramp is completely linear. The XCF of the XCF of each tilted image [1] with the axial one is given by:

[1] For expediency, the image captured when the electron beam is tilted will be referred to as a tilted image and the image captured without injected beam tilt will be referred to as an axial image

$$\big(I^*(-\tau)I(0)\big)^*.I^*(0)I(\tau) = I_-I_0^{*2}I_+ \tag{3.81}$$

in which case the transfer function term can be derived from equation (3.16):

$$P_-P_0^{*2}P_+ = e^{-i(O_-+O_+-2O_0)}\sin E_-\sin E_+\sin^2 E_0 \tag{3.82}$$

From equation (3.22), the exponent term is:

$$\begin{aligned}\tfrac{2\pi}{\lambda}\Re\big((A'_{0-}+A'_{0+}-2A_0)\omega^* + \tfrac{1}{3}(A'_{2-}+A'_{2+}-2A_2)\omega^{*3} \\ + \tfrac{1}{3}(B'_{2-}+B'_{2+}-2B_2)\omega^{*2}\omega\big)\end{aligned} \tag{3.83}$$

Now, from equation (3.20),

$$\begin{aligned}A'_{0-} &= A_0 - A_1\tau^* - C_1\tau + A_2\tau^{*2} + \tfrac{1}{3}B_2^*\tau^2 + \tfrac{2}{3}B_2\tau^*\tau - C_3\tau^*\tau^2\\ A'_{0+} &= A_0 + A_1\tau^* + C_1\tau + A_2\tau^{*2} + \tfrac{1}{3}B_2^*\tau^2 + \tfrac{2}{3}B_2\tau^*\tau + C_3\tau^*\tau^2\\ A'_{2-} &= A'_{2+} = A_2\\ B'_{2-} &= B_2 - 3C_3\tau\\ B'_{2+} &= B_2 + 3C_3\tau\end{aligned} \tag{3.84}$$

so that the exponent term simplifies to give:

$$\tfrac{2\pi}{\lambda}\Re\big(2A_2\tau^{*2} + \tfrac{2}{3}B_2^*\tau^2 + \tfrac{4}{3}B_2\tau^*\tau\big)\omega^* \tag{3.85}$$

This is different from the term in equation (3.80) because the non-linear $\omega^{*2}\omega$ term has disappeared. Applying phase doubling to accommodate the π shifts of the sin factors and examining the XCF should thus lead to an unambiguous value for the displacement. This, in turn, should provide an estimate for A_2 and B_2 in combination if applied to more than one value for τ.

Koster et al. (Koster, A.J. et al., 1989) assume that A_2 is negligible so that only three images are needed to estimate the position of the coma-free

axis. To simplify the maths, the tilts can be assumed to be either in the x- or y-direction. Considering x-direction tilts only, where $\tau = \pm t_1$, the displacement, $\delta_x = d_{xx} + id_{xy}$ is simply:

$$\delta_x = \tfrac{2}{3}B_2^* t_1^2 + \tfrac{4}{3}B_2 t_1^2 \tag{3.86}$$

Expanding the B_2 terms gives:

$$\delta_x = \tfrac{t_1^2}{3}\big(2(B_{2r} - iB_{2i}) + 4(B_{2r} + iB_{2i})\big) = \tfrac{t_1^2}{3}(6B_{2r} + 2iB_{2i}) \tag{3.87}$$

Substituting $\tau_c = t_{cx} + it_{cy}$ from equation (3.25):

$$\begin{aligned} d_{xx} &= -6t_1^2 C_3 t_{cx} \\ d_{xy} &= -2t_1^2 C_3 t_{cy} \end{aligned} \tag{3.88}$$

Therefore the estimate for the coma-free axis is given by

$$\begin{aligned} \hat{t_{cx}}|_{\tau=\pm t_1} &= -\frac{\hat{d_{xx}}}{6t_1^2 C_3} \\ \hat{t_{cy}}|_{\tau=\pm t_1} &= -\frac{\hat{d_{xy}}}{2t_1^2 C_3} \end{aligned} \tag{3.89}$$

As τ_c is the position of the coma-free axis relative to the arbitrary origin chosen for the axial image, the misalignment is $-\tau_c$.

If the tilts are in the y-direction ($\tau = \pm it_1$), the resultant displacement, $\delta_y = d_{yx} + id_{yy}$ is similarly derived:

$$\begin{aligned} d_{yx} &= -2t_1^2 C_3 t_{cx} \\ d_{yy} &= -6t_1^2 C_3 t_{cy} \end{aligned} \tag{3.90}$$

and

$$\begin{aligned} \hat{t_{cx}}|_{\tau=\pm it_1} &= -\frac{\hat{d_{yx}}}{2t_1^2 C_3} \\ \hat{t_{cy}}|_{\tau=\pm it_1} &= -\frac{\hat{d_{yy}}}{6t_1^2 C_3} \end{aligned} \tag{3.91}$$

If A_2 in equation (3.85) is taken into account, it is necessary to use five tilts: 0, $\pm t_1$ and $\pm it_1$. For the tilts in the x-direction:

$$\delta_x = 2A_2 t_1^2 + \tfrac{2}{3} B_2^* t_1^2 + \tfrac{4}{3} B_2 t_1^2 \tag{3.92}$$

so that:

$$\delta_x = \tfrac{t_1^2}{3}(6A_{2r} + 6iA_{2i} + 6B_{2r} + 2iB_{2i}) \tag{3.93}$$

Similarly, for tilts in the y-direction:

$$\delta_y = \tfrac{t_1^2}{3}(-6A_{2r} - 6iA_{2i} + 2B_{2r} + 6iB_{2i}) \tag{3.94}$$

Therefore:

$$\begin{aligned} d_{xx} &= 2t_1^2(A_{2r} + B_{2r}) \\ d_{xy} &= 2t_1^2(A_{2i} + \tfrac{1}{3}B_{2i}) \\ d_{yx} &= 2t_1^2(-A_{2r} + \tfrac{1}{3}B_{2r}) \\ d_{yy} &= 2t_1^2(-A_{2i} + B_{2i}) \end{aligned} \tag{3.95}$$

Solving for A_2 and B_2 gives:

$$\begin{aligned} \hat{B_{2r}} &= \frac{3\hat{d_{xx}} + 3\hat{d_{yx}}}{8t_1^2} \\ \hat{B_{2i}} &= \frac{3\hat{d_{xy}} + 3\hat{d_{yy}}}{8t_1^2} \\ \hat{A_{2r}} &= \frac{\hat{d_{xx}} - 3\hat{d_{yx}}}{8t_1^2} \\ \hat{A_{2i}} &= \frac{3\hat{d_{xy}} - \hat{d_{yy}}}{8t_1^2} \end{aligned} \tag{3.96}$$

The coma-free axis is thus:

$$\hat{t_{cx}} = -\frac{\hat{d_{xx}} + \hat{d_{yx}}}{8C_3 t_1^2}$$
$$\hat{t_{cy}} = -\frac{\hat{d_{xy}} + \hat{d_{yy}}}{8C_3 t_1^2} \tag{3.97}$$

The three-fold astigmatism terms in equation (3.96) are independent of the injected tilt. If A_2 is mistakenly ignored, estimates of the position of the coma-free axis, $\tau_c|_{\tau=\pm t_1}$ and $\tau_c|_{\tau=\pm it_1}$ will in fact find an axis a distance $\epsilon = e_x + ie_y$ from the coma-free. e_x and e_y will be different depending on whether the three tilts are in the x- or y-direction. For tilts in the x-direction:

$$e_x|_{\tau=\pm t_1} = \hat{t_{cx}}|_{\tau=\pm t_1} - \hat{t_{cx}} = -\frac{A_{2r}}{3C_3}$$
$$e_y|_{\tau=\pm t_1} = \hat{t_{cy}}|_{\tau=\pm t_1} - \hat{t_{cy}} = -\frac{A_{2i}}{C_3} \tag{3.98}$$

Similarly for tilts in the y-direction:

$$e_x|_{\tau=\pm it_1} = \hat{t_{cx}}|_{\tau=\pm it_1} - \hat{t_{cx}} = \frac{A_{2r}}{C_3}$$
$$e_y|_{\tau=\pm it_1} = \hat{t_{cy}}|_{\tau=\pm it_1} - \hat{t_{cy}} = \frac{A_{2i}}{3C_3} \tag{3.99}$$

If $|A_2|$ is $1\mu m$ and C_3 is $1mm$ then $|\epsilon|$ is greater than 1 mrad.

The second step in the three-tilt method is to calculate A_1 and C_1 by injecting two tilts at right angles to each other (Koster, A.J. et al., 1989) and measuring the displacements. This step assumes that the misalignment has been correctly estimated in the first step. Clearly, even for relatively small values of A_2, residual misalignment would be present if A_2 is ignored as in the three-tilt method.

Even if the 'five-tilt' method is used, the lack of incorporation of linear specimen drift into the analysis will lead to unreliable aberration estimates.

3.6.2.2 Least squares displacement fit

As noted in section 3.6, aberrations of interest can be also ascertained by a linear least squares procedure if sufficient observations of image displacements resulting from predefined injected tilts are available. Indeed, greater reliability and accuracy is expected if more data are exploited. Saxton (Saxton, W.O., 1994c) describes a method for fitting defocus, two-fold astigmatism, misalignment and spherical aberrations using this method. This method is unreliable because it does not include the effects of three-fold astigmatism.

Assume that N predetermined tilts are injected. The resulting displacements are obtained by differentiating $W(x,y)$ from equation (3.3):

$$\begin{aligned} S_{r_n} &= d_x(\mathbf{t_n}+\mathbf{m}) - d_x(\mathbf{m}) \\ &= A_{1r}t_{x_n} + A_{1i}t_{y_n} + C_1 t_{x_n} + n\Delta C_1 t_{x_n} \\ &+ A_{2r}(t_{x_n}^2 - t_{y_n}^2) + 2A_{2i}t_{x_n}t_{y_n} \\ &+ \tfrac{1}{3}B_{2r}(3t_{x_n}^2 + t_{y_n}^2) + \tfrac{2}{3}B_{2i}t_{x_n}t_{y_n} \\ &+ A_{3r}((t_{x_n}^2 - t_{y_n}^2)t_{x_n} - 2t_{x_n}t_{y_n}^2) + A_{3i}(3t_{x_n}^2 t_{y_n} - t_{y_n}^3) \\ &+ B_{3r}t_{x_n}^3 + \tfrac{1}{2}B_{3i}(3t_{x_n}^2 t_{y_n} + t_{y_n}^3) \\ &\qquad + C_3(t_{x_n}^2 + t_{y_n}^2)t_{x_n} + n\Delta x \end{aligned} \tag{3.100}$$

$$\begin{aligned} S_{i_n} &= d_y(\mathbf{t_n}+\mathbf{m}) - d_y(\mathbf{m}) \\ &= -A_{1r}t_{y_n} + A_{1i}t_{x_n} + C_1 t_{y_n} + n\Delta C_1 t_{y_n} \\ &- 2A_{2r}t_{x_n}t_{y_n} + A_{2i}(t_{x_n}^2 - t_{y_n}^2) \\ &+ \tfrac{2}{3}B_{2r}t_{x_n}t_{y_n} + \tfrac{1}{3}B_{2i}(t_{x_n}^2 + 3t_{y_n}^2) \\ &+ A_{3r}((t_{y_n}^2 - t_{x_n}^2)t_{y_n} - 2t_{x_n}^2 t_{y_n}) + A_{3i}(t_{x_n}^3 - 3t_{x_n}t_{y_n}^2) \\ &- B_{3r}t_{y_n}^3 + \tfrac{1}{2}B_{3i}(t_{x_n}^3 + 3t_{x_n}t_{y_n}^2) \\ &\qquad + C_3(t_{x_n}^2 + t_{y_n}^2)t_{y_n} + n\Delta y \end{aligned} \tag{3.101}$$

where

$$\mathbf{t_n} = \begin{pmatrix} t_{x_n} \\ t_{y_n} \end{pmatrix} \tag{3.102}$$

n is the tilt number (0..N-1), ΔC_1 is the focus drift per image and $(\Delta x, \Delta y)$ is the linear drift per image.

The aberrations are estimated relative to the arbitrary origin provided by the initial beam direction. The observed displacements are entirely linear functions of the aberrations so that, given a sufficient number of independent observations for different injected tilts, a least squares fit can find a unique solution for all aberration coefficients, the position of the coma-free axis and the aberrations coefficients relative to the axis from equations (3.30) and (3.18).

3.6.3 Effective defocus and astigmatism estimation

The diffractogram of a tilted image is similar to that of an axial image with additional astigmatism and defocus (3.31). This approximation applies if envelope terms are neglected which is reasonable near the centre of the diffractogram. This property can be used to ascertain aberration parameters for images in which an amorphous region is present.

In a tilted image, a section through the diffractogram at an angle of θ_i will have an effective defocus of:

$$C_{1i} = C_1 + A_{1r}(\cos^2\theta_i - \sin^2\theta_i) + 2A_{1i}\cos\theta_i \sin\theta_i \tag{3.103}$$

which is found by fitting the positions of maxima and minima. If several observations for C_{1i} are made for different values of θ_i it is possible to estimate C_1, A_{1r} and A_{1i} by a least squares method (Saxton, W.O., Private Communication).

Define the squared error as:

$$E = \sum_i [C_{1i} - C_1 + A_{1r}(\cos^2\theta_i - \sin^2\theta_i) + 2A_{1i}\cos\theta_i \sin\theta_i]^2 \tag{3.104}$$

Minimising this error involves differentiating E with respect to C_1, A_{1r} and A_{1i}. The least squared solution obtained in this way is:

$$\sum_i \begin{pmatrix} 1 & p_i & q_i \\ p_i & p_i^2 & p_i q_i \\ q_i & p_i q_i & q_i^2 \end{pmatrix} \begin{pmatrix} C_1 \\ A_{1r} \\ A_{1i} \end{pmatrix} = \sum_i \begin{pmatrix} C_{1i} \\ p_i C_{1i} \\ q_i C_{1i} \end{pmatrix} \tag{3.105}$$

where

$$\begin{aligned} p_i &= \cos^2\theta_i - \sin^2\theta_i \\ q_i &= 2\cos\theta_i \sin\theta_i \end{aligned} \tag{3.106}$$

3.6.4 Analysing effective defocus and astigmatism observations

3.6.4.1 The four-tilt diffractogram method

If four images are captured at tilts of $(\pm t_1, 0)$ and $(0, \pm t_1)$ then it is possible to estimate A_1, C_1, B_2, A_2 and C_3 by measuring the effective defocus and astigmatism of each diffractogram:

$$\begin{aligned} A_{1r_{+x}} &= A_{1r} + 2A_{2r}t_1 + \tfrac{2}{3}B_{2r}t_1 + C_3t_1^2 \\ A_{1r_{-x}} &= A_{1r} - 2A_{2r}t_1 - \tfrac{2}{3}B_{2r}t_1 + C_3t_1^2 \\ A_{1r_{+y}} &= A_{1r} + 2A_{2i}t_1 - \tfrac{2}{3}B_{2i}t_1 - C_3t_1^2 \\ A_{1r_{-y}} &= A_{1r} - 2A_{2i}t_1 + \tfrac{2}{3}B_{2i}t_1 - C_3t_1^2 \end{aligned} \tag{3.107}$$

$$\begin{aligned} A_{1i_{+x}} &= A_{1i} + 2A_{2i}t_1 + \tfrac{2}{3}B_{2i}t_1 \\ A_{1i_{-x}} &= A_{1i} - 2A_{2i}t_1 - \tfrac{2}{3}B_{2i}t_1 \\ A_{1i_{+y}} &= A_{1i} - 2A_{2r}t_1 + \tfrac{2}{3}B_{2r}t_1 \\ A_{1i_{-y}} &= A_{1i} + 2A_{2r}t_1 - \tfrac{2}{3}B_{2r}t_1 \end{aligned} \tag{3.108}$$

$$\begin{aligned} C_{1_{+x}} &= C_1 + \tfrac{4}{3}B_{2r}t_1 + 2C_3t_1^2 \\ C_{1_{-x}} &= C_1 - \tfrac{4}{3}B_{2r}t_1 + 2C_3t_1^2 \\ C_{1_{+y}} &= C_1 + \tfrac{4}{3}B_{2i}t_1 + 2C_3t_1^2 \\ C_{1_{-y}} &= C_1 - \tfrac{4}{3}B_{2i}t_1 + 2C_3t_1^2 \end{aligned} \tag{3.109}$$

Hence:

$$\begin{aligned}
\hat{A_{1r}} &= \tfrac{1}{4}(A_{1r_{+x}} + A_{1r_{-x}} + A_{1r_{+y}} + A_{1r_{-y}}) \\
\hat{A_{1i}} &= \tfrac{1}{4}(A_{1i_{+x}} + A_{1i_{-x}} + A_{1i_{+y}} + A_{1i_{-y}}) \\
\hat{C_1} &= \tfrac{1}{4}(C_{1_{+x}} + C_{1_{-x}} + C_{1_{+y}} + C_{1_{-y}} - 8\hat{C_3}t_1^2) \\
\hat{B_{2r}} &= \tfrac{3}{8t_1}(C_{1_{+x}} - C_{1_{-x}}) \\
\hat{B_{2i}} &= \tfrac{3}{8t_1}(C_{1_{+y}} - C_{1_{-y}}) \\
\hat{C_3} &= \tfrac{1}{t_1^2}(A_{1r_{+x}} + A_{1r_{-x}} - A_{1r_{+y}} - A_{1r_{-y}}) \\
\hat{A_{2r}} &= \tfrac{1}{8t_1}(A_{1r_{+x}} - A_{1r_{-x}} - A_{1i_{+y}} + A_{1i_{-y}}) \\
\hat{A_{2i}} &= \tfrac{1}{8t_1}(A_{1i_{+x}} - A_{1i_{-x}} + A_{1r_{+y}} + A_{1r_{-y}})) \qquad (3.110)
\end{aligned}$$

In this method, C_1 and C_3 are unfortunately strongly interdependent so that it is usually necessary to determine a value for C_3 independently and fix it. The values of $\hat{A_{1r}}$, $\hat{A_{1i}}$ and $\hat{C_1}$ can be adjusted to be those relative to the coma-free axis from equations (3.30) and (3.18). Although this method is not susceptible to specimen drift, it relies on the assumption that the injected tilts are orthogonal to each other and have the same magnitude, t_1.

Krivanek and Fan (Krivanek, O.L. and Fan, G.Y., 1992) and Krivanek and Leber (Krivanek, O.L. and Leber, M.L., 1994) derive terms based on measurements of effective two-fold astigmatism only. The inclusion of effective defocus measurements in the above analysis makes it more robust.

3.6.4.2 Least squares diffractogram fit

The least squares fit described above for displacement measurements can be used to find aberrations using diffractogram measurements. Greater accuracy and reliability is expected as entire data set is taken into account.

$$\begin{aligned}
A'_{1r_n} &= A_{1r} + 2A_{2r}t_{x_n} + 2A_{2i}t_{y_n} + \tfrac{2}{3}B_{2r}t_{x_n} - \tfrac{2}{3}B_{2i}t_{y_n} \\
&\quad + 3A_{3r}(t_{x_n}^2 - t_{y_n}^2) + 6A_{3i}t_{x_n}t_{y_n} \\
&\quad + \tfrac{3}{2}B_{3r}(t_{x_n}^2 + t_{y_n}^2) + C_3(t_{x_n}^2 - t_{y_n}^2) \qquad (3.111)
\end{aligned}$$

$$\begin{aligned} A'_{1i_n} &= A_{1i} + 2A_{2i}t_{x_n} - 2A_{2r}t_{y_n} + \tfrac{2}{3}B_{2i}t_{x_n} + \tfrac{2}{3}B_{2r}t_{y_n} \\ &\quad + 3A_{3i}(t_{x_n}^2 - t_{y_n}^2) - 6A_{3r}t_{x_n}t_{y_n} \\ &\quad + \tfrac{3}{2}B_{3i}(t_{x_n}^2 + t_{y_n}^2) + 2C_3t_{x_n}t_{y_n} \end{aligned} \tag{3.112}$$

$$\begin{aligned} C'_{1_n} &= C_1 + n\Delta C_1 + \tfrac{4}{3}B_{2r}t_{x_n} + \tfrac{4}{3}B_{2i}t_{y_n} \\ &\quad + \tfrac{3}{2}B_{3r}(t_{x_n}^2 - t_{y_n}^2) + 3B_{3i}t_{x_n}t_{y_n} + 2C_3(t_{x_n}^2 + t_{y_n}^2) \end{aligned} \tag{3.113}$$

As for the displacement least squares fit, the coma-free axis position and the aberration coefficients relative to it can be found from equations (3.30) and (3.18).

3.6.5 Tilt coil calibration

The calibration of x- and y-tilt coils usually involves the assumption that they are orthogonal and have the same strength. It is possible to test the assumption by including a tilt coil calibration (TCC) matrix, $\mathbf{T_{cc}}$, which models the deviation of the tilt coil current directions and strengths from those expected (Saxton, W.O. et al., 1994). If the expected injected tilt is $\mathbf{t_n}$ and the actual injected tilt is $\mathbf{t'_n}$ then:

$$\mathbf{t'_n} = \mathbf{T_{cc}}\mathbf{t_n} \tag{3.114}$$

$$\mathbf{T_{cc}} = \begin{pmatrix} t_{xx} & t_{xy} \\ t_{yx} & t_{yy} \end{pmatrix} \tag{3.115}$$

Figure 3.12 indicates how the elements of the TCC matrix can be interpreted in the general case (Saxton, W.O., Private Communication):

$$\begin{pmatrix} t'_1 \\ t'_2 \end{pmatrix} = \begin{pmatrix} a\cos\theta_1 & b\cos\theta_2 \\ a\sin\theta_1 & b\sin\theta_2 \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \tag{3.116}$$

Therefore

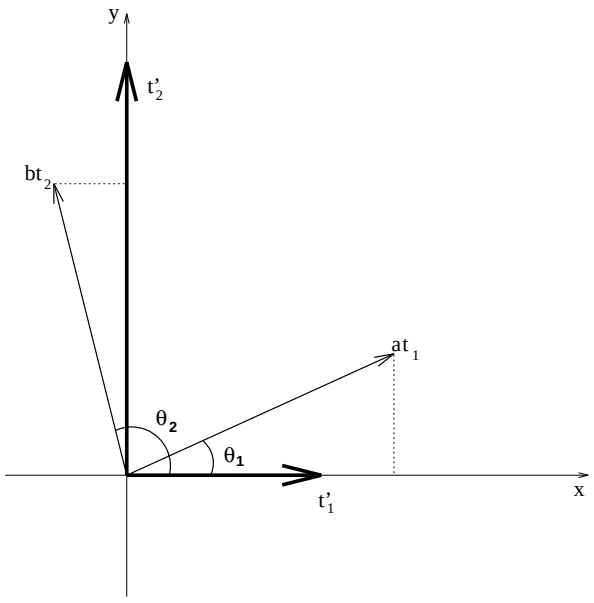


Figure 3.12: Modeling the tilt coil calibration (TCC) matrix

$$\begin{aligned} \theta_1 &= \tan^{-1} \frac{t_{yx}}{t_{xx}} \\ \theta_2 &= \tan^{-1} \frac{t_{yy}}{t_{xy}} \\ a &= \sqrt{{t_{xx}}^2 + {t_{yx}}^2} \\ b &= \sqrt{{t_{xy}}^2 + {t_{yy}}^2} \end{aligned} \tag{3.117}$$

The above four-element form of the TCC matrix will lead to a non-linear fit where all four elements of $\mathbf{T_{cc}}$ are fitted. However, a unique solution cannot be obtained if all the aberrations and TCC matrix elements are fitted: multiplying the injected tilt in equation (3.20) by a real factor f gives:

$$
\begin{aligned}
A_0' &= A_0 + fA_1\tau^* + fC_1\tau + f^2A_2\tau^{*2} + \tfrac{1}{3}f^2B_2^*\tau^2 + \tfrac{2}{3}f^2B_2\tau^*\tau + f^3C_3\tau^*\tau^2 \\
A_1' &= A_1 + 2fA_2\tau^* + \tfrac{2}{3}fB_2\tau + C_3f^2\tau^2 \\
C_1' &= C_1 + \Re\left(\tfrac{4}{3}fB_2\tau^*\right) + f^2C_3\tau^*\tau \\
A_2' &= A_2 \\
B_2' &= B_2 + 3fC_3\tau \\
C_3' &= C_3
\end{aligned} \tag{3.118}
$$

For displacement fits, the A_0' term can completely accommodate f by multiplying A_1 and C_1 by $\frac{1}{f}$, A_2 and B_2 by $\frac{1}{f^2}$ and C_3 by $\frac{1}{f^3}$. For diffractogram fits, the A_1' and C_1' terms can completely accommodate f by multiplying A_2 and B_2 by $\frac{1}{f}$ and C_3 by $\frac{1}{f^2}$.

Therefore, a non-linear fit of all four elements of the TCC matrix and a linear fit of all the aberrations in A_0' or A_1' and C_1' will lead to a continuum of solutions. It is worth noting, however, that fitting a combination of displacement and diffractogram observations and all four TCC matrix elements will provide a unique solution because the adjustments required to compensate for f in A_0' are different from those required in A_1' and C_1'. Rather than doing a complicated combination fit, it is possible to obtain a unique solution by fixing at least one coefficient. This can be an aberration (for example C_3) or elements in the TCC matrix. Possible variants of the general TCC matrix are:

$$
\mathbf{T}_{\mathbf{cc_3}} = \begin{pmatrix} a\cos\theta & -b\sin\theta \\ a\sin\theta & b\cos\theta \end{pmatrix} \tag{3.119}
$$

$$
\mathbf{T}_{\mathbf{cc_{2a}}} = \begin{pmatrix} r\cos\theta & -r\sin\theta \\ r\sin\theta & r\cos\theta \end{pmatrix} \tag{3.120}
$$

$$
\mathbf{T}_{\mathbf{cc_{2b}}} = \begin{pmatrix} \cos\theta_1 & \cos\theta_2 \\ \sin\theta_1 & \sin\theta_2 \end{pmatrix} \tag{3.121}
$$

$$
\mathbf{T}_{\mathbf{cc_{1a}}} = \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \tag{3.122}
$$

$$
\mathbf{T}_{\mathbf{cc_{1b}}} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{3.123}
$$

Only $\mathbf{T_{cc_{2b}}}$ and $\mathbf{T_{cc_{1b}}}$ will lead to a unique solutions if all the aberrations are fitted.

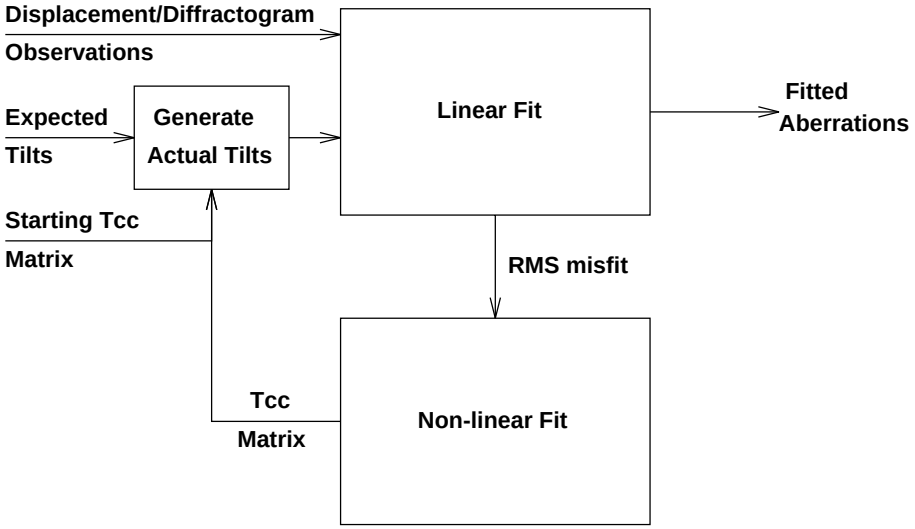


Figure 3.13: Hybrid linear and non-linear least squares fit

The actual fitting procedure is a hybrid linear and non-linear least squares one essentially due to Saxton (Saxton, W.O., Private Communication) and is illustrated in figure 3.13. The linear fit module estimates the aberrations using linear least squares for a given TCC matrix and set of observations and tilts. It returns an r.m.s. misfit between predicted and observed values which is passed on to a non-linear fit module which seeks to minimise it with respect to the elements of the TCC matrix. The method used for the non-linear fit is Powell's direction set method (Press, W.H. et al., 1992).

A physical understanding of the TCC matrix is currently not possible although it is known to vary with illumination conditions (C2 aperture size and excitation) due to flux leakage from the condenser area to the objective area.

3.6.6 X-Y anisotropy

It has been assumed that the framestore pixel aspect ratio is 1:1 i.e. the x-sampling interval is the same as the y-sampling interval. If these differ then

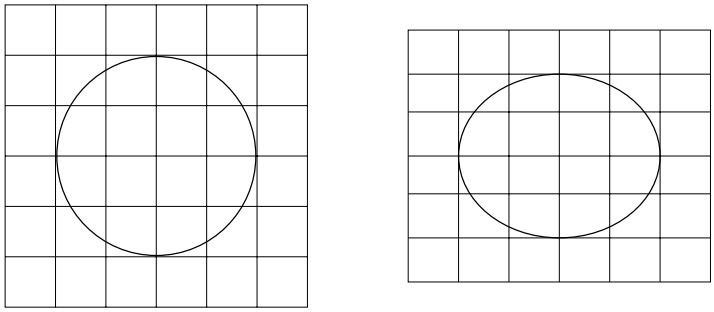


Figure 3.14: Effect of different sampling intervals

the x-y anisotropy needs to be accommodated in the fit. Let the y-sampling interval be greater than the x-sampling one by a factor f ($f > 1$). As illustrated in figure 3.14, this will lead to vertical shifts being underestimated by a factor f. Accordingly, a diffractogram would appear stretched vertically by a factor f. This can be interpreted as a change in A_1' and C_1' in equation (3.20) if it is assumed that the tilts and effective defocus values are large enough for the C_3 term to be negligible:

$$C_1'\omega^*\omega + A_1'\omega^{*2} = C_{1_f}'\omega_f^*\omega_f + A_{1_f}'\omega_f^{*2} \tag{3.124}$$

Here, $\omega_f = x + ify$ and C_{1_f}' and A_{1_f}' are the new effective defocus and astigmatism values arising from the different sampling intervals. Converting the above equation into Cartesian form gives:

$$\begin{aligned} C_1'(x^2+y^2) + A_{1r}'(x^2-y^2) + 2A_{1i}'xy = \\ C_{1_f}'(x^2+f^2y^2) + A_{1r_f}'(x^2-f^2y^2) + 2A_{1i_f}'xfy \end{aligned} \tag{3.125}$$

which has the solution:

$$C_1' = \tfrac{1}{2}\big(C_{1_f}'(f^2+1) - A_{1r_f}'(f^2-1)\big) \tag{3.126}$$

$$A_{1r}' = \tfrac{1}{2}\big(A_{1r_f}'(f^2+1) - C_{1_f}'(f^2-1)\big) \tag{3.127}$$

$$A_{1i}' = fA_{1i_f}' \tag{3.128}$$

Therefore, given observations of C'_{1_f} and A'_{1_f}, from a diffractogram fit, it is possible to derive the actual C'_1 and A'_1 which compensate for the x-y anisotropy.

Chapter 4

Experimental Results

4.1 Introduction

After a description of TEM calibration procedures, this chapter presents extensive results from the fitting of parameters from displacement and diffractogram measurements using linear and non-linear fitting techniques. It reports the practical aspects of using phase compensating filters in displacement estimation, given limited *a priori* information about the imaging conditions. Finally, evidence is presented for a significant level of three-fold astigmatism in the TEM available to this project.

4.2 TEM calibration

Calibration of the TEM is necessary before any experiments can be performed. Chau (Chau, K.-L., 1993) details the calibration procedure for the JEOL 4000 EXII. It is possible to adjust lens currents for this microscope in a precise and reproducible way through the use of digital to analogue converters (DACs). Remote entry of the DAC input values from a computer via a serial interface thus permits accurate control of lens currents from the computer. The following calibration procedures are required:

1. The magnification from specimen level in nm to framestore level in pixels.

2. The conversion factor for tilt coil strength (DAC units to mrad) and orientation relative to framestore axes.

3. The conversion factor for stigmator coil strength (DAC units to nm) and orientation relative to framestore axes.

4. The conversion factor for defocus (DAC units to nm).

The last of these is available from the microscope instruction manual but was still verified by diffraction pattern fitting. The magnification conversion factor was found by imaging a <100> Gold film at different electron-optical magnifications. The value at the typically used magnification of 500 KX was determined. Tilt coil strength calibration can be performed by superposing electron diffraction patterns from a Gold <100> film with no tilt, tilt in the x-direction and tilt in the y-direction. From these the tilt coil calibration can be found. An approximate value for tilt coil orientation with respect to the framestore was found by fitting diffractograms and also by directly observing the diffraction pattern. An accurate value for this misalignment is not necessary if the tilt coil calibration matrix is fitted along with lens aberrations. Examining the effects of injected two-fold astigmatism on diffractograms allowed the determination of the stigmator coil strength and orientation.

4.3 XCF Performance

4.3.1 Normal XCF

Three methods can be used to improve peak definition and resolution when cross correlating TEM images. The first is the removal of background variations by least squares fitting the coefficients of a linear function, $Ax+By+C$, to the image and subtracting this ramp from the image. The second is the windowing of the image to avoid spectral leakage (Bendat, J.S. and Piersol, A.G., 1971). The third is the masking out of low frequency components in the reference image to improve contrast.

Figure 4.1 shows the two images used for comparing XCF techniques. The images are those of Gold particles on an amorphous Germanium support film taken from a JEOL 4000 EXII electron microscope (set 7x7c1 - see appendix A). The image on the left was captured at approximately Scherzer defocus

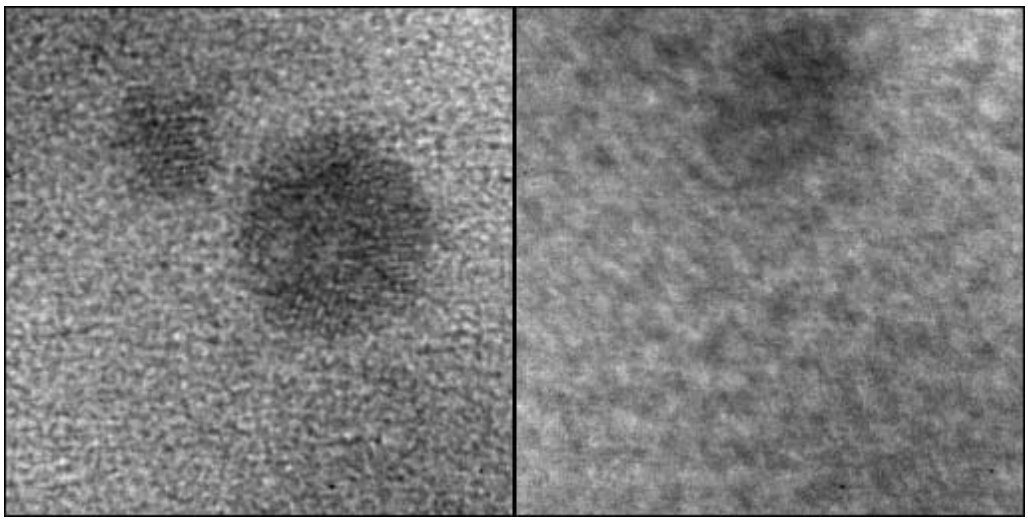


Figure 4.1: Axial and tilted images: gold particles on an amorphous germanium film captured on the JEOL 4000 EXII microscope at 400 kV at 1 Sch underfocus by Dr. A.I. Kirkland. Cs=0.9 mm and field of view is 14.5 nm. The image on the left has no injected tilt and the image on the right has an injected tilt of roughly 17 mrad at an azimuth of $\frac{3\pi}{4}$ from the x-axis.

and close to the voltage centre (<1 mrad misalignment) and that on the right at a tilt of roughly 17 mrad at an azimuth of $\frac{3\pi}{4}$ from the x-axis. The magnification is 500 KX so that the larger particle is about 60 Å in diameter. The TEM accelerating voltage used was 400kV and its spherical aberration constant, C_3, is 0.9mm.

Figure 4.2 shows the XCF of the tilted image with the axial one whilst varying the radius of a circular mask, centred on the origin, in the Fourier transform of latter. The mask simply sets the low frequency components it covers to zero and each image is labeled with its mask radius in pixels. It is noticed that an elliptical peak profile or motif becomes more evident as the mask radius is increased. This motif corresponds to the cross-correlation of the different PSFs in the axial and tilted images as denoted by $c_p(\mathbf{x})$ in equation (3.55). The location of the peak value in the XCFs is marked in the figure. In the last image, fixed noise (from the scintillator for example) has led to a false peak at the centre. It will be shown later (section 4.7.3) that the actual displacement is associated with the centre of the motif. Clearly

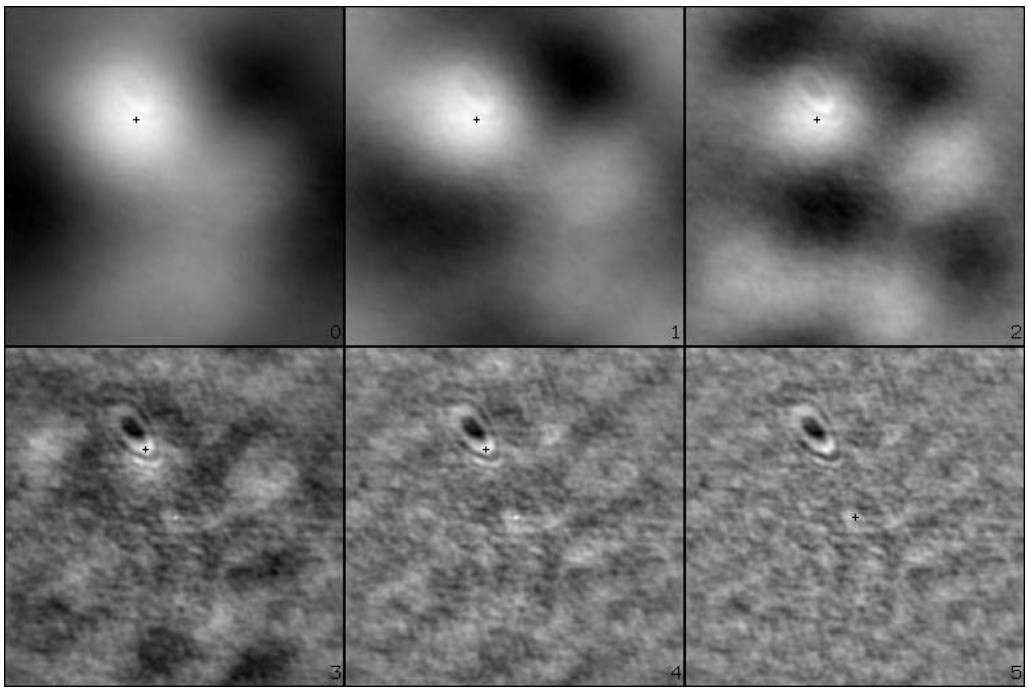


Figure 4.2: Effect on the XCFs of varying the circular mask radius in the Fourier transform of the axial image. The images being cross-correlated are those in figure 4.1. The mask radius is indicated in the bottom right hand corner of the XCF and the peak height is marked.

without some form of phase compensation, a false peak is found because the XCF of the PSFs in each image (centred on the origin) does not have its maximum value at the centre.

A mask radius of 2-4 pixels improves the motif contrast in the XCF whilst retaining a maximum value close to the motif. For a mask radius of 5, the centre pixel has a larger value than those on the 'rim' of the motif and will therefore lead to a highly unreliable estimate of the displacement. Figure 4.3 is a close-up of the images shown in figure 4.2 centred on the motif. Contours have been superimposed to show how the mask radius affects the region around the motif. The maximum value in the region is marked.

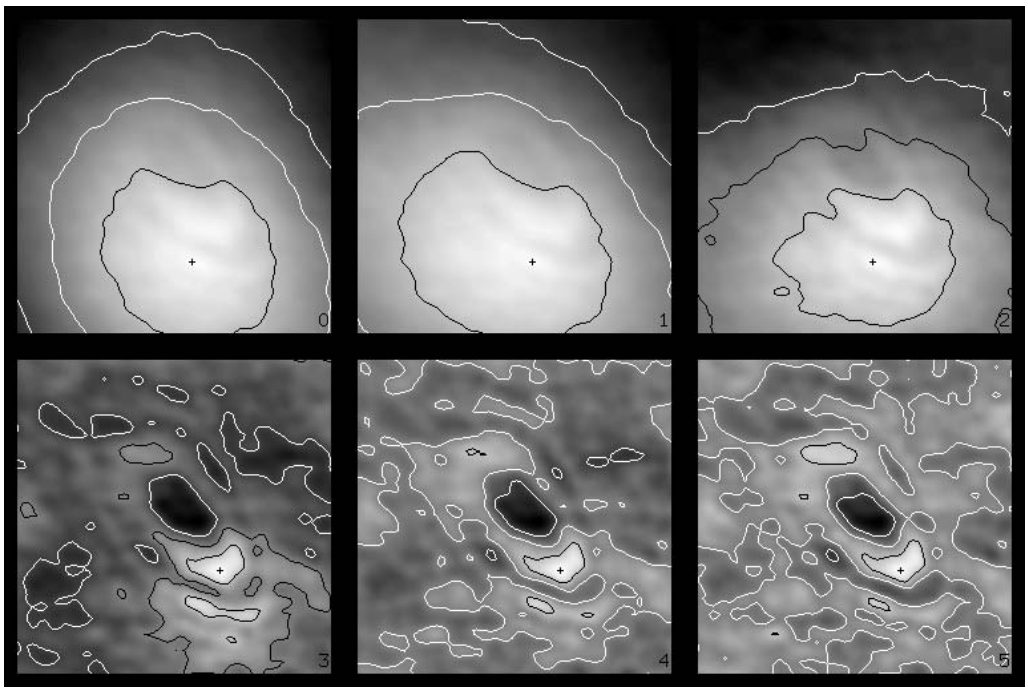


Figure 4.3: Close-up of the effect of varying the mask radius shown in figure 4.2. The peak height in the region is marked.

4.3.2 Generalised XCFs

The denominators for the generalised correlation processors in equations (3.46), (3.49) and (3.52) are ensemble averages. In practice only two observations, $i_1(\mathbf{x})$ and $i_2(\mathbf{x})$ are available so that smoothing these observations is necessary to get a processor estimate. The fact that this provides unsatisfactory results coupled with the knowledge that the processors do not take transfer functions into account has led to their use being discontinued in favour of phase compensated XCFs. The nested-correlation function (Chand, G. and Holburn, D.M., 1993) and the mutual-correlation function (van Heel, M. et al., 1992) are other correlation function derivatives which fail to take transfer functions into account and hence were not considered for analysis. Ref. (Chand, G. and Holburn, D.M., 1993) is reprinted in appendix D.

4.3.3 Phase compensated XCFs

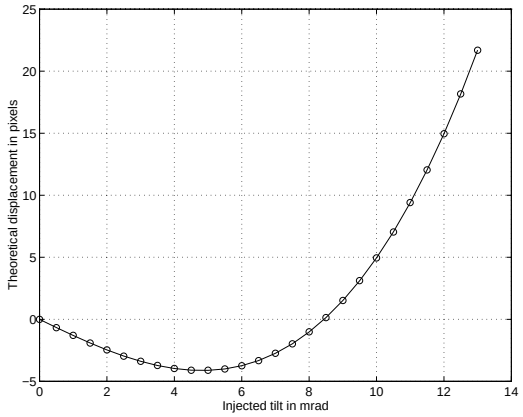


Figure 4.4: Theoretical displacement curve for simulated images of amorphous germanium with injected tilts in the range 0 to 13 mrad in the x-direction at 2 Sch underfocus.

Of the various forms of phase compensated XCFs described in 3.6.1.3, the ST-XCF form (equation 3.68) was chosen for investigation. It is a convenient form as it does not depend upon the estimate of an ϵ SNR value. This particular modified correlation function is equivalent to a double XCF; that of the XCF of the two images with the XCF of their PSFs. The last XCF is subsequently referred to as a template PSF-XCF and it is defined by G_{p1p2} in equation 3.68. Clearly, some *a priori* knowledge of the imaging conditions is necessary to generate this template. It is found, however, that an approximate knowledge of C_1 and τ_c is adequate to compensate the XCF phase. To illustrate this, consider a set of *simulated images* of amorphous germanium with injected tilts in the range 0 to 13 mrad in the x-direction at 2 Sch underfocus. Assume that both the two-fold and three-fold astigmatisms are zero and that the sampling interval is 0.565 Å/pixel. Let C_3 be 0.9mm and the accelerating voltage 400kV, which are the values for the JEOL 4000 EXII. Typical values of beam divergence and focal spread for this instrument are 0.4 mrad and 100 Å respectively. Figure 4.4 shows the theoretical displacement in pixels for the above range of tilts.

Before template PSF-XCFs can be used, they need to be modeled accurately. This includes modeling the effect of the camera modulation transfer

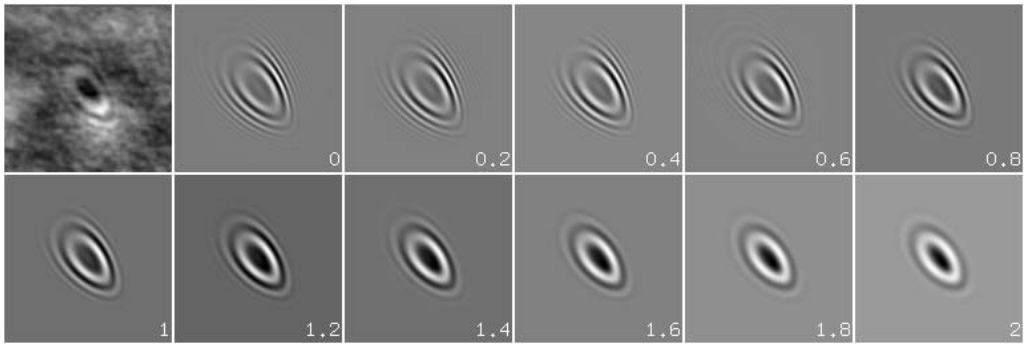


Figure 4.5: Modeling camera MTF using specimen vibration. The magnitude of the vibration is indicated in r.m.s. Å. The top left image is the close-up of the PSF-XCF peak profile (motif) in the XCF of the two images in 4.1. The magnitude of vibration chosen is 1 Å.

.

function (MTF). Assuming that the MTF can be approximated by a Gaussian distribution (Pinson, L.J., 1985), its effect can be simulated by increasing the specimen vibration value used to generate the template PSF-XCFs. Normally the r.m.s. specimen vibration value is approximately 0.5 Å based on matching simulated images with experimental ones. Figure 4.5 shows the effect of the camera MTF as modeled by the specimen vibration for values 0 to 2 Å. The top left image is a close-up of the PSF-XCF peak profile (motif) in the XCF of the two images in 4.1 with a Fourier plane mask radius of 3 pixels. The other images are template PSF-XCFs with the specimen vibration values and hence effective MTF values in r.m.s. Å indicated. A value of 1 Å for the specimen vibration was chosen as a suitable value for subsequent generation of template PSF-XCFs.

Figure 4.6 illustrates how template PSF-XCFs can be used to compensate the phase in the original XCF for the simulated azimuthal tilt set described above. The columns correspond to tilt magnitudes in steps of 1 mrad from 0 mrad on the left to 11 mrad on the right. The first row is the raw XCF of each tilted image with the axial. The second is the PSF of each image. The third is the XCF of each PSF with the axial PSF (the template PSF-XCF). The fourth is the ST-XCF based on the template PSF-XCF and the raw XCF. Note that the highest point in the ST-XCF is always at the centre.

Figures 4.7 and 4.8 show how the error between the ST-XCF peak position

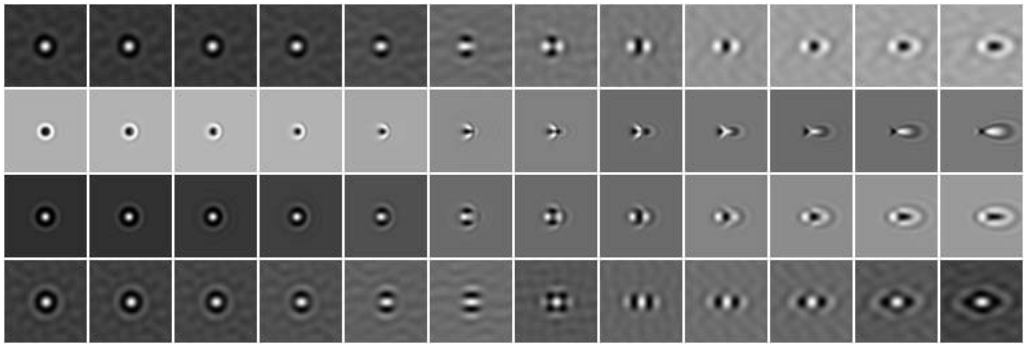


Figure 4.6: Components of ST-XCFs for the simulated images of amorphous germanium. The columns correspond to tilt magnitudes in steps of 1 mrad from 0 mrad on the left to 11 mrad on the right. The first row is the raw XCF of each tilted image with the axial. The second is the PSF of each image. The third is the XCF of each PSF with the axial PSF (the template PSF-XCF). The fourth is the ST-XCF based on the template PSF-XCF and the raw XCF.

and the analytical value varies with tilt. Ideally, the ST-XCF peak position should be zero as the template PSF-XCF is not centred on the origin. Thus the error shown in these figures is actually the peak position in the ST-XCF with non-zero centred template PSF-XCFs. The XCF peak position error is also shown from which it evident that the XCF is highly unreliable especially for large tilt values. This is due to the XCF peak region becoming an elliptical ring structure with its highest value on the ring rather than at the centre. Figure 4.9 shows how the peak height in the ST-XCF and the raw XCF varies with tilt. The XCF value drops with tilt as the peak region spreads out more. The analysis above assumes that the defocus value is known in advance.

Figures 4.10,4.11 and 4.12 show how the ST-XCF performs when the template PSF-XCFs are calculated at defoci other than 2 Sch. Values 0.5 Sch either side of this value have been used. The tolerance of the ST-XCF to such large errors in defocus estimates is high : approximately a 6 Å error in displacement for a 200 Å error in defocus.

Figures 4.13,4.14 and 4.15 show how the ST-XCF performs when the template PSF-XCFs are calculated with injected x-direction misalignments other than 0.25 Gl/Sch, for which purpose a new set of simulated images was generated. Values of 0 and 0.125 Gl/Sch were used. Again, the tolerance of

the ST-XCF to misalignment estimation errors is high: an error of roughly 6 Å in displacement for a 0.8 mrad error in misalignment.

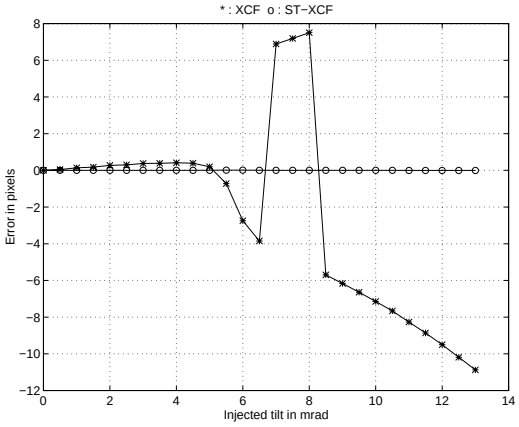


Figure 4.7: X difference

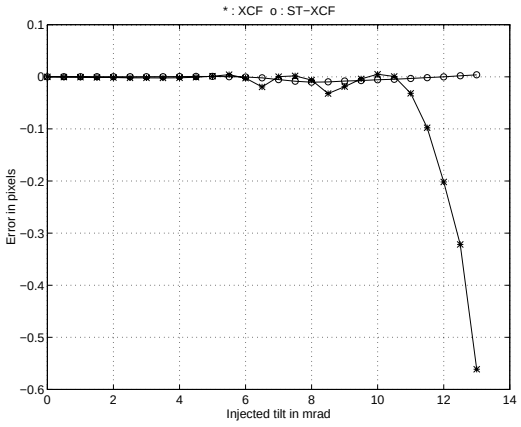


Figure 4.8: Y difference

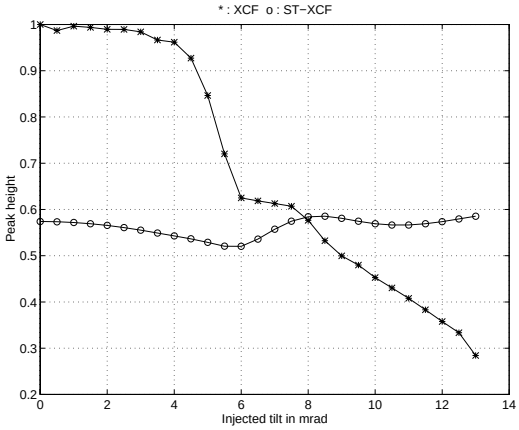


Figure 4.9: ST-XCF and XCF Peak heights

Figure 4.16 shows how the PC-XCF (equation 3.62) and TD-XCF (equation 3.71) compare to the ST-XCF. The XCF being phase compensated is that between the two images in figure 4.1. A specimen vibration value of

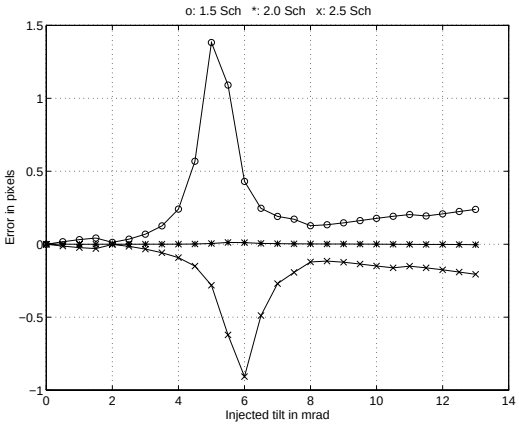


Figure 4.10: X difference

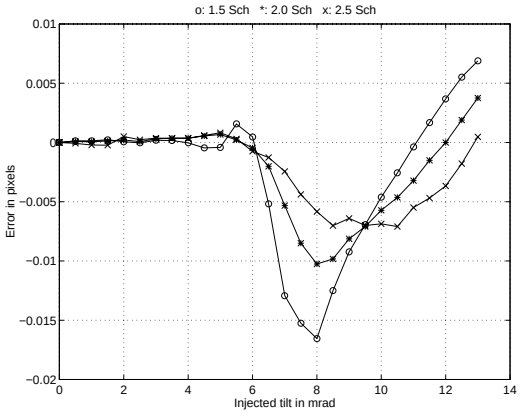


Figure 4.11: Y difference

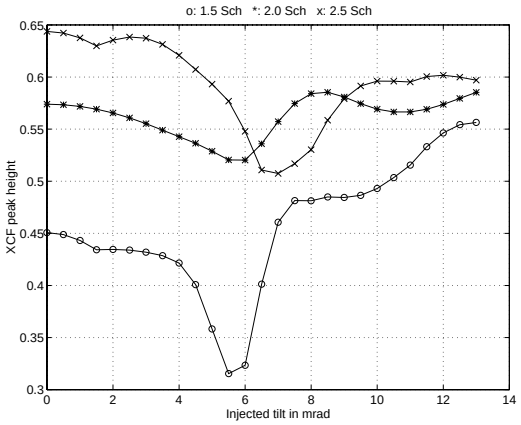


Figure 4.12: ST-XCF Peak heights for different defoci

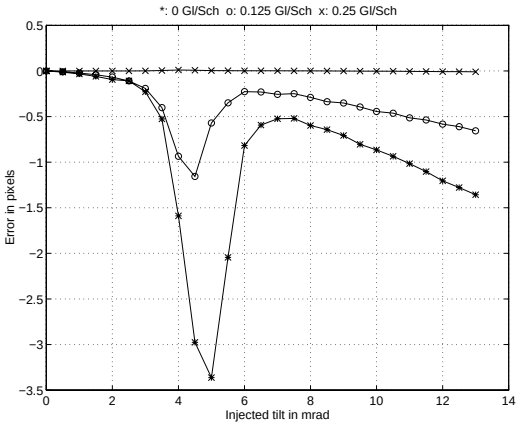


Figure 4.13: X difference

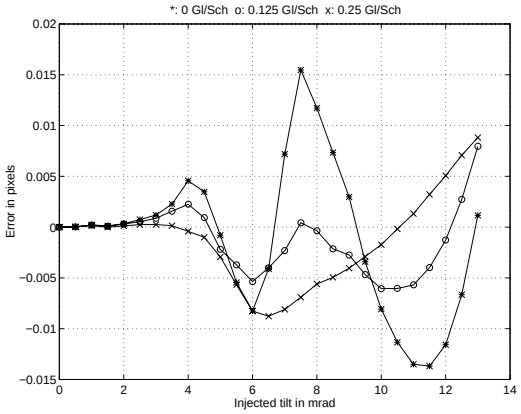


Figure 4.14: Y difference

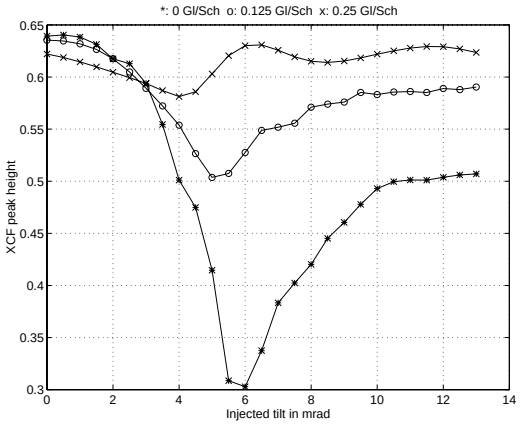


Figure 4.15: ST-XCF Peak heights for different tilts

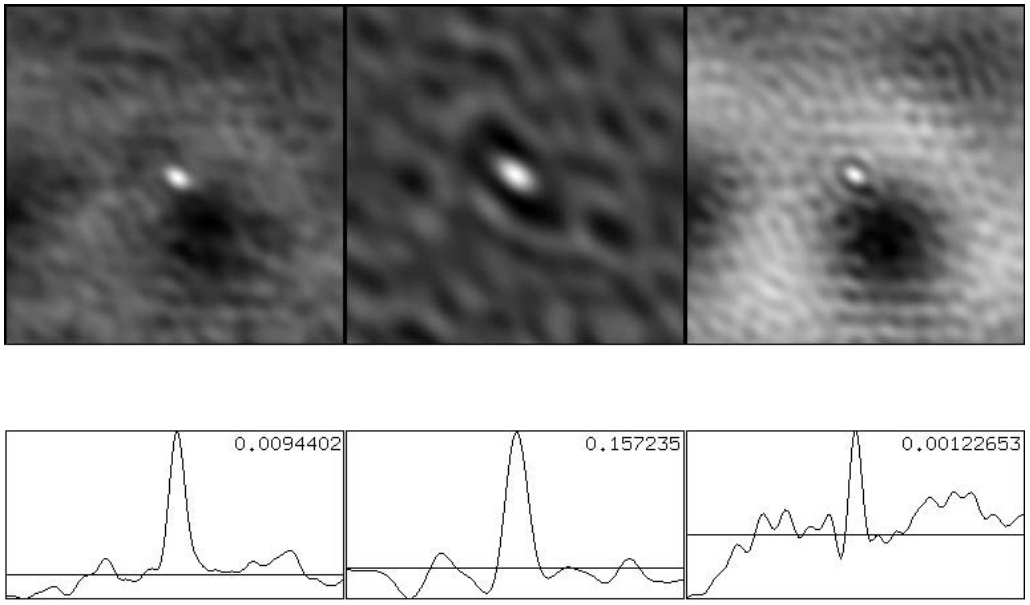


Figure 4.16: Left to right : PC-XCF, ST-XCF and TD-XCF of the two images in figure 4.1

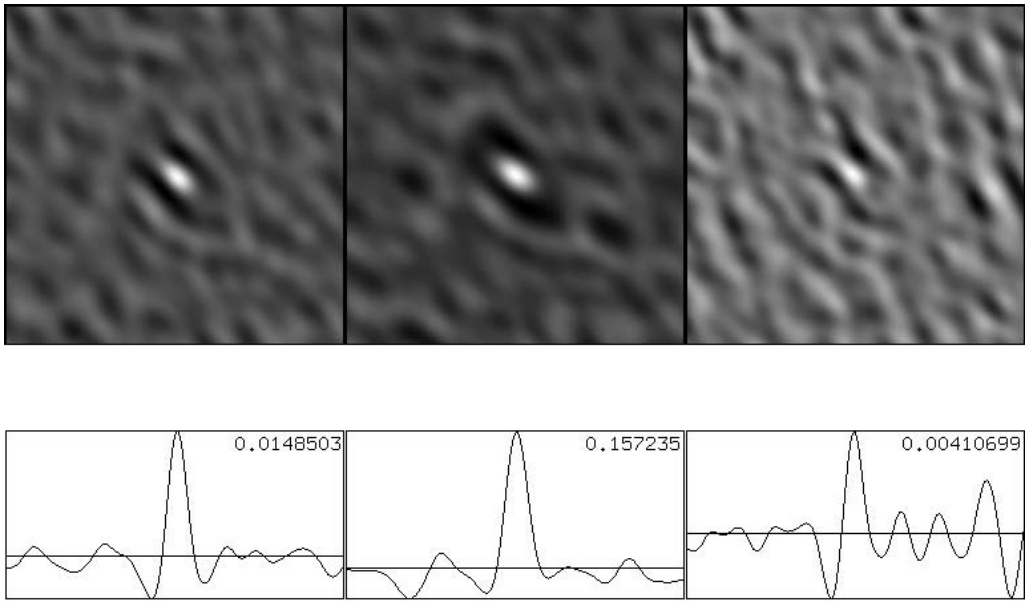


Figure 4.17: Left to right : Modified PC-XCF, ST-XCF and modified TD-XCF of the two images in figure 4.1

1 Å was used to generate the template PSF-XCFs. The top row shows a close-up of each transform centred on location of the peak value and the bottom row shows a cross-section through the peak region along with a numerical value for the peak height. The left column shows the PC-XCF with $\epsilon = 0.1$, the middle, the ST-XCF and the right, the TD-XCF with $\epsilon = 1$. As expected, the ST-XCF is pessimistic compared to the PC-XCF while the TD-XCF is optimistic. Indeed, due to the noisy nature of the TD-XCF, a value of 0.1 for ϵ resulted in a false peak being located.

Figure 4.17 shows how the modified forms of the PC-XCF (equation 3.65) and TD-XCF (equation 3.74), which do not require ϵ, compare to the ST-XCF. It is evident that the modified PC-XCF and especially the modified TD-XCF peak location estimates are less reliable than the unmodified transforms due to the effects of the object functions in the denominators.

The effect of phase doubling and transfer squaring on axial XCFs is shown in figure 4.18. The top row shows typical XCFs obtained from a JEOL 4000 EXII focal series with defocus values in the range 1 Sch underfocus to 2 Sch overfocus. The middle row shows the phase-doubled XCF and the bottom row, the transfer squared XCF. The peak in each modified XCFs is ringed. It is disappointing to see the middle transfer squared XCF failing to produce a peak in the right place considering that the transform is expected to be more resistant to the effects of noise than the phase-doubled XCF.

4.4 Cross-power Spectrum Phase Angles

As noted in section 3.6.1.4, the linear phase ramp in the cross-power spectrum is adversely affected by the different coherence envelopes. Figure 4.19 shows a section through the cross-power phase spectrum of the two images in figure 4.1. The section is at the same azimuth with a line in the XCF which joins the origin and the PSF-XCF motif. The sampling interval is about 0.04 Gl^{-1}. The phase spectrum appears to be so noisy that a ramp would be difficult to fit even after phase doubling.

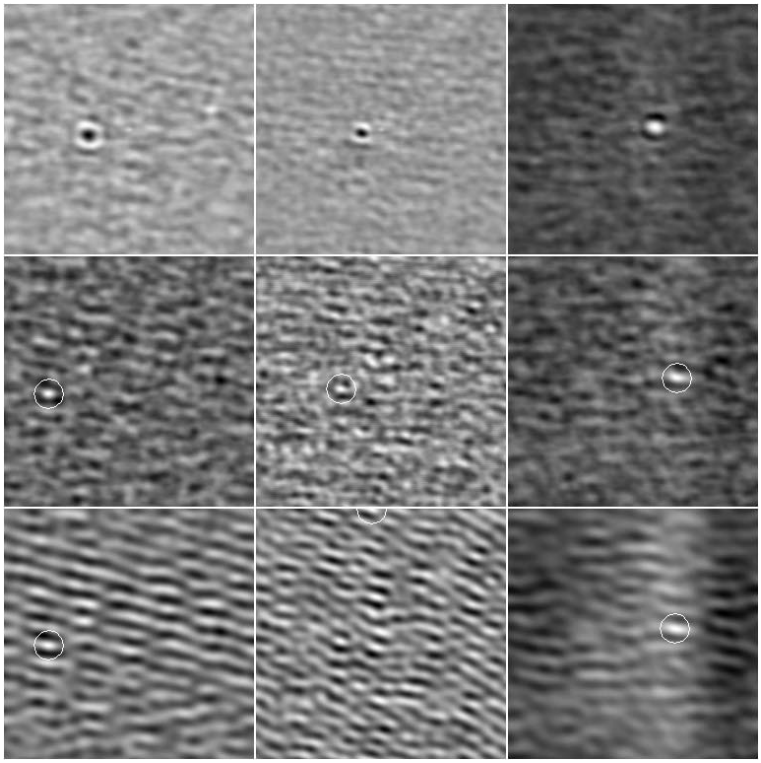


Figure 4.18: Phase doubling and transfer squaring for typical axial XCFs. The top row shows XCFs obtained from a JEOL 4000 EXII focal series with defocus values in the range 1 Sch underfocus to 2 Sch overfocus. The remaining microscope parameters are identical to those shown in figure 4.1. The middle row shows the phase-doubled XCF and the bottom row, the transfer squared XCF.

4.5 The Five-tilt Displacement Method

To investigate this method, two sets of images were examined. Both sets contain 10 images - two axial captured at the beginning and end, four at orthogonal tilts of 1 Gl/Sch and four at orthogonal tilts of 2 Gl/Sch. One set was captured at 4 Sch and the other at 8 Sch underfocus (10mx4u and 10mx8u : see appendix A).

Only the 1 Gl/Sch tilt 'crosses' were analysed. Figures 4.20 and 4.21 show the steps involved in the method. The top row is for the processing of tilts in the x-direction and the bottom, for the tilts in the y-direction. The

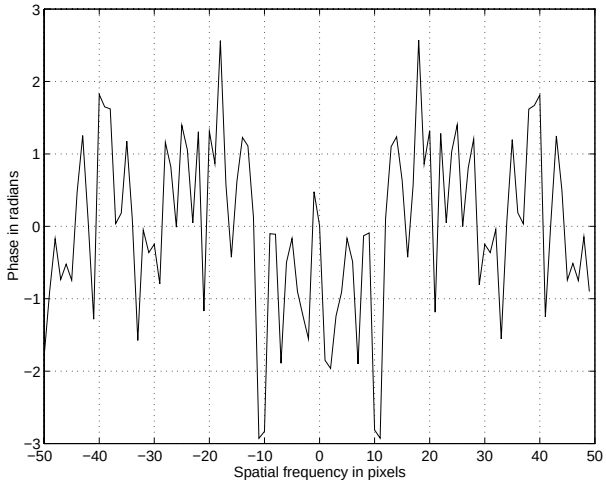


Figure 4.19: The phase of the cross-power spectrum for the two images in figure 4.1

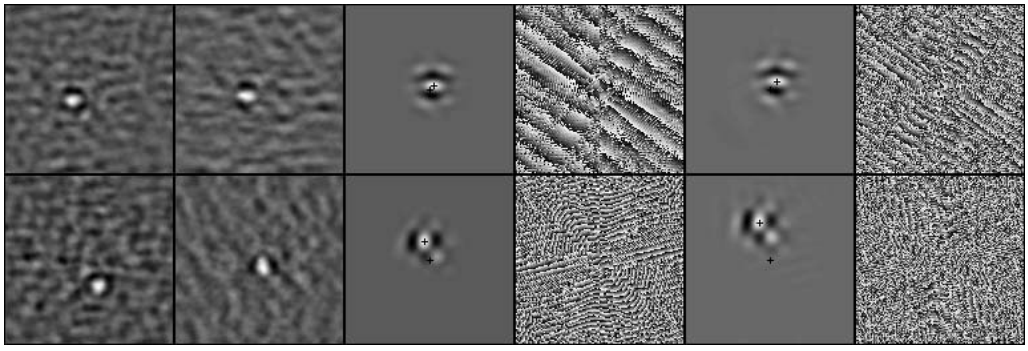


Figure 4.20: Five-tilt method applied to 4 Sch set shown in figure A.2. The top row is for processing tilts in the x-direction and the bottom, for tilts in the y-direction. The first two columns are the XCFs of each tilted image with the axial image. The third and fourth columns are the XCF of these XCFs and the phase spectrum respectively. The fifth and sixth columns are the results of phase doubling this nested XCF and its phase spectrum respectively.

first two columns are the XCFs of each tilted image with the axial image. The third and fourth columns are the XCF of these XCFs and the phase spectrum respectively. The area outside a circular region of radius 32 pixels

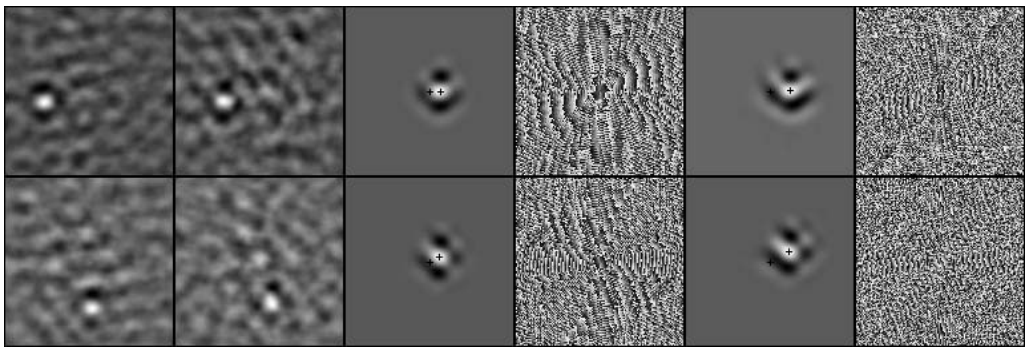


Figure 4.21: Five-tilt method applied to 8 Sch set. The parameters are identical to those in A.2 except that the defocus was set to 8 Sch underfocus. For details of the image descriptions see figure 4.20.

.

around the peak in each of the first two XCFs was masked out in order to give the third 'nested' XCF better definition. The fifth and sixth columns are the results of phase doubling this nested XCF and its phase spectrum respectively. The peak location in each XCF is marked as is its origin. The *linear* phase ramp in the phase spectra of the nested XCFs is quite evident.

	pixels				mrad		μm	
Set	d_{xx}	d_{xy}	d_{yx}	d_{yy}	$\hat{\tau_{cx}}$	$\hat{\tau_{cy}}$	$\hat{A_{2r}}$	$\hat{A_{2i}}$
4 Sch	5.11	6.45	-8.69	28.38	0.33	-3.22	2.60	-0.75
8 Sch	14.68	1.19	13.85	8.04	-2.64	-0.85	-2.23	-0.37

Table 4.1: Five-tilt method results

The results of the five-tilt method are shown in table 4.1. The values for d_{xx}, d_{xy}, d_{yx}, d_{yy} are obtained by simply halving the phase doubled XCF peak location estimate. Linear specimen drift, to which this method is susceptible, is probably the cause of the disagreement between the two sets and, as other methods will show later, the unreliability of the results. The images were examined off-line so that the second step in the five-tilt method i.e. the elucidation of defocus and two-fold astigmatism could not be performed.

4.6 Least Squares Displacement Fit

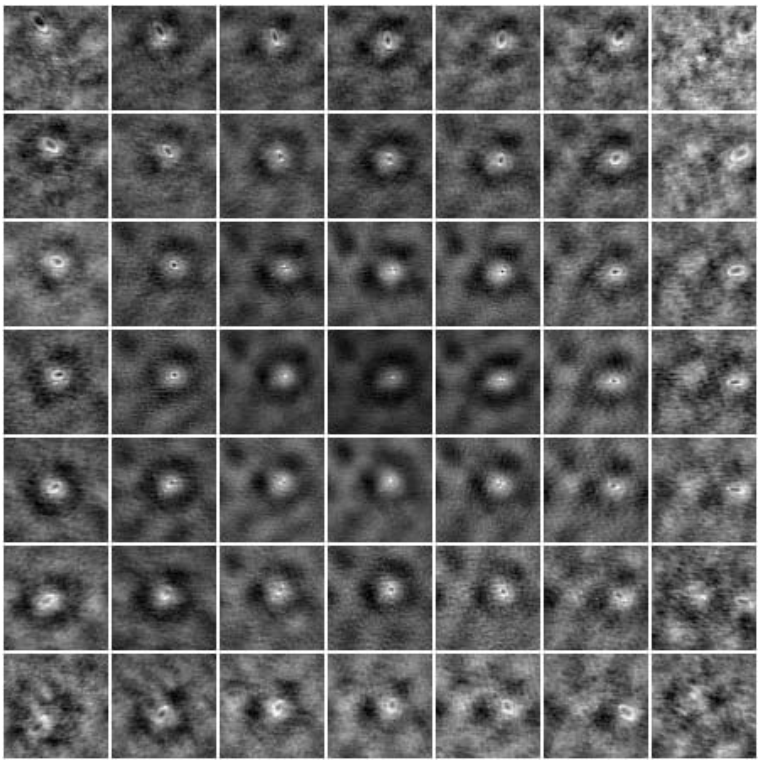


Figure 4.22: Raw XCF for tilt set 7x7c1 whose parameters are defined in figure A.1

A 51 member set (7x7c1) and a 10 member set (10mx8u) (see appendix A) were used to explore the least squares (LS) fitting of displacements. The Semper program used to capture the data was based on one provided early on in the project by Saxton (Saxton, W.O., Private Communication). The larger of the two sets is considered first: its raw XCF is shown figure 4.22. Clearly, using the location of the peak values from the XCFs in a LS fit would lead to unreliable estimates of aberration coefficients as pointed out in section 4.3.3. From the same section, it can be concluded that template PSF-XCFs can be used to compensate the phase in the XCF even if the defocus and misalignment are not known accurately. As the microscope was aligned to the voltage centre (< 1 mrad from the coma-free axis) and

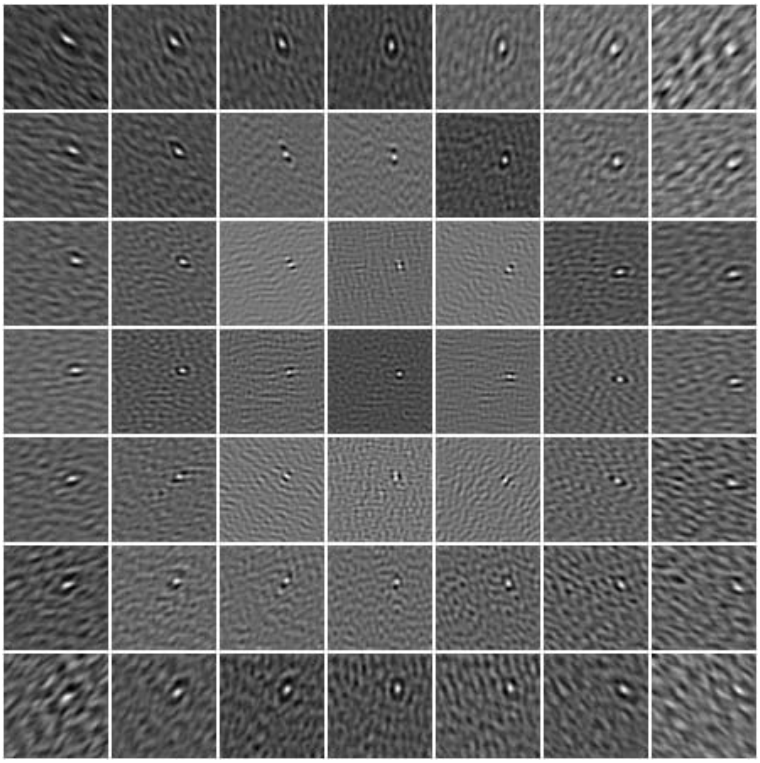


Figure 4.23: ST-XCF for tilt set 7x7c1 whose parameters are defined in figure A.1

set to approximately 1 Sch underfocus, template PSF-XCFs were calculated based on zero misalignment and 1 Sch underfocus. ST-XCFs were then generated in which the peak location is likely to be a good estimator of image displacement. This peak position was located with sub-pixel accuracy by fitting a biquadratic function (Chau, K.-L., 1993) to the square block of 5x5 pixels around the pixel with the highest value. This corresponds to an accuracy of better than 0.15 Å in displacement measurement. The ST-XCF tableau for the set is shown in figure 4.23.

It is interesting to examine how the peak height of the ST-XCF varies across the set. Figure 4.24 represents this peak height in such a way that black is 0.03 and white is 0.13. It is noticed that low peak heights can be found around the central zero-tilted image. These tilts correspond to small displacements and it can be deduced from figures 4.4, 4.10 and 4.13 that

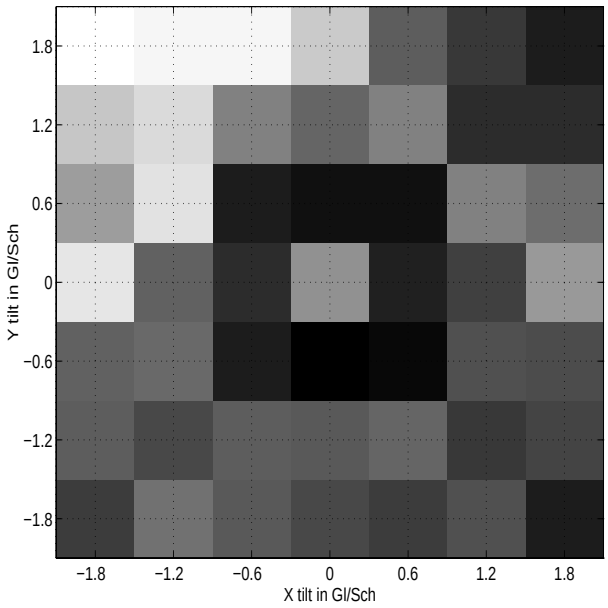


Figure 4.24: ST-XCF peak heights for tilt set 7x7c1 where black is 0.03 and white is 0.13

the use of approximate template PSF-XCFs is likely to lead to inaccurate displacement values at these tilts.

Before the ST-XCF peak positions or displacements can be analysed, drift compensation needs to be performed unless drift is fitted, as proposed earlier. This ensures that the displacement recorded for axial images is zero and is most easily accomplished by examining the values for displacement returned by the ST-XCF of the original axial image with the two other axial images in the set. Two linear ramps can then be fitted between these displacement values for the x- and y-directions as indicated in figure 4.25. In this figure, tilted images 25 and 50 correspond to the second and last axial images in the set. The ramps obtained this way are then subtracted from the displacement measurements from the other images in the set. Ideally, the ramps should be collinear indicating that drift is predominantly linear. However for this set, as the figure shows, the drift in the x-direction is not linear and in the worst case, drift compensation might result in additional drift being introduced to the displacements.

A Fortran program to perform linear and non-linear least squares fitting of

Coeff.	Units	Fit I	Fit II	Fit III	Fit IV	Fit V	Fit VI	Fit VII	Fit VIII
A_{1r}	nm	-0.4	-0.2	7.0	7.9	3.4	3.1	19.5	5.5
A_{1i}	nm	-16.9	19.3	8.4	8.4	5.6	5.5	13.1	17.8
C_1	nm	-8.6	-9.0	-19.8	-33.2	-28.7	-28.3	-37.4	-31.1
A_{2r}	μm	-	1.38	0.53	0.60	0.79	0.79	0.74	0.54
A_{2i}	μm	-	-1.31	-1.70	-1.64	-1.72	-1.78	-1.56	-1.61
τ_{cx}	mrad	-0.30	-0.44	-0.49	-0.51	-0.86	-0.85	-1.34	-0.46
τ_{cy}	mrad	-0.42	-0.55	-0.44	-0.74	-0.52	-0.57	0.97	-0.73
θ_1/θ	$^\circ$	-	-	-12.3	-12.3	-10.8	-12.6	-4.3	-13.6
θ_2	$^\circ$	-	-	75.5	75.4	77.1	75.3	65.5	-
a/r		-	-	1.01	1.01	1.02	1.02	1.10	1.04
b		-	-	1.07	1.07	1.06	1.06	1.11	-
A_{3r}	μm	-	-	-	-	-	-35	-40	-
A_{3i}	μm	-	-	-	-	-	-19	-18	-
B_{3r}	μm	-	-	-	-	-	-	-287	-
B_{3i}	μm	-	-	-	-	-	-	-359	-
ΔC_1	nm	-	-	-	0.53	0.29	0.38	0.65	0.46
Δx	nm	-	-	-	-	-0.006	-0.006	-0.004	-
Δ_y	nm	-	-	-	-	0.001	0.001	0.002	-
E_{rms}	nm	0.427	0.383	0.218	0.218	0.205	0.198	0.197	0.224

Table 4.2: 7x7c1 displacement fits

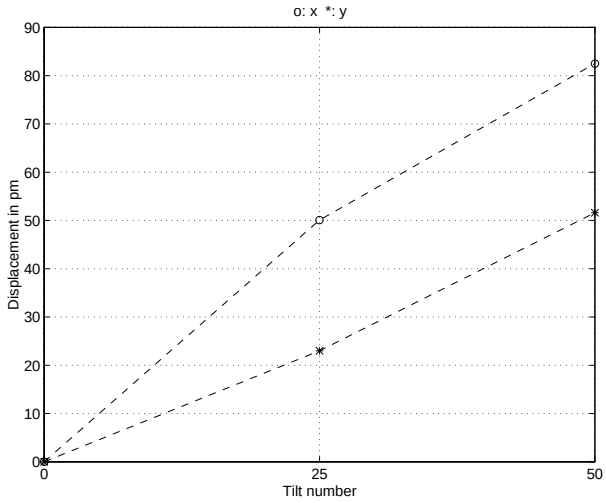


Figure 4.25: Drift compensation in 7x7c1 for x- and y-directions

aberrations (see section 3.6.5) using both the displacement and diffractogram data was provided by Saxton (Saxton, W.O., Private Communication) at the initial stages of the project. This was ported to C and enhanced to incorporate A_3, B_3, ΔC_1, Δ_x and Δ_y. Table 4.2 shows the results from the fit of drift compensated displacements. Different numbers of parameters have been fitted and the r.m.s. error returned by the least squared fit, E_{rms}, is indicated. The dashes in the table show which parameters have been fixed at zero for the particular fit. The value for C_3 was fixed at 0.9 mm throughout and the first tilted image is what the focus and linear specimen drift have as their reference. It is clear from fits II and III that the TCC matrix has a dramatic effect in reducing the r.m.s. error. The values for residual drift, Δx and Δy show that Δx is the larger of the two which is what one would expect based on the previous discussion related to the possible non-linear nature of the drift in the x-direction. The values for four-fold astigmatism are likely to be more accurate than those for axial star aberration considering the adverse effect the latter has on fitted parameter values in fit VII. This is probably because the axial star aberration coefficients are effectively soaking up the noise present in the set. It should also be noted that the inclusion of A_3 and B_3 does not lead to a significant reduction in E_{rms}. The TCC matrix values indicate an effective rotation of about 12° clockwise. It will be shown

later that the typical value for this rotation obtained from diffractogram fits is about 4°. One possible explanation for this rotation is the non-linear fit finding a false minimum. This can be investigated by using a two parameter non-linear fit based on $\mathbf{T_{cc_{2a}}}$ in equation (3.120).

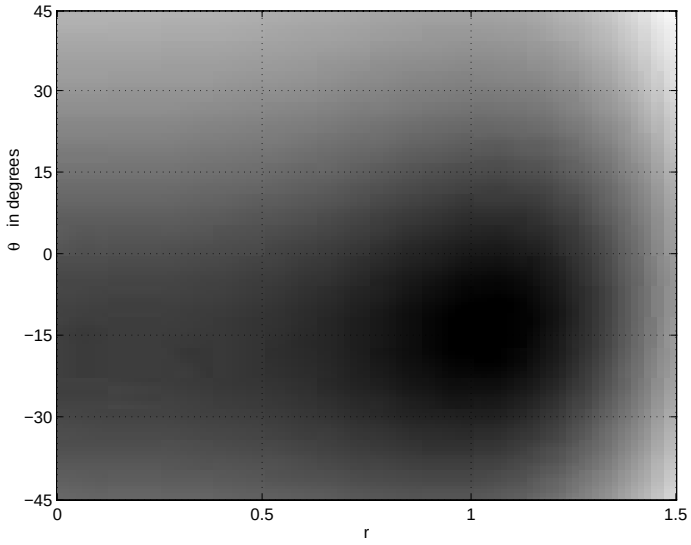


Figure 4.26: 7x7c1 r.m.s. error for two parameter non-linear fit

Figure 4.26 shows how the r.m.s. error varies with r and θ when $\mathbf{T_{cc_{2a}}}$ is fitted. The fitted values for these two variables are 1.04 and $-13.6°$ as shown in Fit VIII in table 4.2. The minimum found is clearly a global minimum over the large range of values for r and θ shown in the figure. Assuming that a similar 4-dimensional minimum exists when the general TCC matrix is fitted, which is reasonable considering how close the above fitted results are to fit IV, false minima can probably be ruled out as explanations for TCC matrix rotations.

An alternative explanation is inadequate initial drift compensation due to its non-linear nature. From table 4.3, fit III for the 10 member set results in an effective TCC matrix rotation of about 15° anti-clockwise. When this is compared with 12° *clockwise* for set 7x7c1, non-linear drift is not ruled out as the cause of the rotation. The problem is discussed further in section 4.8 but remains unresolved.

Figure 4.27 shows how the error in the fit of displacements (as far as

Coeff.	Units	Fit II	Fit III	Fit VIII
A_{1r}	nm	3.2	0.4	3.9
A_{1i}	nm	-4.8	2.9	-4.0
C_1	nm	-304.8	-313.9	-313.6
A_{2r}	μm	1.63	1.83	1.86
A_{2i}	μm	-1.18	-0.52	-0.49
τ_{cx}	mrad	-0.38	-0.43	-0.37
τ_{cy}	mrad	-0.59	-0.59	-0.57
θ_1/θ	$^\circ$	-	17.7	15.3
θ_2	$^\circ$	-	102.8	-
a/r		-	1.02	1.02
b		-	1.02	-
E_{rms}	nm	0.351	0.087	0.088

Table 4.3: 10mx8u displacement fits

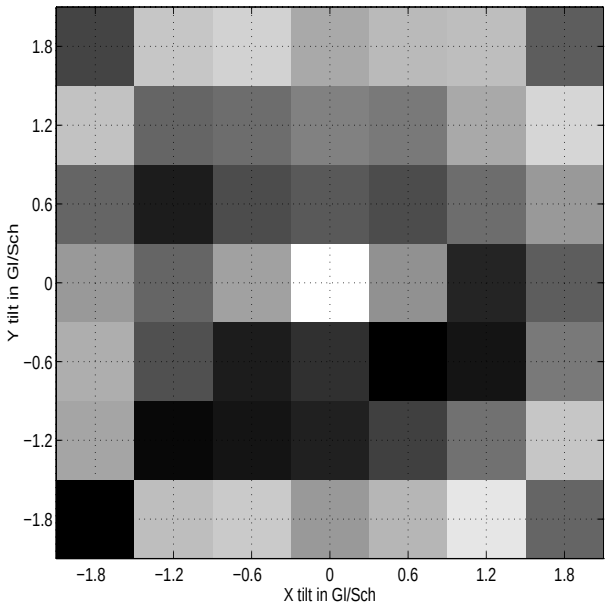


Figure 4.27: Modulus of 7x7c1 displacement fit error for fit 4 where white is 0 nm and black is 0.54 nm

moduli are concerned) varies with position for fit IV of the large set. White corresponds to an error of 0 nm and black to an error of 0.54 nm so that this figure can be compared with 4.24 directly. It is noticed that some of the largest 'misfits' occur around the central axial value just like the smaller ST-XCF peak heights.

4.7 Diffractogram Fitting

4.7.1 Manual implementation

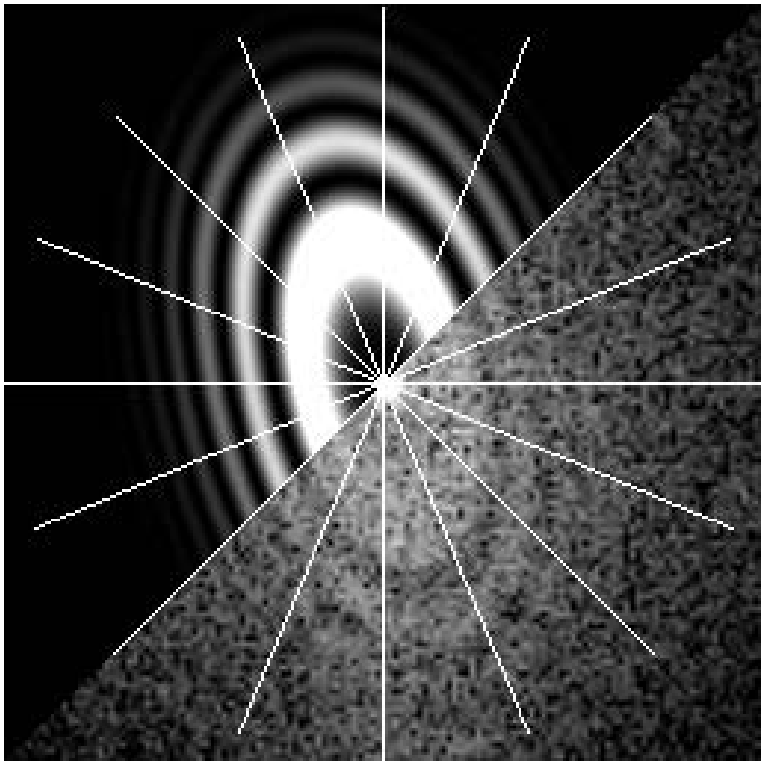


Figure 4.28: The fitting of a typical diffractogram where the lines indicate the different azimuths at which the ring positions in the theoretical diffractogram have been matched by eye to those in the experimental one.

Equation (3.103) implies that effective defocus values have to be measured at a number of different azimuths in order to derive values for C_1' and A_1'.

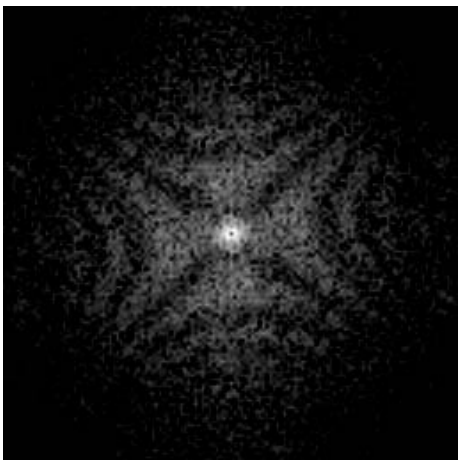


Figure 4.29: A Maltese Cross diffractogram

One way to implement this is the manual comparison of a experimental diffractogram with a theoretical one.

Figure 4.28 illustrates the fitting of a typical diffractogram. The lines indicate the different azimuths at which the ring positions in the theoretical diffractogram have been matched by eye to those in the experimental one. The procedure takes 2-3 minutes depending on how accurately the diffractogram needs to be fitted.

There are some diffractograms, shaped like Maltese Crosses, which are very difficult to fit accurately. An example is shown in figure 4.29 which is the diffractogram for a 2 Gl/Sch injected tilt at a defocus of 8 Sch underfocus from set 10mx8u (see Appendix A). The defocus in these diffractograms is effectively overfocus along one axis of symmetry and underfocus along the other axis of symmetry. Another source of problems is the fact that overfocus and underfocus diffractograms look almost identical far from focus. This means that an approximate knowledge of imaging conditions and tilts is necessary before diffractograms are fitted. Diffractograms captured from an overfocused microscope at arbitrary tilt will always have an effective overfocus. As it is possible to set the microscope to overfocus with a high degree of confidence, it should be possible to get diffractograms without a defocus ambiguity.

4.7.2 The four-tilt diffractogram method

Coeff.	Units	Fit 1	Fit 2	Fit 3
t_1	mrad	3.9	7.8	11.8
A_{1r}	nm	0.0	-3.6	-4.5
A_{1i}	nm	-1.8	1.8	-9.8
C_1	nm	1.3	2.8	10.4
A_{2r}	μm	1.33	1.69	1.67
A_{2i}	μm	-0.99	-1.13	-1.37
C_3	mm	0.93	0.94	0.93
τ_{cx}	mrad	-0.64	-0.64	-0.68
τ_{cy}	mrad	-0.16	-0.26	-0.29

Table 4.4: 7x70 4-tilt fits

Coeff.	Units	c0 Fit 1	c0 Fit 2	c0 Fit 3	c1 Fit 3	c2 Fit 3
t_1	mrad	3.9	7.8	11.8	11.8	11.8
A_{1r}	nm	-2.8	-4.2	-6.7	-3.9	-6.8
A_{1i}	nm	-0.9	2.1	5.0	9.0	3.5
C_1	nm	4.3	4.5	24.6	-28.5	-69.3
A_{2r}	μm	1.42	1.68	1.33	1.49	1.40
A_{2i}	μm	-1.13	-1.13	-1.35	-0.99	-1.23
C_3	mm	0.94	0.97	0.89	0.92	0.94
τ_{cx}	mrad	-0.55	-0.54	-0.43	-0.49	-0.43
τ_{cy}	mrad	-0.21	-0.22	-0.42	-0.39	-0.31

Table 4.5: 7x7c0, 7x7c1 and 7x7c2 4-tilt fits

Tables 4.4 and 4.5 show the results of applying the four-tilt diffractogram method to orthogonal tilts present in sets 7x70, 7x7c0, 7x7c1 and 7x7c2 (see appendix A). Three values for the injected tilt, t_1, can be extracted from each set. However only one value for t_1 (11.8 mrad) could be used for sets 7x7c1 and 7x7c2 because of the difficulty involved in reliably fitting the diffractograms for other values of t_1 in these sets.

For the 7x70 and 7x7c0 sets, the most reliable estimates are likely to be those for t_1=3.9 and 7.8 mrad as their defocus estimate is closer to the expected value. The reasons for this are the omission of focal drift and tilt coil calibration in the method which will lead to the greatest errors for large magnitude injected tilts. Unfortunately such observations are the only ones available for sets 7x7c1 and 7x7c2 so that the estimates obtained from them are expected to be unreliable.

All the estimates for A_2 have been plotted in figure 4.30 and those for τ_c in 4.31. It is very encouraging to see that both these sets of estimates have a low variance. The low variance for τ_c is expected because the microscope was aligned to the voltage centre for each of these sets.

4.7.3 Least squares diffractogram fit

Clearly, in order to make the most of diffractogram measurements, the results should be least squares fitted. Table 4.6 shows the results from fitting the diffractograms from the typical set, 7x7c0. This can be compared with table 4.2. The accuracy of the effective defocus and astigmatism determination from the diffractograms is roughly 10 nm . It is found that the majority of diffractograms in sets at Gaussian focus can be fitted. However, due to the low contrast of such sets, displacement measurements are not possible. This is the reason why set 7x7c1 has its displacement fit tabulated and set 7x7c0 has its diffractogram fit tabulated. The dashes in the table indicate parameters which have been fixed at zero for the particular fit. C_3 is fixed at 0.9mm and the first axial image is what the focus and linear specimen drift have as their reference. As for least squares displacement fitting, the inclusion of TCC matrix elements has a considerable effect on reducing the r.m.s. error. The effect of fitting axial star aberration is not as detrimental as in the displacement fit. This is probably because diffractograms fits are less prone to errors caused by noise and specimen drift.

Table 4.7 shows the results of a fit of type IV applied to all the 7 by 7 sets. Fit IV was chosen as it avoids possible errors arising from the fit of A_3 and B_3. Table 4.8 shows the results of fit V as far as A_3 is concerned and table 4.9 shows the results of fit VI as far as A_3 and B_3 are concerned. There is little consistency in the estimates of A_3 and B_3 from set to set and it is thought that these variables simply reflect the errors from the diffractogram fits. However, the estimates of A_3 at least indicate its maximum magnitude which is less

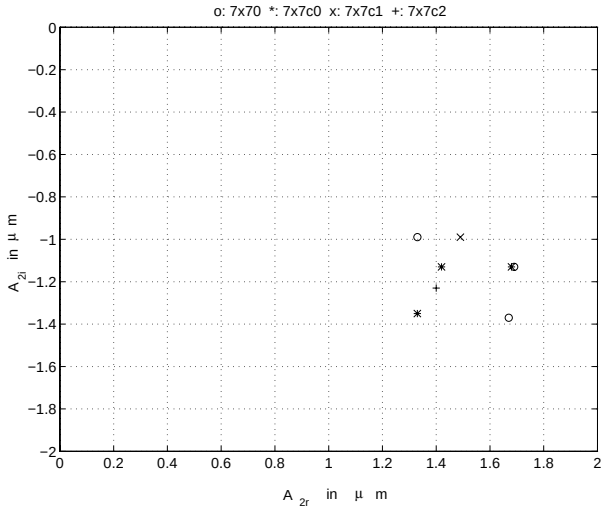


Figure 4.30: Estimates of A_2 from 4-tilt fits

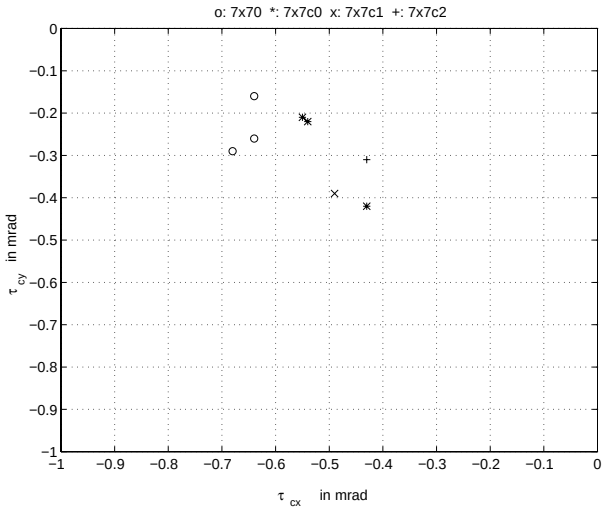


Figure 4.31: Estimates of τ_c from 4-tilt fits

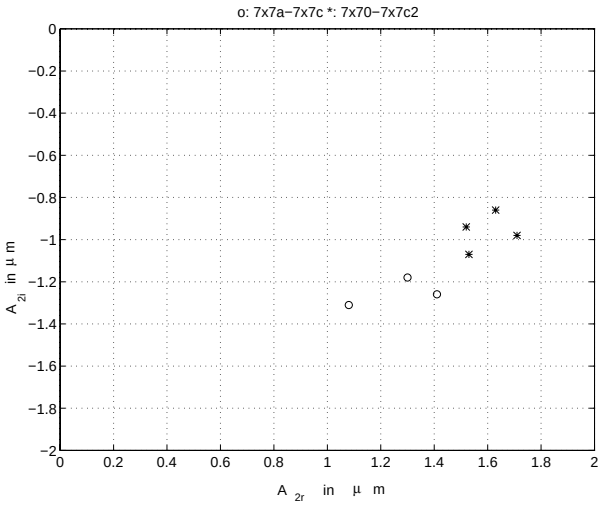


Figure 4.32: Estimates of A_2 from fit IV

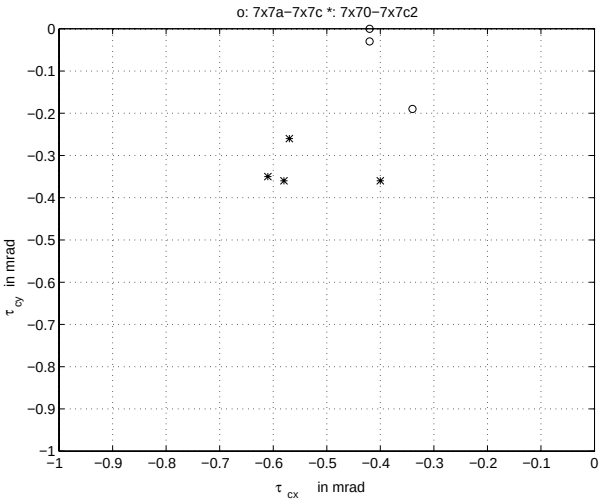


Figure 4.33: Estimates of τ_c from fit IV

Coeff.	Units	Fit I	Fit II	Fit III	Fit IV	Fit V	Fit VI
A_{1r}	nm	-1.3	-2.5	-0.8	-0.9	-0.9	-5.1
A_{1i}	nm	3.1	5.2	-1.9	-2.0	-2.0	1.0
C_1	nm	19.3	19.3	4.8	8.2	8.4	8.0
A_{2r}	μm	-	1.52	1.52	1.52	1.52	1.51
A_{2i}	μm	-	-1.01	-0.94	-0.94	-0.94	-0.93
τ_{cx}	mrad	-0.57	-0.57	-0.56	-0.57	-0.57	-0.58
τ_{cy}	mrad	-0.30	-0.30	-0.30	-0.26	-0.26	-0.27
θ_1	$^\circ$	-	-	3.8	3.8	3.8	4.9
θ_2	$^\circ$	-	-	90.2	90.3	90.2	89.4
a		-	-	1.03	1.03	1.03	1.00
b		-	-	1.04	1.04	1.04	1.07
A_{3r}	μm	-	-	-	-	-1.5	-1.6
A_{3i}	μm	-	-	-	-	-1.5	-2.8
B_{3r}	μm	-	-	-	-	-	55.7
B_{3i}	μm	-	-	-	-	-	-34.9
ΔC_1	nm	-	-	-	-0.14	-0.14	-0.13
E_{rms}	nm	29.16	17.94	12.62	12.60	12.59	12.41

Table 4.6: 7x7c0 diffractogram fits

then 10 μm for most of the sets. The TCC matrix varies significantly for sets 7x7a-7x7c which implies that it is a function of condenser lens strength and condenser aperture size.

All the fitted values for A_2 from table 4.7 are plotted in figure 4.32 and those for τ_c in 4.33. Like the four-tilt estimates, both parameters have a reasonably low variance. There is a definite anti-clockwise rotation of about 15° in the A_2 estimates from sets 7x70-7x7c2 compared with sets 7x7a-c. There is also an almost negligible change in mean magnitude over the period (from 1.77 μm to 1.87 μm). As there was an eight month period between these two groups of sets it is likely that A_2 has actually changed its direction and possibly its magnitude over this period.

Coeff.	Units	7x7a	7x7b	7x7c	7x70	7x7c0	7x7c1	7x7c2
n		39	38	41	51	51	24	20
A_{1r}	nm	-6.0	-2.0	-4.4	-0.9	-0.9	-7.1	-7.3
A_{1i}	nm	-9.2	-1.1	-3.1	-1.1	-2.0	-5.6	-0.7
C_1	nm	13.2	1.9	15.8	-1.8	8.2	-25.6	-73.9
A_{2r}	μm	1.08	1.30	1.41	1.63	1.52	1.71	1.53
A_{2i}	μm	-1.31	-1.18	-1.26	-0.86	-0.94	-0.98	-1.07
τ_{cx}	mrad	-0.42	-0.42	-0.34	-0.61	-0.57	-0.58	-0.40
τ_{cy}	mrad	-0.03	0.00	-0.19	-0.35	-0.26	-0.36	-0.36
θ_1	$^\circ$	6.1	0.8	4.0	4.7	3.8	3.8	3.8
θ_2	$^\circ$	84.7	88.8	89.1	90.3	90.3	90.6	90.7
a		0.95	1.00	1.01	1.02	1.03	1.00	1.01
b		0.90	1.01	1.01	1.03	1.04	1.03	1.03
ΔC_1	nm	0.38	-0.05	-0.06	0.17	-0.14	-0.12	-0.22
E_{rms}	nm	19.7	7.3	8.5	10.1	12.6	9.4	7.5

Table 4.7: Fit IV type diffractogram fits

Coeff.	Units	7x7a	7x7b	7x7c	7x70	7x7c0	7x7c1	7x7c2
A_{3r}	μm	19.4	-5.1	6.2	0.6	-1.5	3.7	-0.5
A_{3i}	μm	-26.1	5.8	-16.1	-6.8	-1.5	-3.8	-4.8

Table 4.8: Fit V type diffractogram fits of A_3

Coeff.	Units	7x7a	7x7b	7x7c	7x70	7x7c0	7x7c1	7x7c2
A_{3r}	μm	15.3	-5.0	5.5	1.2	-1.6	0.2	4.6
A_{3i}	μm	-25.2	4.7	-15.3	-6.6	-2.8	7.1	0.1
B_{3r}	μm	-26.3	-15.2	-42.2	37.9	55.7	-172.6	-206.5
B_{3i}	μm	-74.5	-20.6	-18.8	33.5	-34.9	-180.1	-53.2

Table 4.9: Fit VI type diffractogram fits of A_3 and B_3

4.8 Comparison of displacement and diffractogram methods

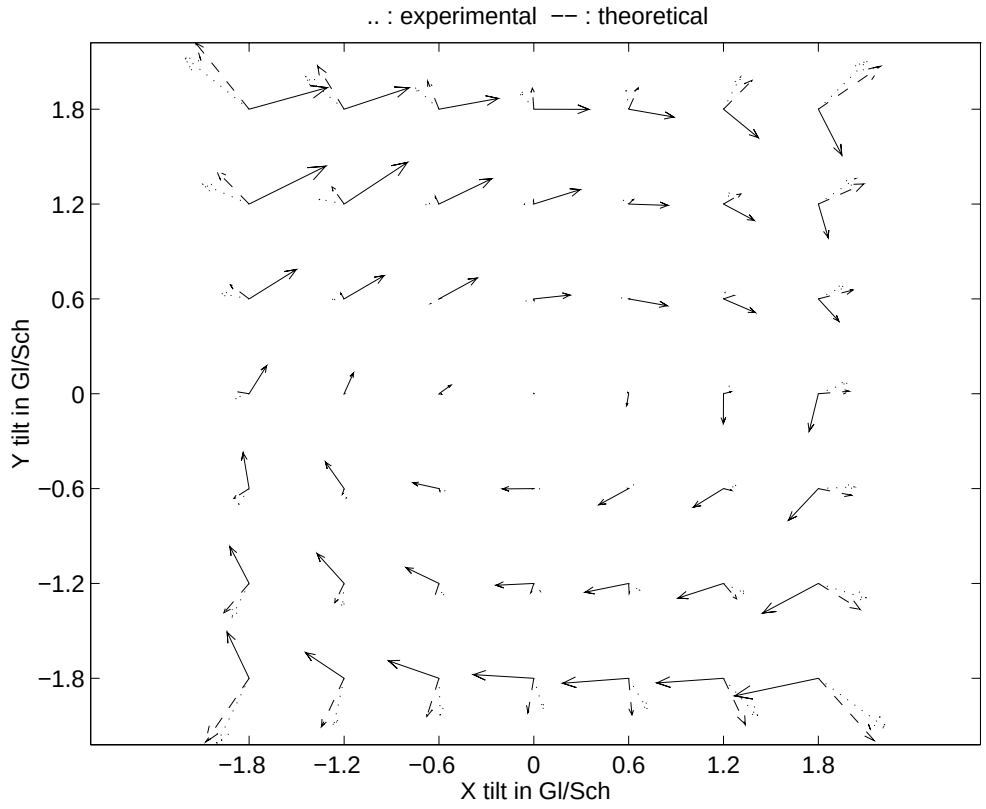


Figure 4.34: Theoretical and experimental displacements for set 7x7c1. The dotted vectors are the experimental displacements returned from the ST-XCF and the dashed vectors are the theoretical displacements generated from the diffractogram fit. The solid vectors (with a different scale) are the difference between the two from which a clockwise rotation is clearly seen.

It is now possible to compare the measured displacements for set 7x7c1 with the displacements one would expect based on the diffractogram fit of this set. Figure 4.34 shows these displacements as vectors: the dotted vectors are the experimental displacements returned from the ST-XCF and the dashed vectors are the theoretical displacements generated from the diffractogram

fit. The solid vectors (with a different scale) are the differences between the two from which a clockwise rotation is clearly seen.

It was mentioned in section 4.6 that inadequate non-linear drift compensation is a possible cause of the effective rotation. Examination of the figure reveals that the rotation does not suffer from a raster related pattern of fluctuations which one might expect if this was indeed the cause.

The misalignment of the framestore with respect to the tilt coils is effectively modeled by the TCC matrix. As the TCC matrices returned by the displacement and diffractogram LS fits are different, the implication is that image shift rather than beam tilt is the cause of the discrepancy. The physical cause of this shift is still under investigation.

Until the cause of the rotations in the displacement measurements is found, only diffractogram methods are available for accurate aberration determination. Even though diffractograms are currently fitted manually, as few as six can be fitted (with the configuration of sets tilts1 to tilts3 : see Appendix A) to elucidate enough aberrations to align a microscope.

4.9 Three-fold astigmatism

The least squares diffractogram fit and displacement fits arrive at very similar values for three-fold astigmatism: about 2 μm at an azimuth of $-\frac{\pi}{5}$. Possible sources of this aberration are the deviation from circularity in the objective lens polepiece bores as a result of the manufacturing process, heterogeneity in the polepiece material and the top and bottom polepieces not being coaxial (Kirkland, A.I. and Saxton, W.O., Private Communication) so that it is likely to be a constant of the microscope. The aberration has been almost totally ignored until quite recently when Krivanek and Leber (Krivanek, O.L. and Leber, M.L., 1993) reported a measurement for a 100 kV showing it to be significant for resolutions around 1.5 Å. As the phase shift is proportional to $\mathbf{k}^3$, its effect grows rapidly with increasing resolution while being negligible for lower resolution. As far back as 1978, Zemlin et al. (Zemlin, F. et al., 1978) derived a figure of around 0.3 μm for this aberration which they concluded has a negligible effect at resolutions of 2-3 Å.

The effect of three-fold astigmatism can be minimised by choice of specimen orientation. To illustrate this consider simulations of Si [110]. The top left image in figure 4.35 is one of 7.68 Å thick silicon expected from the

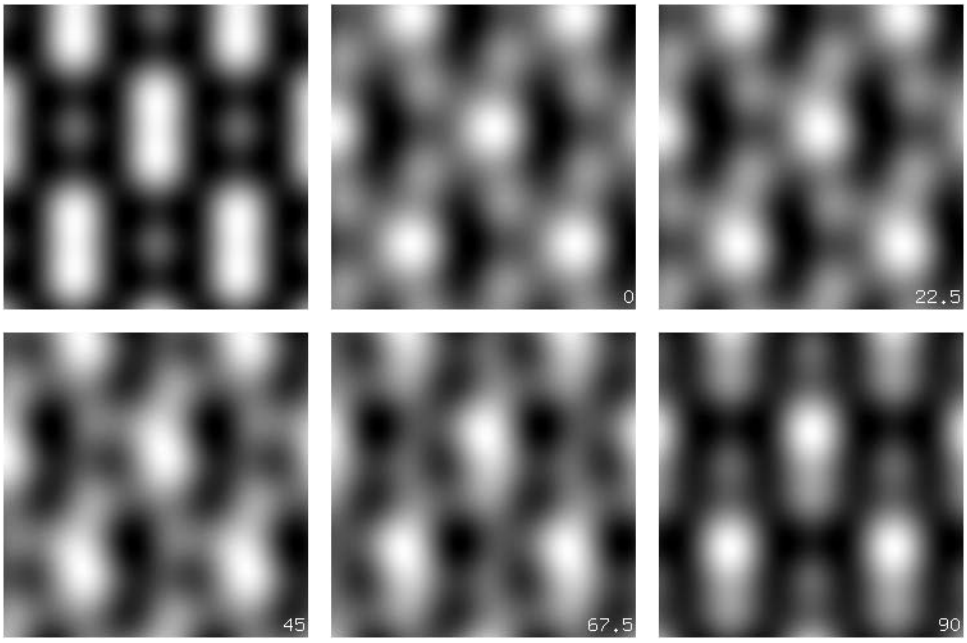


Figure 4.35: Effect of A_2 on simulation of Si [110] with thickness 7.68 Å at a defocus of -75nm. The top left image has no three-fold astigmatism whereas the rest of the images have three-fold astigmatism with magnitude 1.9 μm at the indicated azimuths to the horizontal.

JEOL 4000 EXII microscope at a defocus of -75 nm (about 2 Sch) with no three-fold astigmatism. The beam divergence used is 0.4 mrad and the r.m.s. focus spread is 10 nm. The next image across models the effect of A_2 with typical magnitude 1.9 μm and an azimuth of 0 to the horizontal. The other images in the figure model the effect of A_2 at different azimuths as indicated in their bottom right hand corner in degrees. The distortion caused at an azimuth of 0° almost disappears at an azimuth of 90°. It is therefore conceivable that microscope operators have in the past rejected areas of the specimen adversely affected by the A_2 because it would have been possible to find an area which was at a favourable orientation with respect to A_2. It may also be possible partly to compensate for the effects of A_2 by a suitable combination of A_1 and beam tilt.

The most striking evidence for the existence of A_2 comes from the ex-

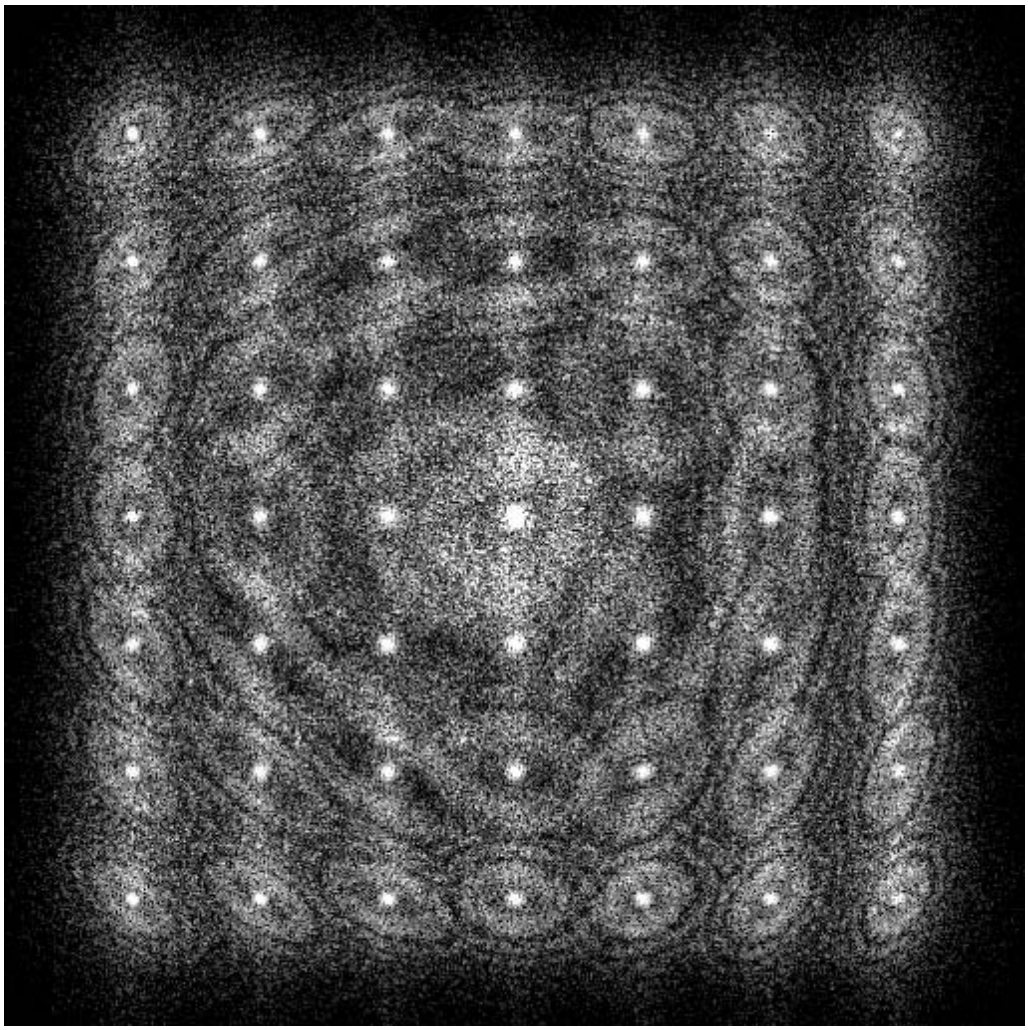


Figure 4.36: Diffractogram tableau of set 7x7c1 clearly showing the presence of *three-fold* astigmatism.

amination of diffractogram tableaus. Figure 4.36 shows the tableau for set 7x7c1. A three-fold symmetry is apparent in the orientation of the medium tilt diffractograms.

This can be explained by deriving the diffractogram orientation as a function of azimuth. From equation (3.20), if $A_1 = 0$, A_1' can be written as:

$$A_1' = 2A_2\tau^* + C_3\tau^{*2} \tag{4.1}$$

If $\tau = te^{i\phi}$, equation (4.1) becomes:

$$A_{1_0}' = 2A_2te^{-i\phi} + C_3t^2e^{2i\phi} \tag{4.2}$$

At an azimuth of $\frac{2\pi}{3}$ from this, $A_1' = A_{1_+}'$:

$$A_{1_+}' = 2A_2te^{-i\phi}e^{-\frac{2\pi}{3}i} + C_3t^2e^{2i\phi}e^{\frac{4\pi}{3}i} \tag{4.3}$$

As $\frac{-2\pi}{3} = \frac{4\pi}{3} - 2\pi$,

$$A_{1_+}' = (2A_2te^{-i\phi} + C_3t^2e^{2i\phi})e^{\frac{4\pi}{3}i} = A_{1_0}'e^{\frac{4\pi}{3}i} \tag{4.4}$$

In other words, A_1' at an azimuth of $\frac{2\pi}{3}$ is rotated from A_1' at an azimuth of 0 by $\frac{4\pi}{3}$. The orientation of a diffractogram is half that of A_1'. Therefore the orientation of a diffractogram at an azimuth of $\frac{2\pi}{3}$ is also $\frac{2\pi}{3}$. This is sufficient for three-fold symmetry to manifest itself in diffractogram orientations. It is clear that if A_2 is ignored, the orientation of the diffractograms will be the same as their tilt azimuths which is what one expects. It is worth noting that the mirror axis for the three-fold symmetry in the diffractogram tableau is close to the expected value of $-\frac{\pi}{5} \div 3 \approx 10^\circ$ where $-\frac{\pi}{5}$ is the typical azimuth for A_2 returned by least squares fits.

4.10 X-Y anisotropy

So far both the displacement and diffractogram fits have assumed that the framestore pixel ratio is 1:1. The ratio can vary by as much as 5% so that

investigation of how this anisotropy affects the fits is necessary. Table 4.10 shows how the factor f described in section 3.6.6 alters the aberration values returned by the diffractogram fit. A_1 and C_1 are the aberrations most sensitive to f. It should be possible to run an iterative procedure whereby f is varied for both the diffractogram and displacement fit of the same set until the difference in the values of the fitted aberrations is minimised. This will only be possible when the displacement method can be made to work reliably.

Coeff.	Units	0.95	1.00	1.05
A_{1r}	nm	5.8	-0.9	-8.1
A_{1i}	nm	-1.9	-2.0	-2.0
C_1	nm	6.2	8.2	10.3
A_{2r}	μm	1.49	1.52	1.55
A_{2i}	μm	-0.91	-0.94	-0.97
τ_{cx}	mrad	-0.58	-0.57	-0.56
τ_{cy}	mrad	-0.25	-0.26	-0.26
θ_1	$^\circ$	3.6	3.8	4.1
θ_2	$^\circ$	90.1	90.3	90.4
a		1.02	1.03	1.03
b		0.99	1.04	1.09
ΔC_1	nm	-0.07	-0.14	-0.21
E_{rms}	nm	12.8	12.6	13.1

Table 4.10: Effect of x-y anisotropy on 7x7c0 fit IV

Chapter 5

Image Reconstruction

5.1 Image Resolution

In an electron microscope, the electron wave suffers phase shifts due, predominantly, to the spherical aberration of the objective lens and the defocus. It is common practice to compensate spherical aberration C_3 by defocus so that for a relatively large range of spatial frequencies, the phase shift is nearly constant and equal to $\frac{\pi}{2}$. Three such defoci which are used in practice are: Scherzer defocus ($C_1 = 1$ Sch), extended Scherzer defocus ($C_1 = \sqrt{1.5}$ Sch) and extended second broadband ($C_1 = \sqrt{3.5}$ Sch). The second of these is commonly used to define different measures of image resolution (Slayter, E.M. and Slayter, H.S., 1992; Smith, D.J., 1989):

- **The resolution limit or interpretable resolution**: This is the position of the first zero of the contrast transfer function (CTF) which represents the widest possible bands of spatial frequencies without phase reversals under axial illumination conditions.
- **The information limit or the instrumental limit**: This is defined in terms of the attenuation produced by the envelope function which includes the effect of mechanical vibration. A cut-off value of e^{-2} or 15% is commonly used.

Spatial frequencies to the left of the resolution limit can be observed directly, whereas frequencies between the resolution limit and the information limit can only be retrieved by image processing techniques.

5.2 Digital Image Reconstruction

The interpretable resolution can be increased by decreasing the spherical aberration coefficient of the objective lens and/or increasing the accelerating voltage. These approaches are reaching their limits so that new approaches are necessary to improve resolution below 1 Å. The fact that large field digital images are available from microscopes which have themselves had their imaging conditions better characterised implies that digital image reconstruction holds most promise as a route to higher interpretable resolution.

Such processing improves the interpretable resolution by compensating for the CTF phase reversals. It also attempts to recover the modulus and the phase components of the specimen exit plane wavefunction. This is important because of the general 'phase problem' in electron microscopy where only image intensities are recorded. Reconstruction can only take place where the effects of the coherence envelope allow so that there is a move towards the use of field emission guns to increase the extent of this envelope.

Three important reconstruction techniques are currently in practice: holography (Lichte, H., 1991), focal series reconstruction (van Dyck, D. and Op de Beeck, M., 1990) and tilt series reconstruction (Saxton, W.O., 1994b).

5.2.1 Holography

Although more than 20 forms of electron holography can be defined, the most fruitful is referred to as off-axis holography (Lichte, H., 1991; McCartney, M.R., 1994). In this, the wave passing through the specimen is made to interfere with a reference wave from the vacuum by means of an electrostatic biprism. The resulting interference pattern allows the recovery of both amplitude and phase information. The effect of inelastic scattering and the uninterpretable non-linear image intensity components can be eliminated in holography but the requirement of carrier frequency fringes two to three times finer than the reconstruction resolution puts high demands on microscope stability and beam coherence. As a consequence field emission gun electron microscopes are pre-requisites for holography.

5.2.2 Focal series reconstruction

Focal series reconstruction attempts to improve interpretable image resolution through the processing of a number of images captured at different defoci under axial illumination conditions. The resolution limit in such restorations is set by the extent of coherence envelope factors which determine the extent of the wavefunction transfer function as shown in equation (3.7). These are mildly focus sensitive as they extend slightly further at higher underfocus values than at Scherzer focus.

Early reconstruction schemes performed complex focal series reconstructions (Schiske, P., 1968; Schiske, P., 1973; Frank, J., 1972) but neglected the effects of coherence envelopes. Saxton et al. (Saxton, W.O. et al., 1977; Saxton, W.O., 1986b; Saxton, W.O., 1986a) and Pitt (Pitt, A.J., 1980) provide more recent examples of focal series reconstructions. As contrast transfer theory was improved to account for the effects of beam divergence and focal spread, appropriately revised reconstruction formulae were developed (Saxton, W.O., 1978; Saxton, W.O., 1988).

Recently, two approaches to focal series reconstruction have been promoted : the 3D paraboloid method and the maximum likelihood method. The paraboloid method (van Dyck, D. and Op de Beeck, M., 1990; van Dyck, D. and Op de Beeck, M., 1993) involves the quasi-continuous recording of the image density as a function of focus over a wide range and the analysis makes conceptual use of three-dimensional Fourier transforms. It is not really a new method other than the fact that it is applicable to a larger number of images than have been used previously (Saxton, W.O., 1994d) and is actually less effective than the formula Schiske gave for focal series reconstruction 25 years ago (Schiske, P., 1968).

The maximum likelihood method introduced by Kirkland (Kirkland, E.J., 1984). It utilises a non-linear search method with respect to a very large number of parameters and its full potential has yet to be realised although it holds great promise. Coene et al. (Coene, W.M.J et al., 1993) have developed a variant of the method relying on a field emission gun for which the simpler image formation theory greatly reduces the computation necessary. This uses a linear reconstruction as the starting point for the non-linear search.

5.2.3 Tilt series reconstruction

Although further improvements in lens design will bring further marginal improvements in resolution in the axial mode, only tilted beam modes hold out any promise for a substantial improvement. Tilt series reconstruction involves the capture of images with different injected tilts. An early example of such a reconstruction can be found in (Jenkins, W.K., 1979). In the tilted imaging mode, for close to the optimum defocus, the envelope is wider in one direction for a given illumination coherence than in the axial mode. In addition, transfer is free of zeros over a substantial part of the area. As the transfer is direction dependent, at least four images are required at different tilt azimuths to cover a reasonable portion of Fourier space. Theoretical improvements in resolution using this technique exceed those for focal series reconstruction (Kirkland, A.I. et al., 1994; Kirkland, A.I. et al., 1995) and because of this, the method will be investigated in detail.

5.3 Theory of Tilt Series Reconstruction

5.3.1 Image formation

From equation (3.6), the image contrast is related to the specimen exit plane wavefunction, $\Psi(\mathbf{k})$, as follows:

$$I(\mathbf{k}) = \Psi(\mathbf{k})T(\mathbf{k}) + \Psi^*(-\mathbf{k})T^*(-\mathbf{k}) \tag{5.1}$$

where $T(\mathbf{k})$ is the wavefunction transfer function or contrast transfer function (CTF) defined by equations (3.7), (3.8) and (3.9).

For each set of images with recorded contrast, $i_1(\mathbf{x}), i_2(\mathbf{x}), ...$, the Fourier components of the contrast are related to those of the specimen exit plane wave function, ψ, via:

$$I_n(\mathbf{k}) = \Psi(\mathbf{k})T_n(\mathbf{k}) + \Psi^*(-\mathbf{k})T_n^*(-\mathbf{k}) \tag{5.2}$$

Adopting a simpler notation one gets:

$$I_n = \Psi T_n + \Psi_-^* T_{n-}^* \tag{5.3}$$

The reconstruction problem is thus the optimum estimation of Ψ from noisy observations of the image data, I_n.

There are in effect two complex unknowns (Ψ and Ψ_-) in equation (5.3) so that two images would provide enough information to determine Ψ. However, more than two images are desirable in practice in order to eliminate singular points and introduce redundancy to combat noise. Then Ψ is over-determined and least squares fitting can be performed.

5.3.2 Least squares fit

There are two ways in which a least squares criterion can be applied:

1. Find the Ψ which minimises the sum of the squares of the residues in equation (5.3) i.e. fit the observed data.

2. Fit the object itself rather than the image data by including the effects of noise.

The result from both these fits is a weighted sum of image components, I_n (Saxton, W.O., 1988):

$$\boxed{\hat{\Psi} = \sum_n R_n I_n} \tag{5.4}$$

where R_n is called a restoring filter. For the case where the observed data is fitted,

$$R_n = \frac{\mathbb{S}_- T_n^* - \mathbb{C}^* T_{n-}}{\mathbb{S}_- \mathbb{S} - |\mathbb{C}|^2} \tag{5.5}$$

$$\mathbb{S} = \sum_n |T_n|^2 \qquad \mathbb{S}_- = \sum_n |T_{n-}|^2 \qquad \mathbb{C} = \sum_n T_n T_{n-} \tag{5.6}$$

The weakness of this approach is the restoring filter diverges, with infinite noise levels, at high and low spatial frequencies where all the transfer functions are small.

A way to avoid this is to include noise in the analysis:

$$I_n = \Psi T_n + \Psi^*_- T^*_{n-} + N_n \tag{5.7}$$

from which the restoring filter using the Wiener criterion (Saxton, W.O., 1988) is:

$$\boxed{R_n = \frac{(\mathbb{S}_- + \nu)T^*_n - \mathbb{C}^* T_{n-}}{(\mathbb{S}_- + \nu)(\mathbb{S} + \nu) - |\mathbb{C}|^2}} \tag{5.8}$$

where $\nu = \frac{N}{\Psi}$, the noise-to-signal ratio.

The effective CTF measures the transfer of Ψ all the way from object to reconstruction and is defined by:

$$\mathfrak{T} = \sum_n R_n T_n \tag{5.9}$$

so that

$$\mathfrak{T} = \frac{(\mathbb{S}_- + \nu)\mathbb{S} - |\mathbb{C}|^2}{(\mathbb{S}_- + \nu)(\mathbb{S} + \nu) - |\mathbb{C}|^2} \tag{5.10}$$

It is clear that $\mathfrak{T}$ will be close to 1 whenever $\mathbb{S}$ is large compared to ν. When this is not the case, one does not expect to be able to recover any information about the specimen wave so that $\mathfrak{T}$ is close to zero.

5.3.3 Tilt angle

There is no obvious optimum tilt series tilt angle for all conditions. Tilting the beam by an angle θ (in Gl/Sch) from the axis and setting the defocus near θ^2 (in Sch) transfers perfectly all the beams at the same angle θ to the optic axis and transfers strongly all beams within the achromatic circle defined by a tilt of θ. This is regardless of the focus spread or beam divergence (Saxton, W.O., 1988). When the beam tilt is too small, there is little improvement over focal series reconstruction; when it is too large, the transfer envelope

falls away within the circular region transferred, to an extent which depends upon the illumination coherence. Large angles are also likely to cause parallax problems if the specimen is too thick. An angle around 1 Gl/Sch is a useful compromise.

5.3.4 Practical Implementation

In the past, the major problems with reconstructions based on tilt series have been the sufficiently accurate determination of imaging conditions and the location of a common origin in the images (registration). The previous chapter has described how lens aberrations, which characterise microscope imaging conditions, can be elucidated by a diffractogram fit, solving the first problem to an extent. The second problem can be solved by using knowledge of imaging conditions to derive phase compensated XCFs. Typical image sets for tilt series reconstruction begin and end with an axial image so that possible specimen damage over the set can be seen by cross-correlating these images.

5.3.4.1 Determination of imaging conditions

Currently, the most reliable aberration elucidation procedure is the least squares fitting of diffractograms. As indicated in the previous chapter, this is still a manual procedure which does not lend itself very well to on-line microscope alignment. Because of this, tilt series reconstructions were performed on images from manually aligned microscopes. Two microscopes were used to investigate tilt series reconstruction: a JEOL 4000 EXII at Cambridge HREM and a JEOL 2010F microscope in Tokyo. For the image set from the '4000 (tilts1), the instrumental alignment process consisted of finding the voltage centre after focusing and stigmating. For the images sets from the '2010F (tilts2-4), the alignment involved finding the coma-free axis using on-line diffractograms as a visual guide (Hosokawa, H. et al., 1993) after focusing and stigmating. The '2010F has a field emission gun (FEG) which provides more rings in the diffractogram than the LaB_6 based '4000 which makes diffractograms easier to fit due to the improved coherence envelopes.

5.3.4.2 Image registration

5.3.4.2.1 Use of phase compensated XCFs

Before any reconstruction can be performed, each image set has to have a common origin. This is trivial in the case of images captured on the '4000 as only drift compensation is required because the images are captured straight into a framestore. Only photographic plates were available for the images from the '2010F so that a more complicated registration procedure had to be used.

The usual technique for finding a common origin is to take the XCF of each tilted image with the first axial image. Section 3.6.1.3 showed how unreliable this technique can be due to the effect of different transfer functions in the images. Consequently, the knowledge of imaging conditions derived from a diffractogram fit can be used to generate the PSF-XCF templates for a ST-XCF, for example.

Images captured on the achromatic circle suffer no displacement relative to axial images. In practice, there is likely to be a displacement present due to the inaccuracies in achieving the achromatic condition. This displacement is compensated for in the reconstruction process so that it is vital for it to be preserved. Fortunately, the use of ST-XCFs in image registration does indeed preserve these displacements.

5.3.4.2.2 Least squares fit of displacements

Using the first axial image as the reference for ST-XCFs leads to an accumulation of errors. It therefore makes sense to take the ST-XCF of each with all the others in the set. This provides an over-determined set of observations which can be entered into a least squares fit process to provide more accurate estimates for image registration (Saxton, W.O., 1994a).

5.4 Tilt Series Reconstruction Results

Sets tilts1-4 (see Appendix A) were used to investigate tilt series reconstruction. Set tilts1 is of the block oxide, $4Nb_2O_5.9WO_3$ (Horiuchi, S. et al., 1978) from the '4000 and sets tilts2-4 are of Gold on Germanium from the '2010F.

5.4.1 Gold on Germanium sets

Tilts2 was captured at 2 Sch underfocus with a tilt magnitude of roughly 1.44 Gl/Sch, tilts3 at 1 Sch underfocus and 1 Gl/Sch tilt and tilts4 at 0.25 Sch underfocus and 0.5 Gl/Sch tilt. Figures 5.1, 5.2 and 5.3 show the diffractograms for the images in these sets. From the position of the (111) and (002) beams in the diffractogram, the sampling interval of the images was determined to be 0.333 Å/pixel.

Coeff.	Units	tilts2	tilts3	tilts4
A_{1r}	nm	-2.0	-2.1	-2.1
A_{1i}	nm	-2.1	1.3	1.6
C_1	nm	-110.1	-57.0	-22.0
A_{2r}	μm	-0.15	0.16	-0.51
A_{2i}	μm	-0.28	-0.18	-0.29
τ_{cx}	mrad	-0.37	-0.26	-0.45
τ_{cy}	mrad	-0.36	-0.38	-0.08
θ_1	$^\circ$	7.7	4.8	7.2
θ_2	$^\circ$	97.4	97.9	92.4
a		0.97	1.05	0.98
b		0.98	1.11	0.97
ΔC_1	nm	0.28	0.14	-0.80
E_{rms}	nm	1.05	1.80	1.06

Table 5.1: Diffractogram fits of tilt sets tilts2-4

5.4.1.1 Determination of imaging conditions

The results of diffractogram fitting each set are shown in table 5.1. C_3 has been fixed at 1.1 mm. 1 Sch is 52.52 nm and 1 Gl/Sch is 6.91 mrad. Only approximate values for injected tilt were used (10, 6 and 3.5 mrad for tilts2,3 and 4 respectively) so that the TCC matrix values are not directly comparable. Both axial diffractograms were fitted in order to get a more reliable value for focal drift.

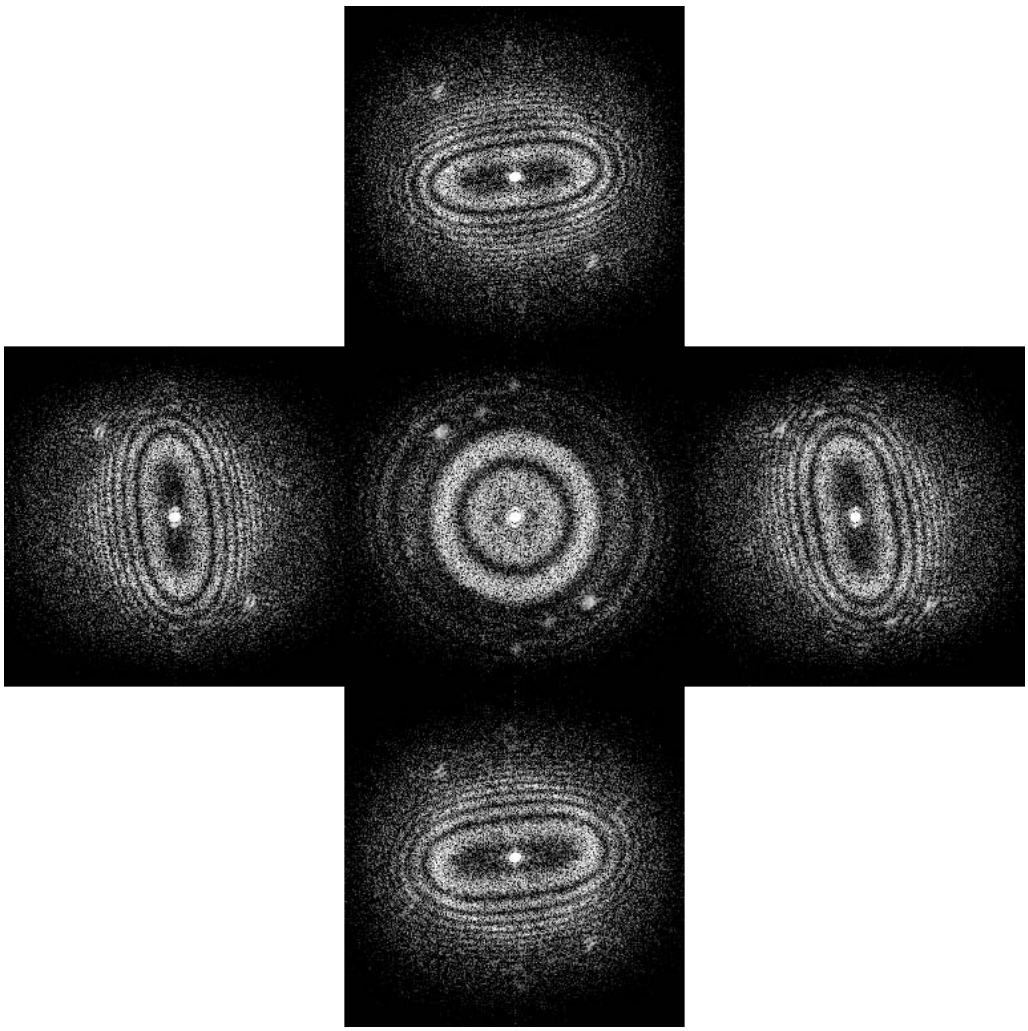


Figure 5.1: Diffractogram tableau for tilt azimuth set tilts2 whose parameters are defined in figure A.4. The edge of each field corresponds to a period of 0.23 nm.

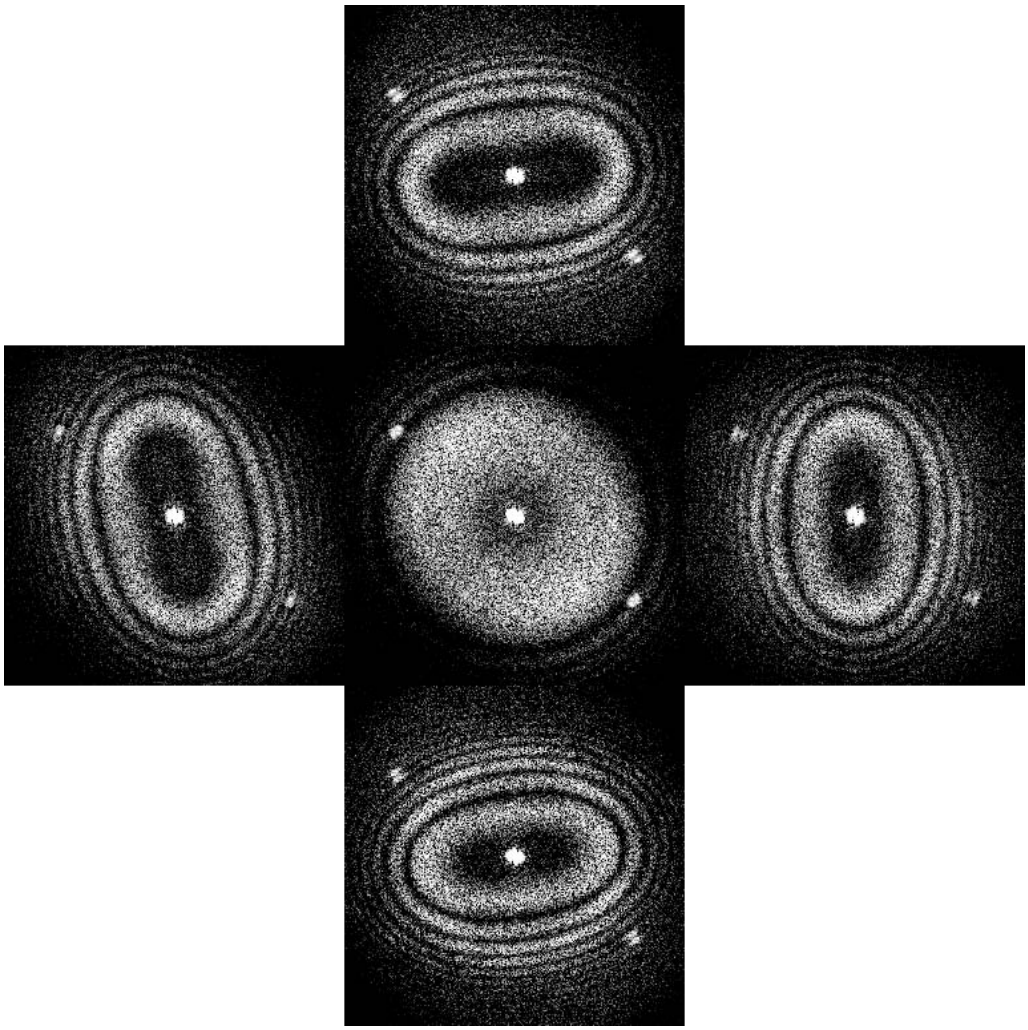


Figure 5.2: Diffractogram tableau for tilt azimuth set tilts3; the parameters of which are identical to those in figure A.4 apart from a defocus value of 1 Sch underfocus and tilt magnitudes of 1 Gl/Sch. The edge of each field corresponds to a period of 0.23 nm.

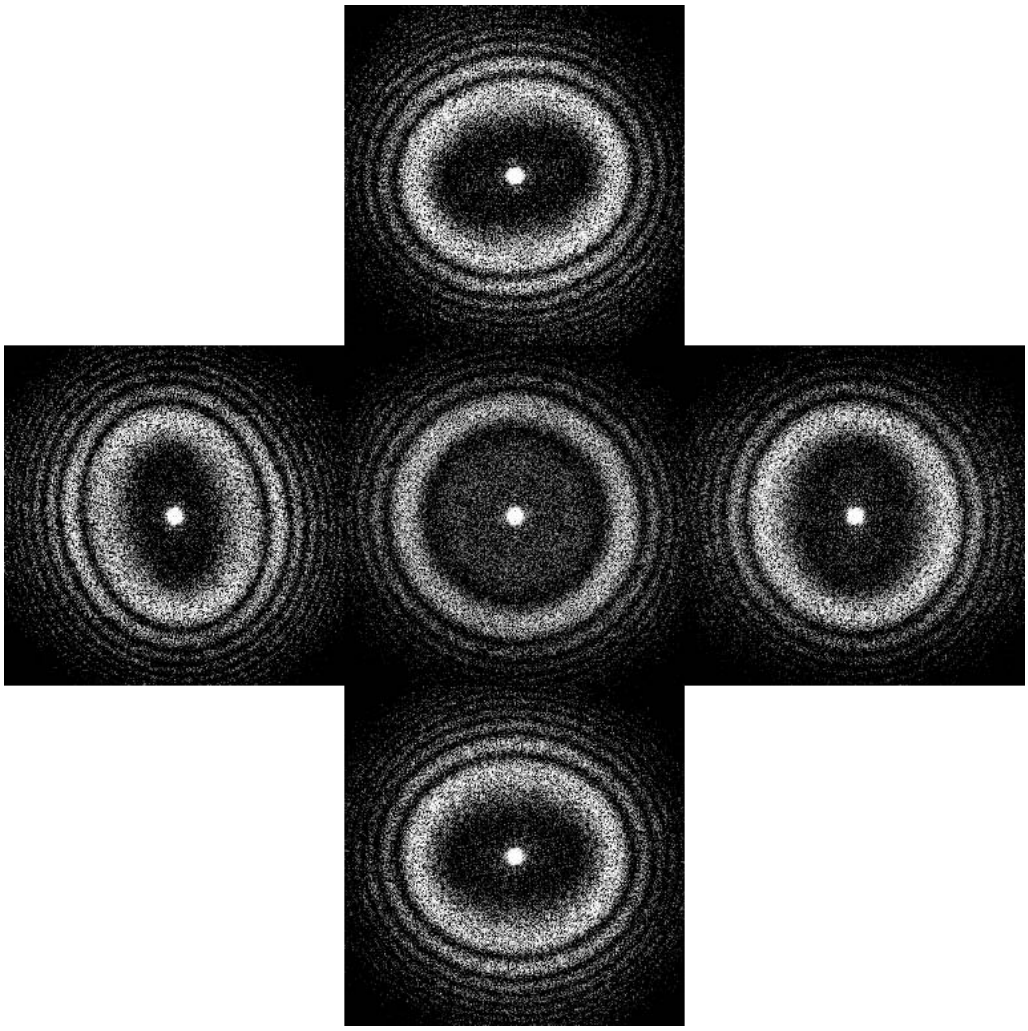


Figure 5.3: Diffractogram tableau for tilt azimuth set tilts4; the parameters of which are identical to those in figure A.4 apart from a defocus value of 0.25 Sch underfocus and tilt magnitudes of 0.5 Gl/Sch. The edge of each field corresponds to a period of 0.23 nm.

5.4.1.2 Image registration

Set tilts2 will be considered in detail because it proved to be the most fruitful of the '2010F sets as the others had unsuitably oriented particles. Each image in the set was cross-correlated with the others in the set. ST-XCFs were then generated from these XCFs using PSF-XCF templates derived from the imaging conditions defined by the aberration estimates shown in table 5.1. Figures 5.4, 5.5 and 5.6 show the XCFs, template PSF-XCFs and ST-XCFs respectively. The column number is one less than the first image in the XCF and the row number is second image in the XCF so that, for example, the image in column 2 and row 1 of figure 5.4 is the XCF of image 3 with image 1. As can be seen by the number of rings in the template PSF-XCFs, the r.m.s. focus spread and beam divergence chosen for the generation of the template PSF-XCFs are probably optimistic. However, this is unlikely to affect the ST-XCF.

The XCFs in 5.4 are phase compensated effectively as the ST-XCFs in 5.6 show. The results of least squares fitting the fifteen ST-XCF displacement estimates is shown in table 5.2. The r.m.s. error for the data is only 0.42 pixels which corresponds to 14 pm. This implies a 10 pm r.m.s. error for each image. Table 5.3 shows the values which were used in the reconstruction of set tilts2.

5.4.1.3 Reconstruction

Appendix B has a listing of the Semper program used to perform the reconstruction which is a modified version of the program listed in (Kirkland, A.I. et al., 1995). ν in equation 5.8 was set to the constant value of 0.1. Figure 5.7 shows the modulus and phase of the restored specimen exit plane wave function Ψ for tilts2. Figure 5.8 shows a reverse video diffractogram of Ψ and the effective CTF, $\mathfrak{T}$. A (113) beam is clearly visible in the diffractogram at about 20° to the horizontal and implies a resolution improvement of more than 65% (2.04 Å to 1.23 Å). The (113) lattice fringes are visible in the bottom right of the particle in the modulus of the reconstruction.

Figure 5.9 attempts to match diffractogram beams with crystal planes. The top left image is the modulus of the reconstruction. The other images represent the inverse Fourier transform of a masked Fourier transform of the modulus. For each of these images, the region outside a circle of radius 5

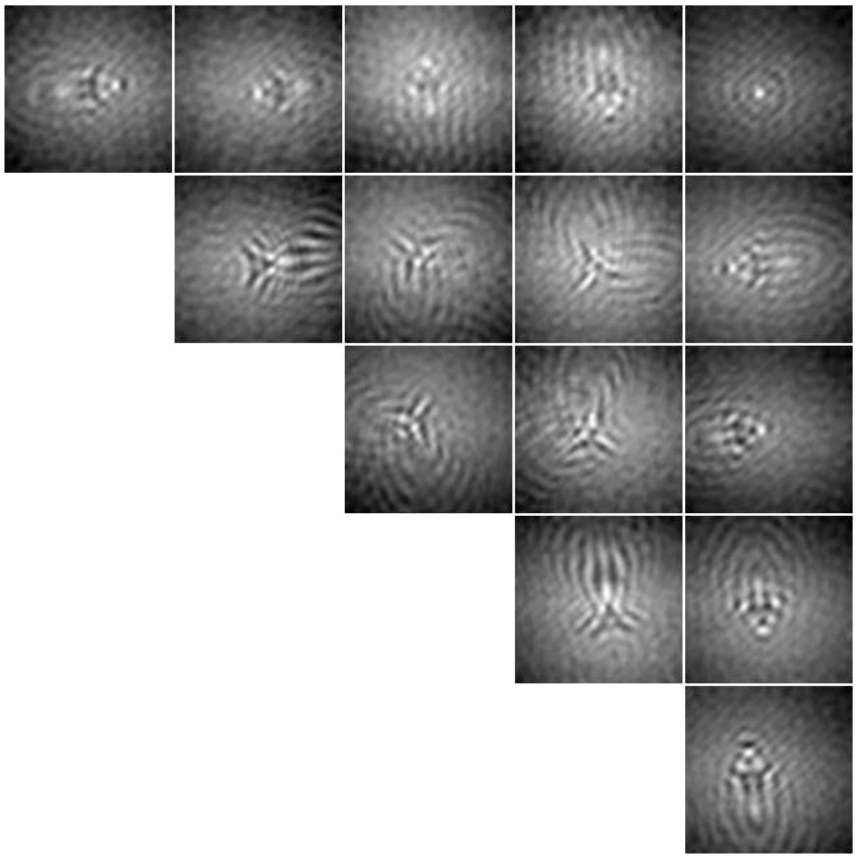


Figure 5.4: XCFs for set tilts2. The XCF in row r and column c is that between image $c + 1$ and image r.

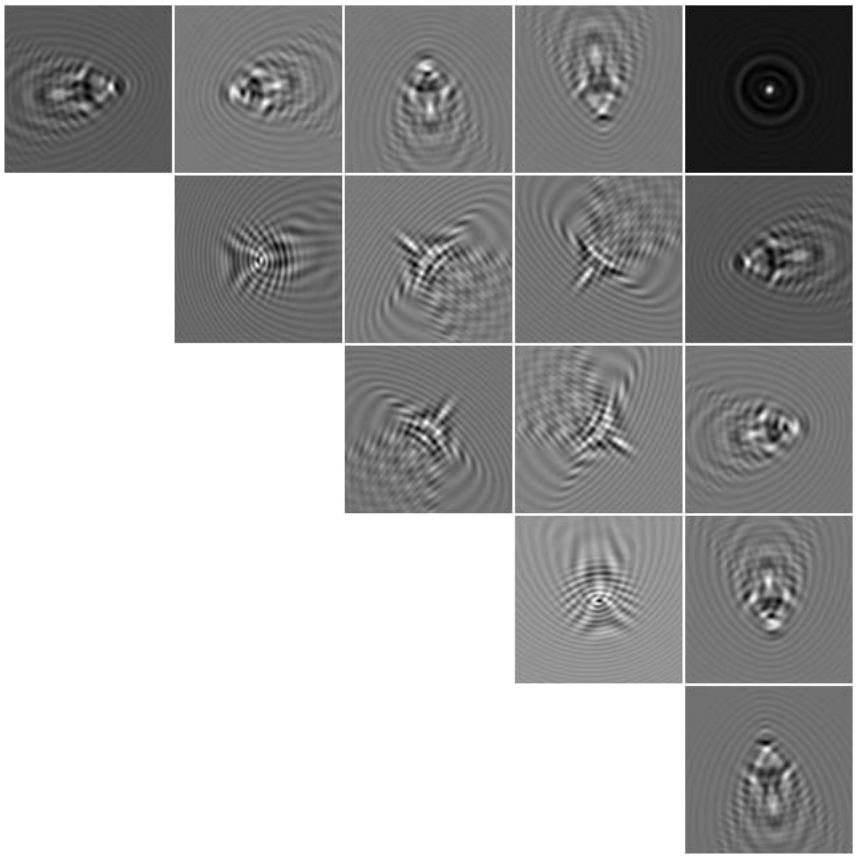


Figure 5.5: Template of PSF-XCFs for set tilts2

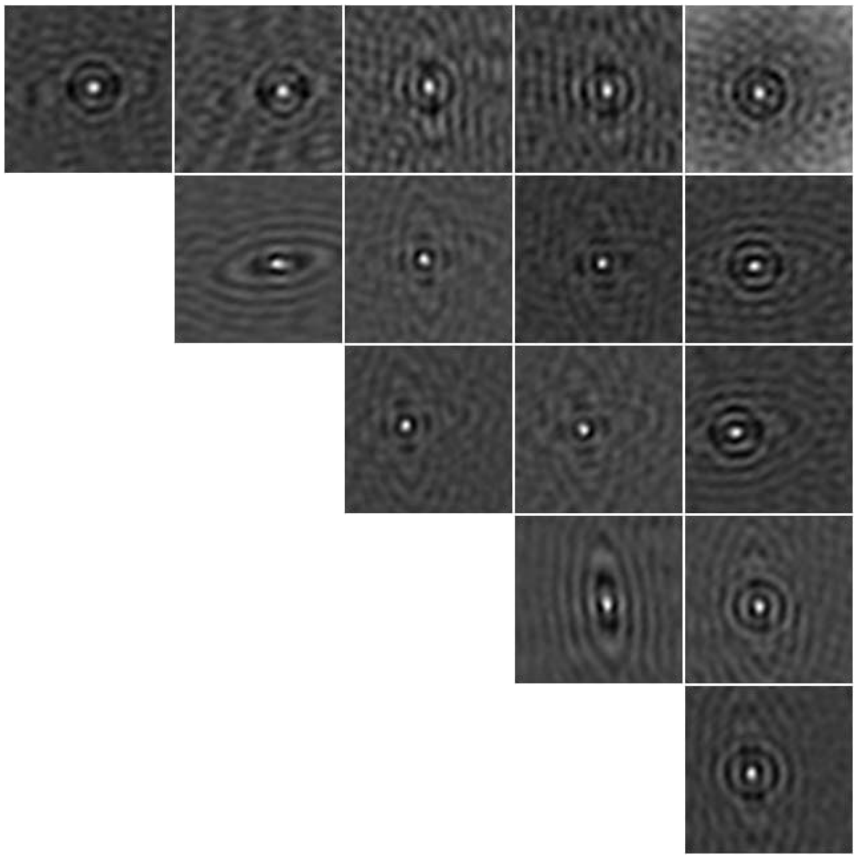


Figure 5.6: ST-XCFs for set tilts2

img. 1	img. 2	obs. x	pred. x	|misfit x|	obs. y	pred. y	|misfit y|
1	2	4	4.00	0.00	2	2.00	0.00
1	3	18	17.33	0.67	-1	-1.17	0.17
1	4	0	0.00	0.00	2	2.33	0.33
1	5	6	6.17	0.17	-1	-1.17	0.17
1	6	-8	-7.50	0.50	-3	-3.00	0.00
2	3	13	13.33	0.33	-3	-3.17	0.17
2	4	-4	-4.00	0.00	0	0.33	0.33
2	5	2	2.17	0.17	-3	-3.17	0.17
2	6	-11	-11.50	0.50	-5	-5.00	0.00
3	4	-17	-17.33	0.33	4	3.50	0.50
3	5	-11	-11.17	0.17	0	0.00	0.00
3	6	-25	-24.83	0.17	-2	-1.83	0.17
4	5	6	6.17	0.17	-4	-3.50	0.50
4	6	-7	-7.50	0.50	-5	-5.33	0.33
5	6	-14	-13.67	0.33	-2	-1.83	0.17

Table 5.2: Least squares fit of ST-XCF displacements for tilts2

	actual tilt		position of origin	
Image	x/mrad	y/mrad	x/pixels	y/pixels
1	-0.370	0.358	0.00	0.00
2	-9.946	-0.940	4.00	2.00
3	9.206	1.655	17.33	-1.17
4	0.899	-9.400	0.00	2.33
5	-1.639	10.115	6.17	-1.17
6	-0.370	0.358	-7.50	-3.00

Table 5.3: Values used for tilts2 reconstruction

pixels centred on a diffractogram beam was masked. The beam used for the top right image is (022) and the beam used for the bottom two images is (113). The (022) beam is very weak so that the crystal plane locations are

not very evident in the modulus image. Both the (113) planes show up very clearly in the modulus although the second (113) beam is only just visible in the diffractogram.

Figure 5.10 compares the modulus of the restored specimen exit plane wave function with the original axial image in order to emphasise the resolution improvement. Figure 5.11 compares two predicted images with the original images they correspond to. The images are those at tilts $(\theta, 0)$ and $(0, \theta)$. Apart from the inevitable smoothing in the predicted images, the agreement is very good indicating that the reconstruction modulus and phase are reliable. Kirkland et al. (Kirkland, A.I. et al., 1995) point out that good agreement between the predicted images and the original images is indicative of linear imaging being the predominant process.

Figures 5.12, 5.13 and 5.14 show typical images from the reconstruction of set tilts3 and 5.15, 5.16 and 5.17 show those for tilts4. It is clear from the effective CTF for tilts4, for which the tilt used was small, that there is likely to be little gain in performing the reconstruction. Also, both these sets do not seem to have beams in the tilted images which are not already present in the axial image as was the case with tilts2. An improvement in the detail of the amorphous region can be noticed despite this.

5.4.1.4 Defocus offsets

The Gold particle in tilts2 is about 50 Å in diameter and can be assumed to be a sphere resting on the amorphous film. The imaging conditions returned by the diffractogram fit are necessarily those of the amorphous film so that the defocus value for the amorphous film is likely to differ from that for the particle by as much as 50 Å . This implies that a defocus offset has to be introduced into the reconstruction which takes this height difference into account. In the case of tilts2, the offset is as low as 0.1 Sch and can be ignored in the reconstruction.

In order exploit the full potential of tilt series reconstruction, it is necessary for the method to work with crystalline objects. Clearly, imaging conditions for pure crystals cannot be derived from diffractograms so that it is important for a small amorphous region to be included in each captured image. For example, in the block oxide tilt set shown in figure A.3, the amorphous region in the top right hand corner of the images is available for diffractogram fitting. It is then assumed that the crystal is wedge shaped at

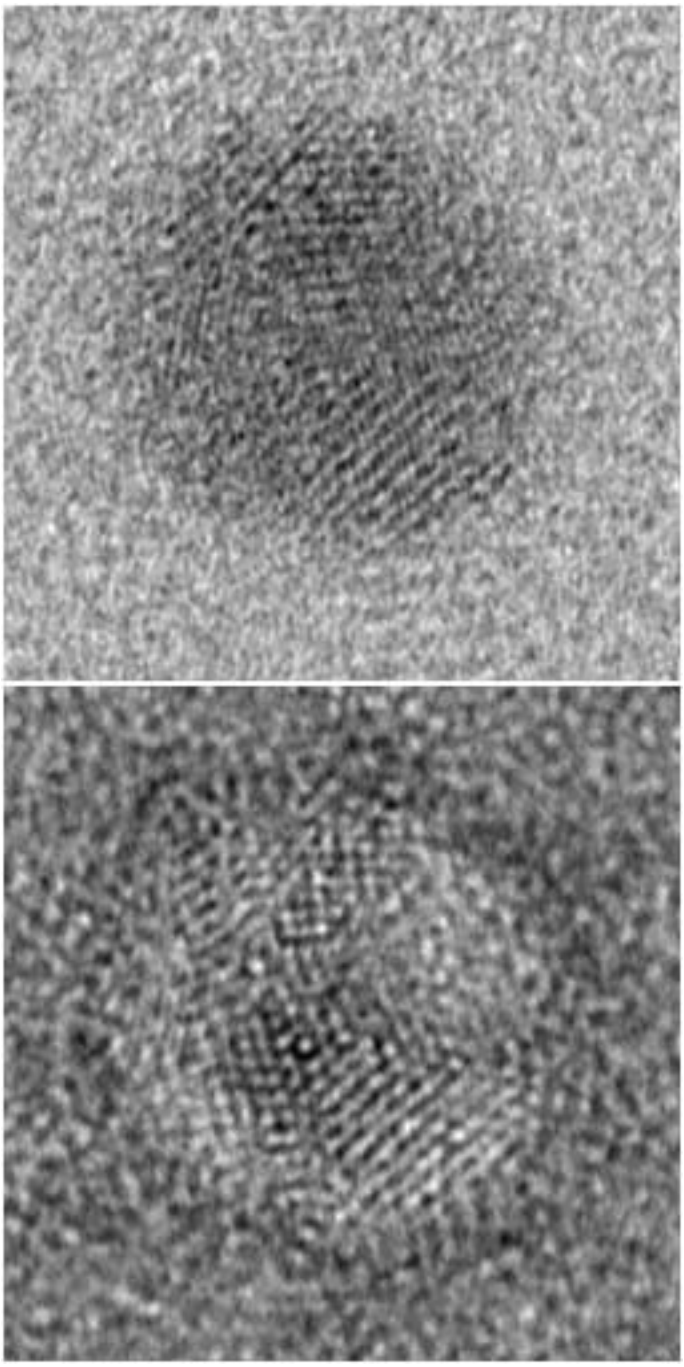


Figure 5.7: Top: modulus and bottom: phase of restored Ψ for tilt azimuth set tilts2. The field of view is 8.5 nm.

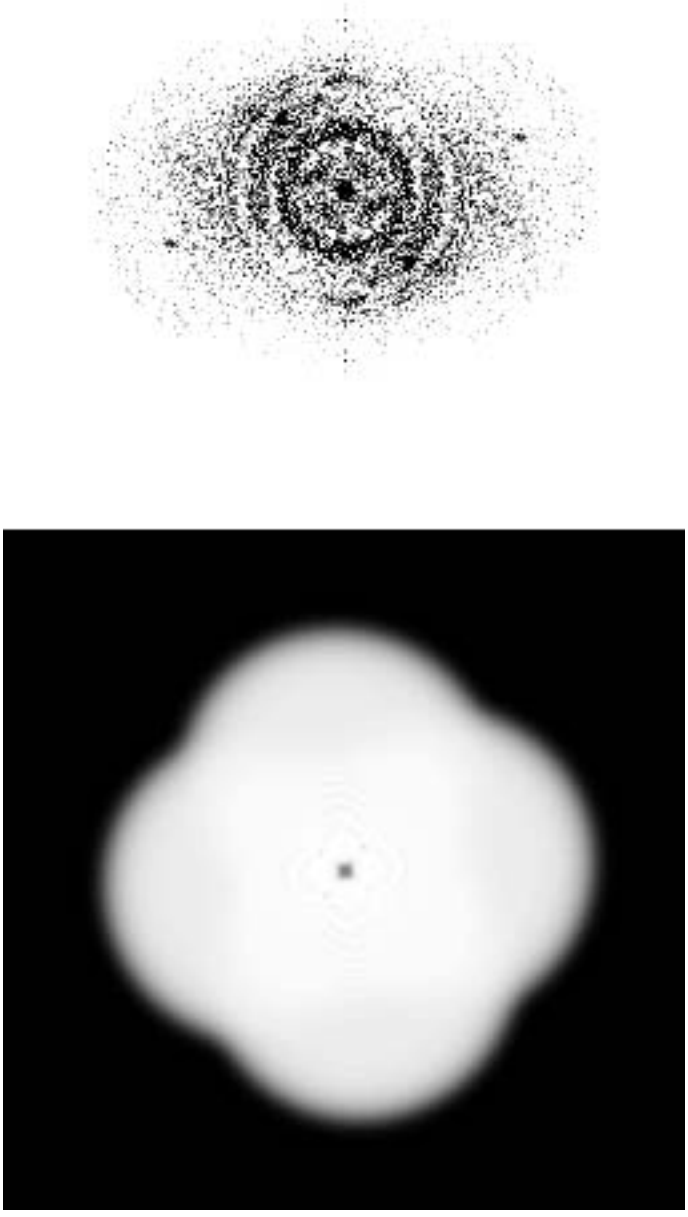


Figure 5.8: Top: diffractogram and bottom: effective CTF of restored Ψ for tilt azimuth set tilts2

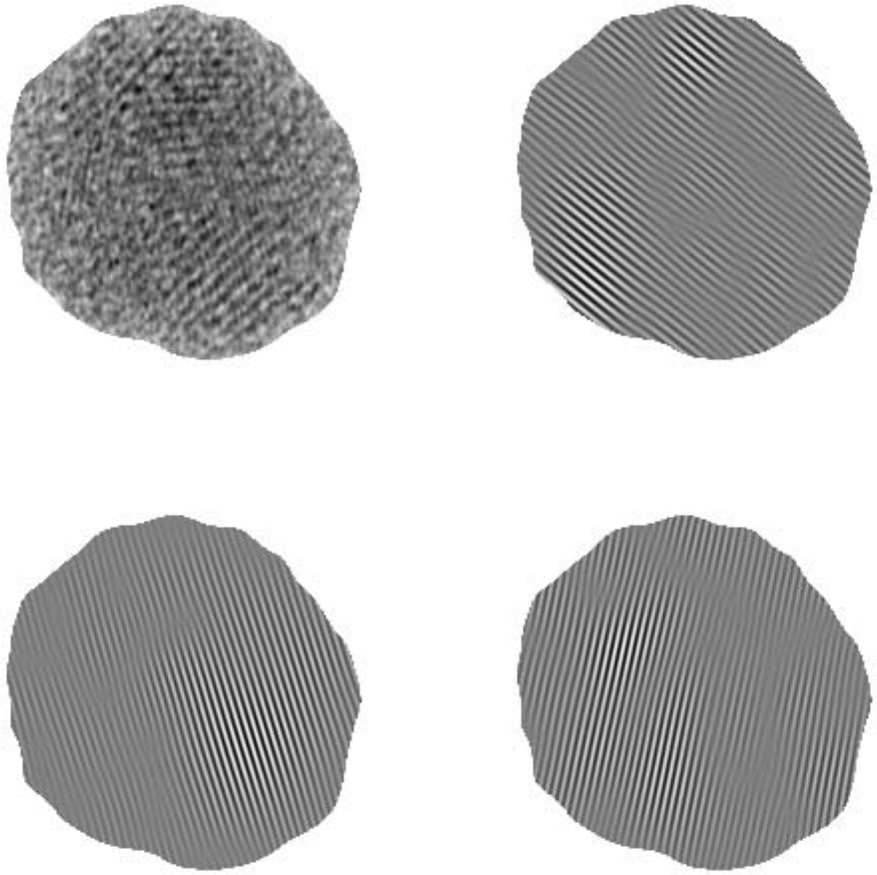


Figure 5.9: Location of crystal planes for tilt azimuth set tilts2. The top left image is the modulus of the reconstruction. The other images represent the inverse Fourier transform of a masked Fourier transform of the modulus. For each of these images, a region outside a circle of radius 5 pixels centred on a diffractogram beam was masked. The beam used for the top right image is (022) and the beam used for the bottom two images is (113). The width of the particle is 5.7 nm.

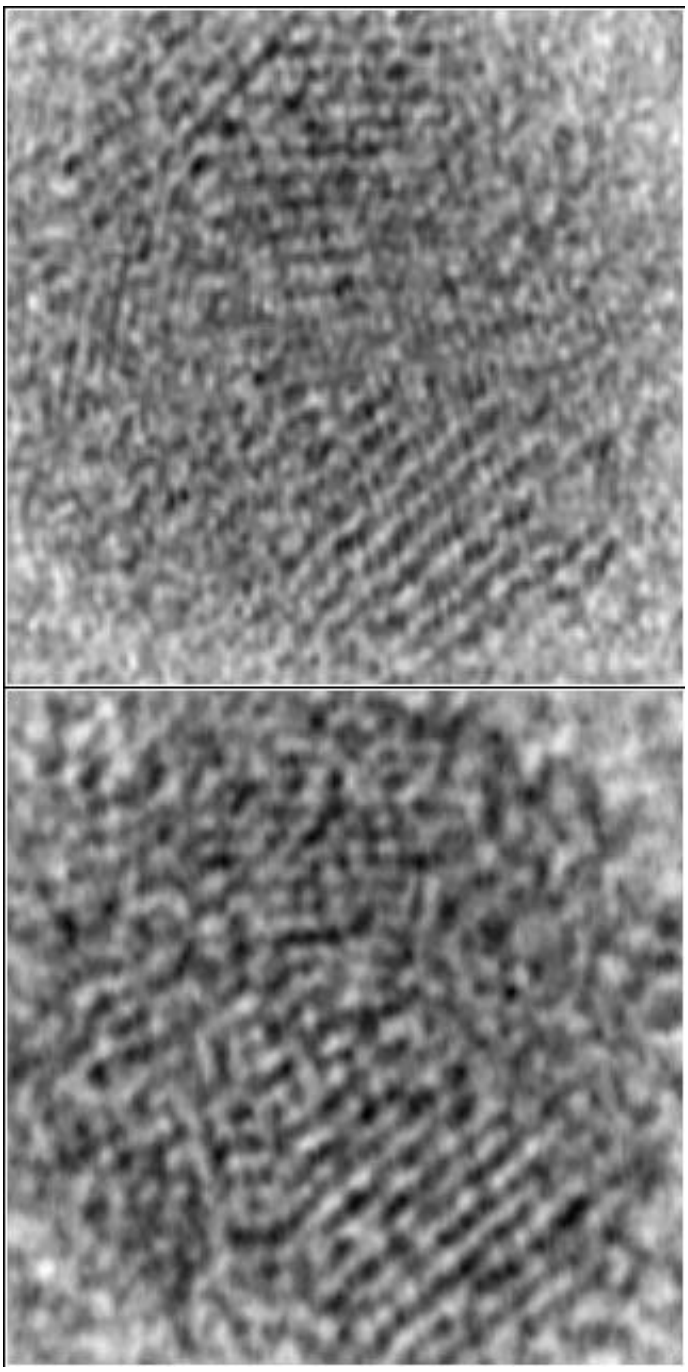


Figure 5.10: Top: modulus of Ψ and bottom: original axial image for tilt azimuth set tilts2. The field of view has been reduced to 4.25 nm. The modulus of Ψ shows signs of (113) lattice fringes (1.23 Å) in the top left and bottom right of the image. The original axial image has (002) lattice fringes (2.04 Å) present so that resolution has been dramatically improved by reconstruction.

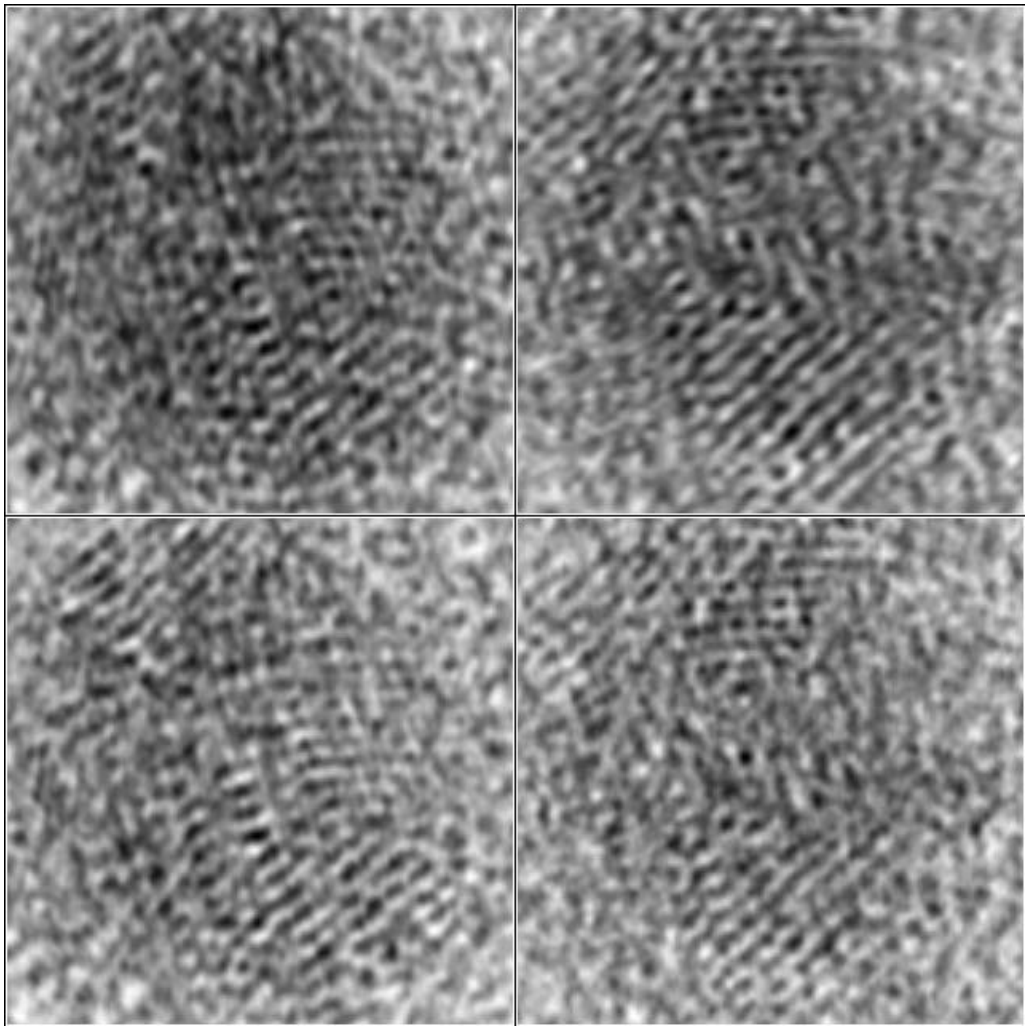


Figure 5.11: Top: predicted images and bottom: original images for two members of tilt azimuth set tilts2. The field of view has been reduced to 4.25 nm.

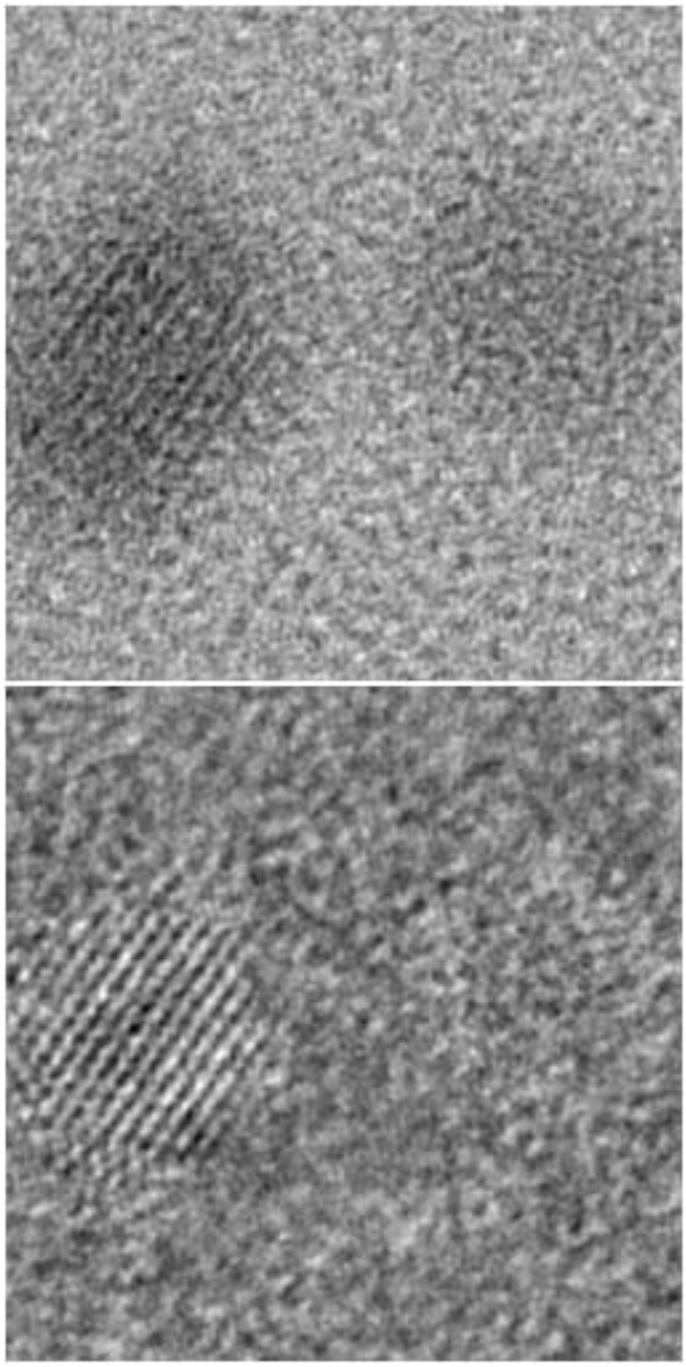


Figure 5.12: Top: modulus and bottom: phase of restored Ψ for tilt azimuth set tilts3. The field of view is 8.5 nm.

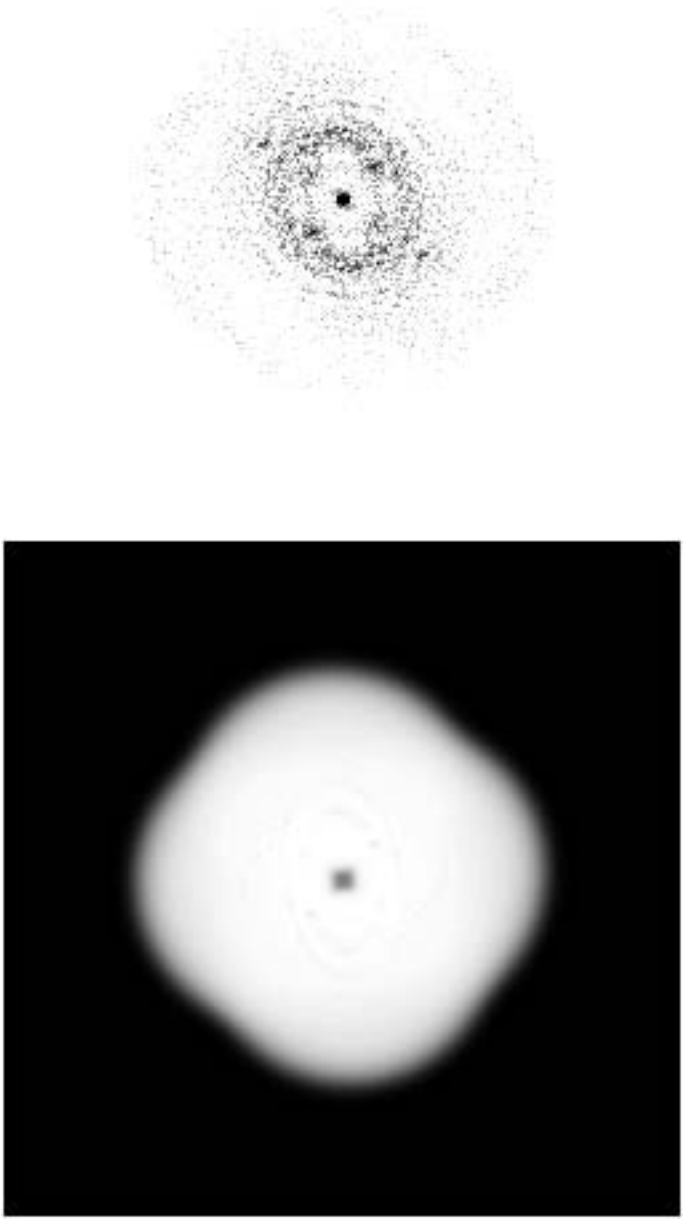


Figure 5.13: Top: diffractogram and bottom: effective CTF of restored Ψ for tilt azimuth set tilts3

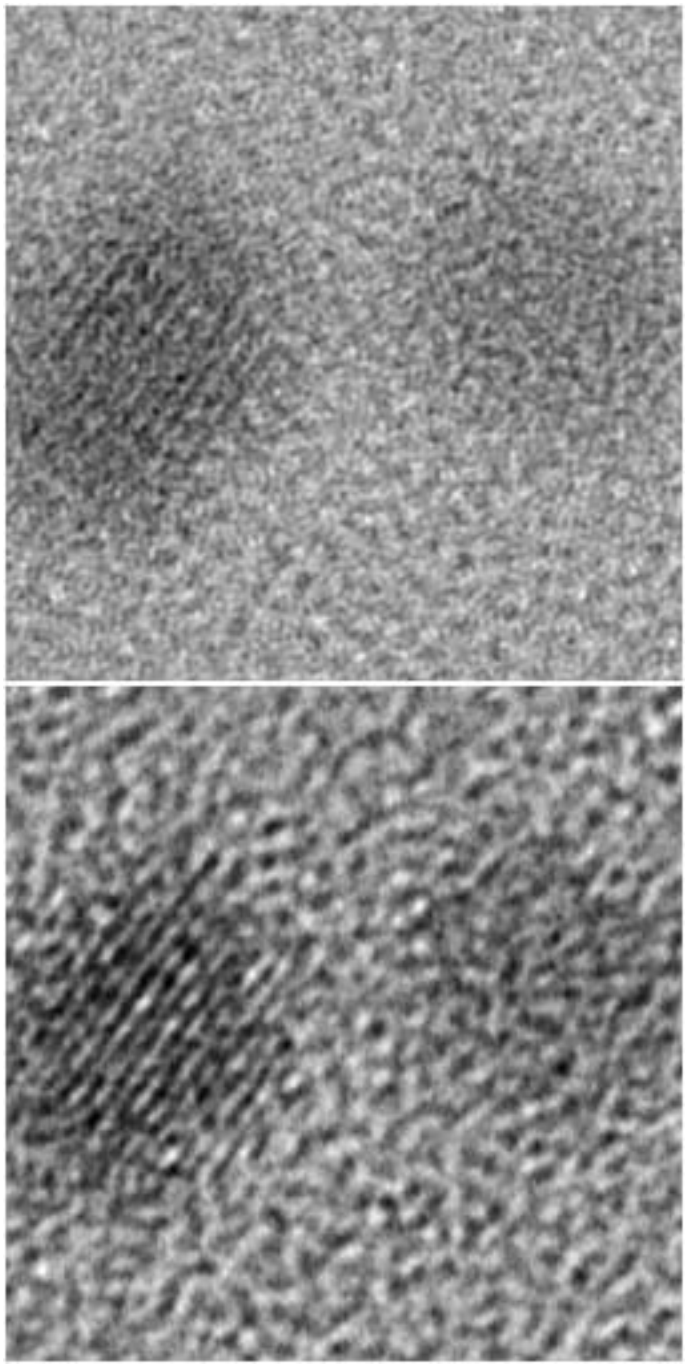


Figure 5.14: Top: modulus of Ψ and bottom: original axial image for tilt azimuth set tilts3. The field of view is 8.5 nm.

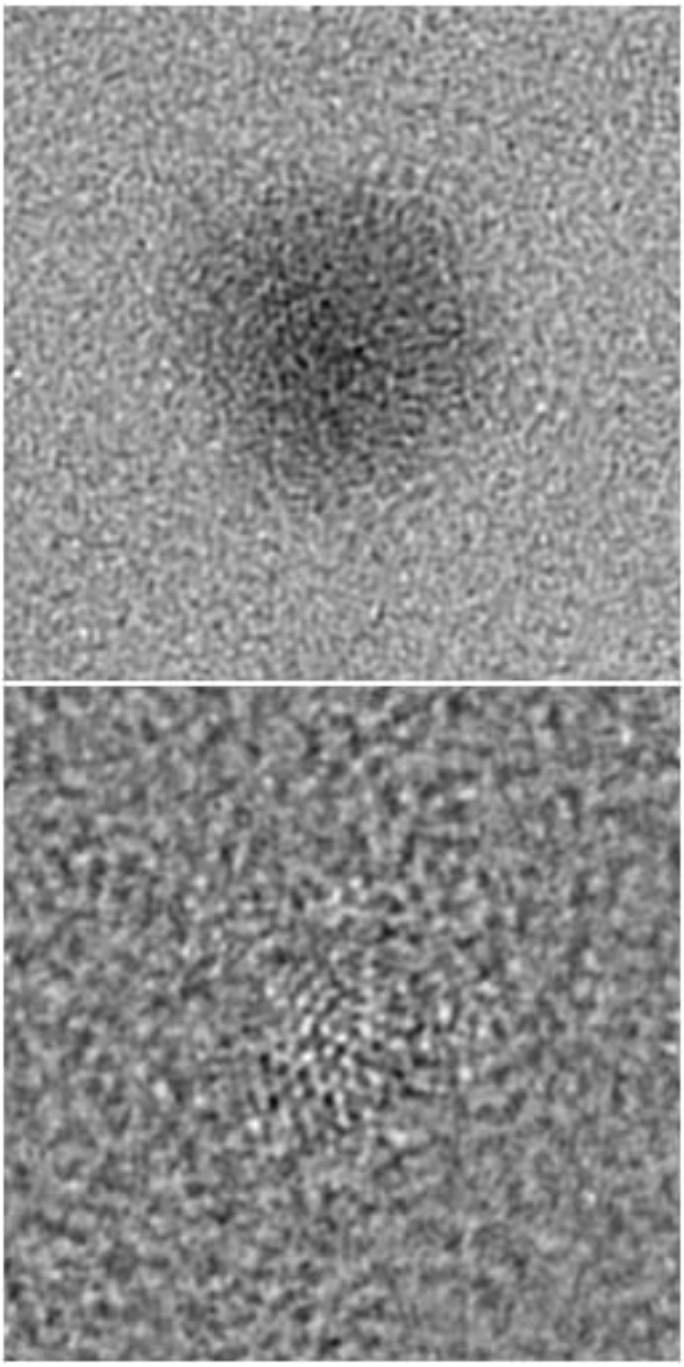


Figure 5.15: Top: modulus and bottom: phase of restored Ψ for tilt azimuth set tilts4. The field of view is 8.5 nm.

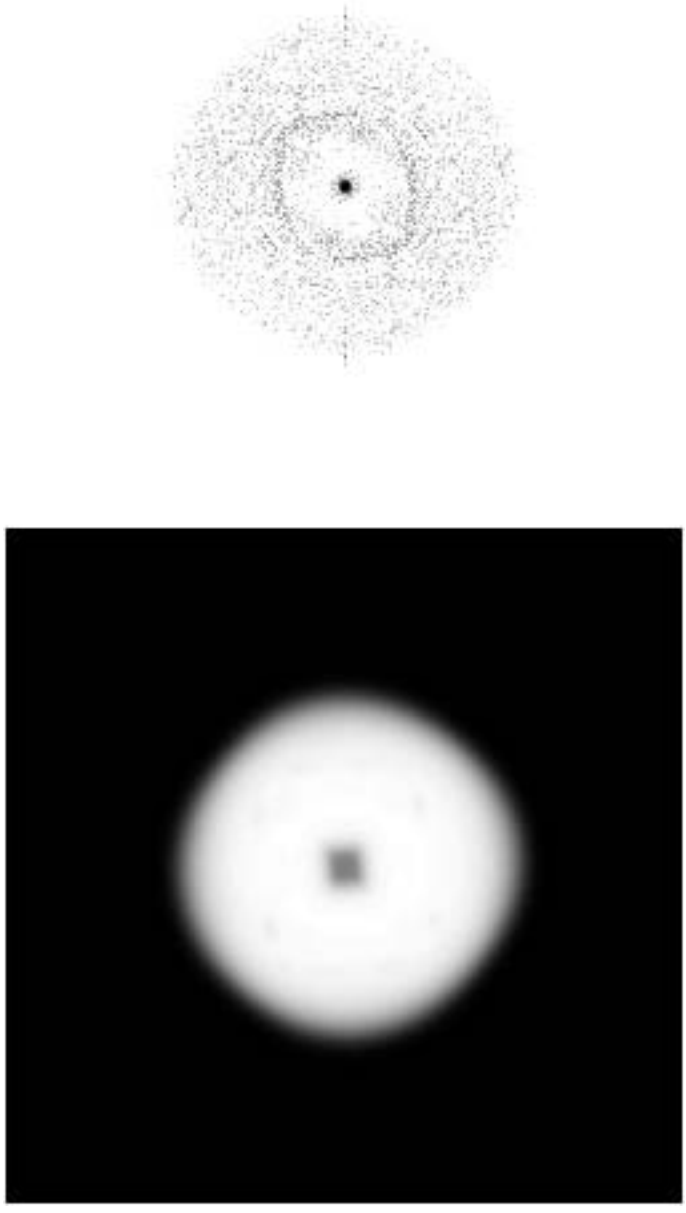


Figure 5.16: Top: diffractogram and bottom: effective CTF of restored Ψ for tilt azimuth set tilts4

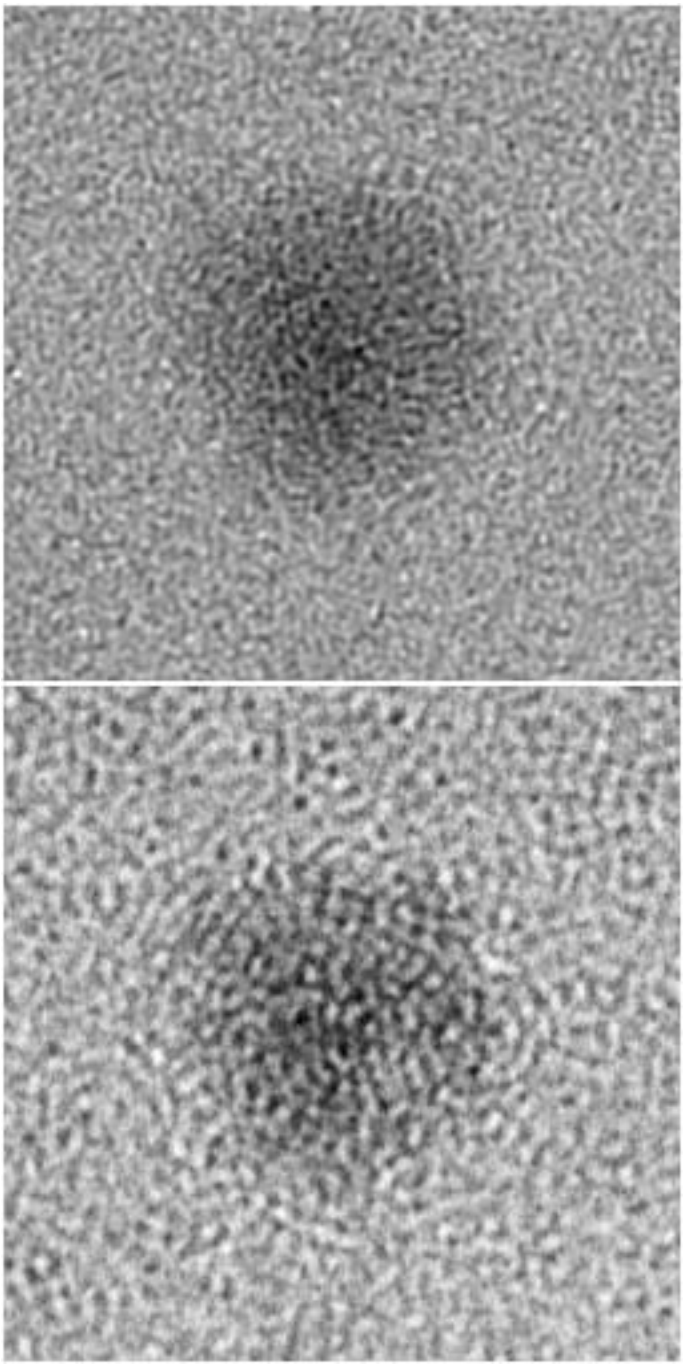


Figure 5.17: Top: modulus of Ψ and bottom: original axial image for tilt azimuth set tilts4. The field of view is 8.5 nm.

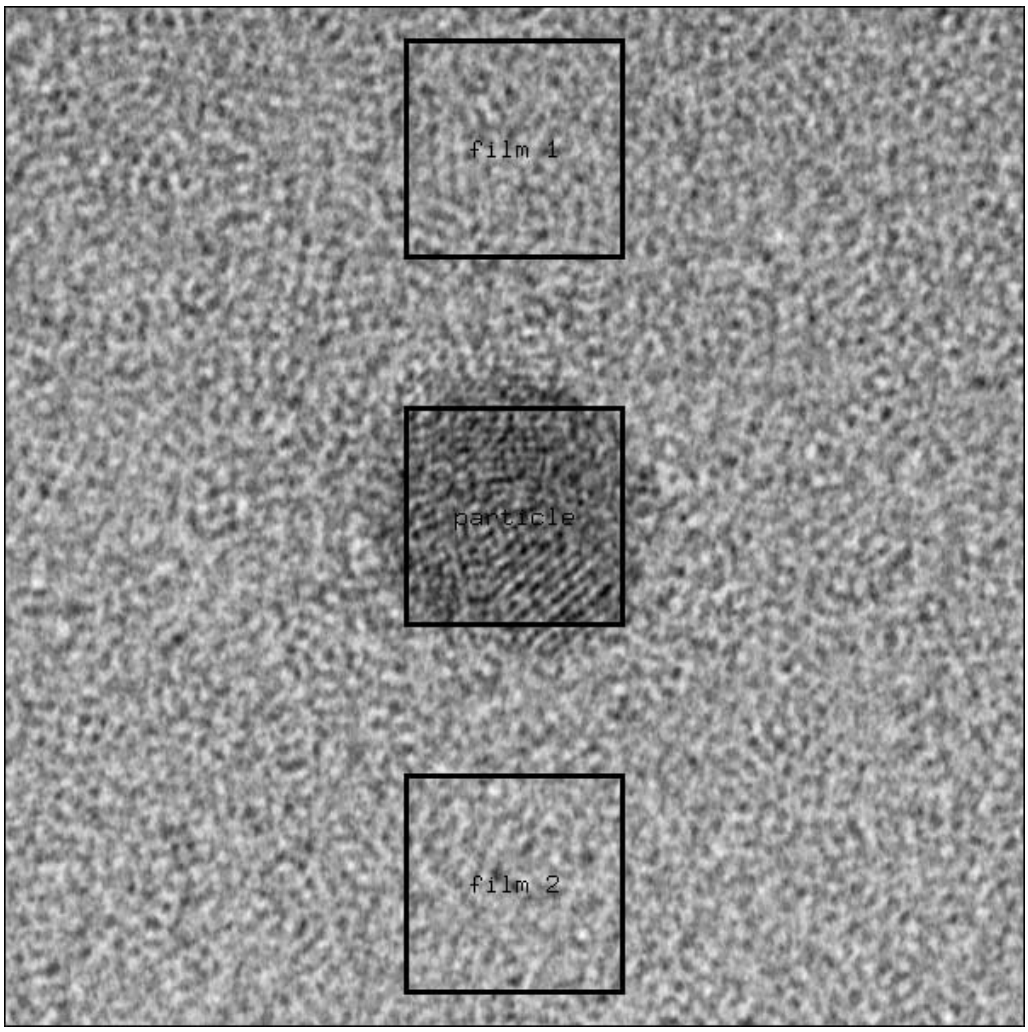


Figure 5.18: Location of height measurement areas examined for tilt azimuth set tilts2. The field of view is 17 nm.

a maximum angle of 30° so that the defocus offset for various parts of the crystal can be estimated.

It is possible to derive the defocus offset *a posteriori* by comparing the predicted images from the reconstruction and the original images in the set. If non-linear effects are ignored, there should be very good agreement between the predicted images and the original images. One measure of this

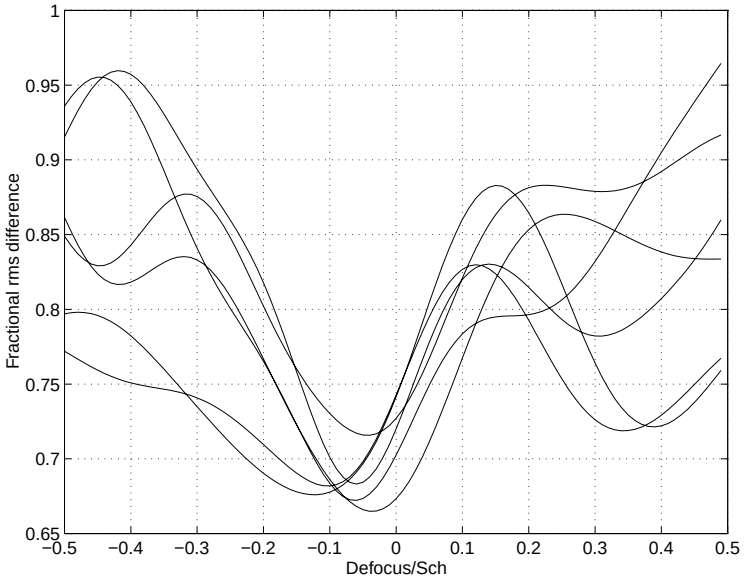


Figure 5.19: Variation of particle r.m.s. difference with defocus for members of tilt azimuth set tilts2

is the standard deviation of the difference between the normalised images, normalisation being achieved first by subtracting the mean from the image and then dividing it by its standard deviation which gives the image zero mean and unity standard deviation.

The reconstruction of different parts of the image can thus be performed at various defocus offsets and the effective r.m.s. difference between the predicted and original images can be measured. The best estimate for the defocus offset is the one which minimises the r.m.s. difference. As the defocus offset is actually the height of the part of the image under consideration relative to the support film, this technique can be used as a crude height measurement tool.

Figure 5.18 shows the three areas in the image considered for height measurement. For a particular area, the variation of r.m.s. difference with defocus offset was plotted for each image in the set. Figure 5.19 shows how the r.m.s. differences vary with defocus for all six images in the set for the area 'particle' marked in figure 5.18. The minima for each of these curves are estimates for the actual defocus offset of the particle compared with the

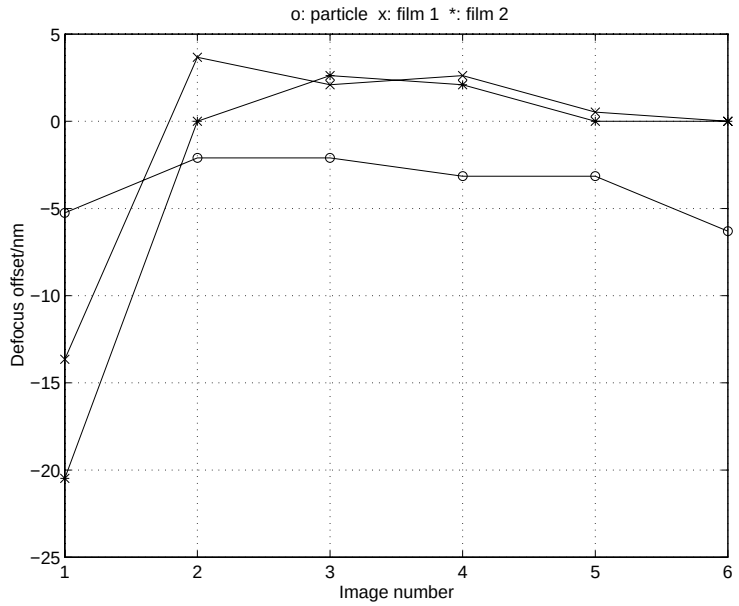


Figure 5.20: Relative variation of defocus offset for areas of tilt azimuth set tilts2

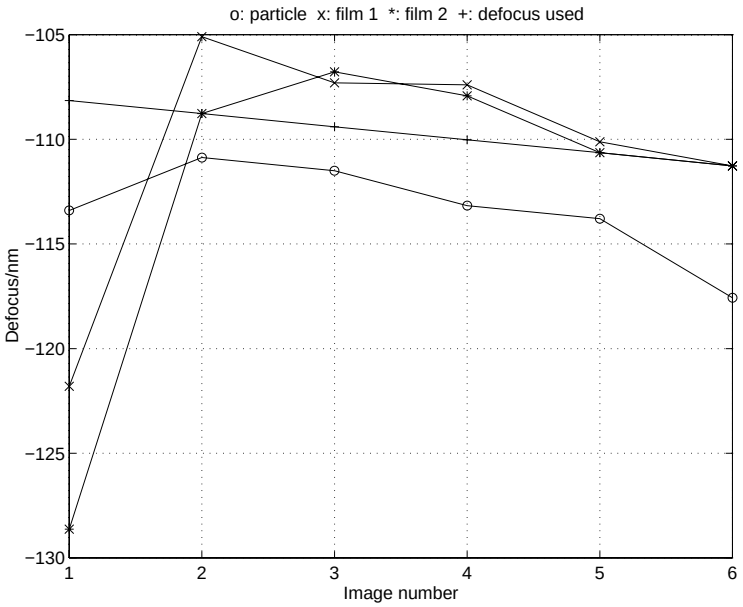


Figure 5.21: Absolute variation of defocus offset for areas of tilt azimuth set tilts2

support film. The variation of this offset over the tilt set for the three areas relative to the actual defocus used in the reconstruction is shown in figure 5.20. The absolute defocus values are shown in figure 5.21 in which the actual defocus used in the reconstruction is also marked.

Figures 5.20 and 5.21 confirm that focal drift is present. The most striking result is the consistent focus difference between the film and the particle which is close to the expected value of 50 Å. The anomaly of the defocus offsets in the first image may be due to a jump in the focus from image 1 to image 2.

5.4.2 Block oxide set

Tilts1 is a set of block oxide images at a defocus of 1 Sch underfocus and a tilt of 1 Gl/Sch. The sampling interval is 0.2825 Å/pixel and the members of the set are shown in figure A.3.

5.4.2.1 Determination of imaging conditions

The diffractograms for the set are shown in figure 5.22. These were very difficult to fit so it was decided to try out a new technique for finding the imaging conditions. The XCF tableau for the amorphous film in the sets is shown in figure 5.23. The XCF peak shapes have very distinctive shapes so that it should be possible to fit these rather than diffractograms. The reference image for each XCF had a high-pass Fourier mask of radius 3 pixels in order to improve contrast.

5.4.2.1.1 Fitting the template PSF-XCFs

The peak height in a ST-XCF is the measure of how well the template PSF-XCFs match the XCFs with which they are being cross-correlated. It is possible to regard this peak height as the value of a multi-dimensional function (the ST-XCF peak height as a function of lens aberrations) which has to be maximised. To investigate this fit, Semper was run in batch mode from inside a non-linear least squares program using Powell's direction set method. The non-linear variables, A_{1r}, A_{1i}, τ_{cx}, τ_{cy} and r.m.s. vibration were passed to the Semper process which generated the ST-XCF for a particular XCF from the XCF tableau in 5.23. The peak height of this ST-XCF was

returned to the non-linear fit program to be maximised with respect to the non-linear variables.

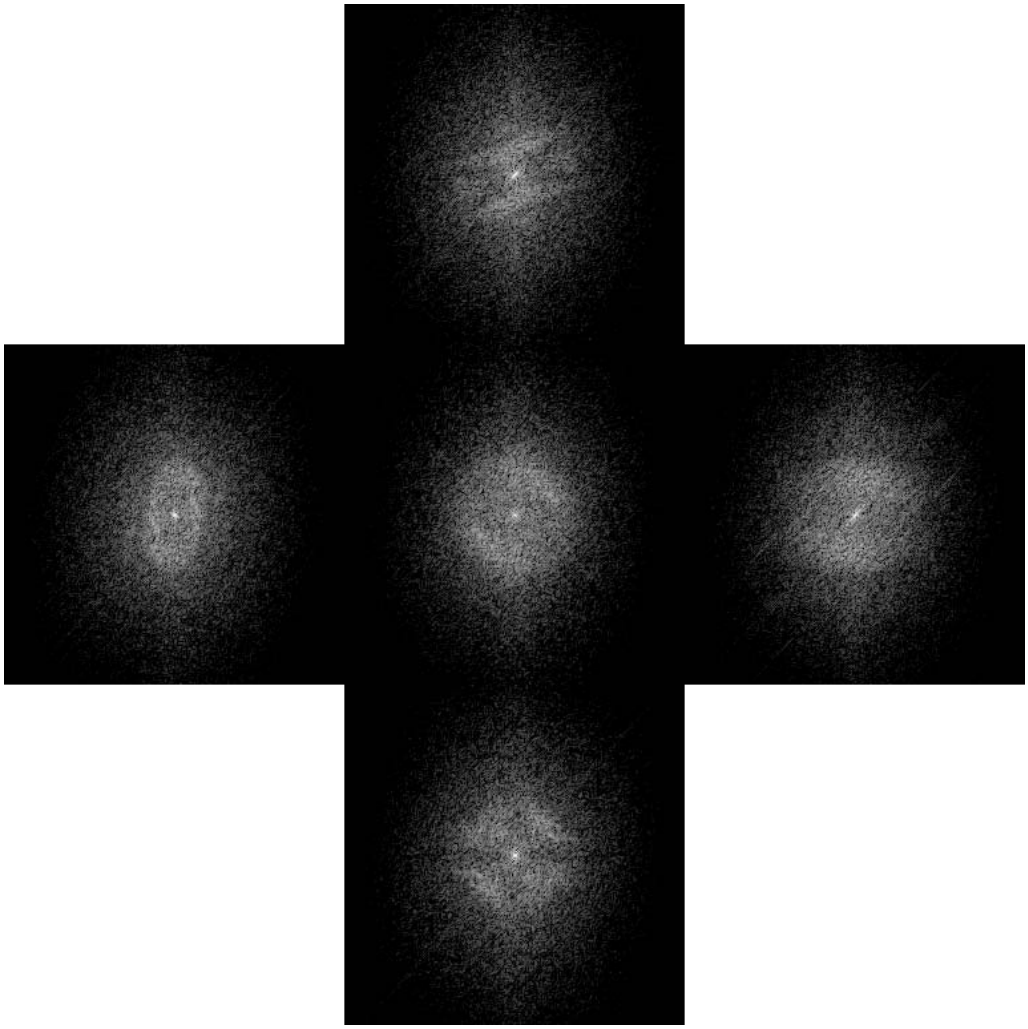


Figure 5.22: Diffractogram tableau for tilt azimuth tilts1 whose parameters are defined by figure A.3. The edge of each field corresponds to a period of 0.14 nm.

The defocus was fixed at -38 nm (1 Sch) and the values of A_{2r} and A_{2i} were fixed at 1.4 and -1.4 μm respectively. The fitted values for the parameters are shown in table 5.4. The significant misalignment returned by this fit is not too

Coefficient	units	fitted value
A_{1r}	nm	7.22
A_{1i}	nm	-15.14
τ_{cx}	mrad	2.11
τ_{cy}	mrad	-0.67
rms vibration	nm	0.13

Table 5.4: Results of fitting template PSF-XCFs for tilts1

surprising after the diffractogram tableau in figure 5.22 has been examined. The template PSF-XCFs and the ST-XCFs based on this fit are shown in figures 5.24 and 5.25 respectively. The XCF chosen for PSF-XCF fitting was that between image 1 and image 5. This XCF and its fitted template PSF-XCF are shown in 5.26 as the left and centre images respectively which indicates a good match. The right image in figure 5.26 shows the effect of ignoring A_2 from which it can be deduced that the aberration has a significant effect on the template PSF-XCFs. It is clear that some template PSF-XCFs in figure 5.24 do not look much like the original XCFs. This is probably the result of fitting only one XCF from the tableau and it is likely that if the sum of the peak heights of all the XCFs were maximised a more reliable fit would be obtained.

5.4.2.2 Reconstruction

Figure 5.27 shows the modulus and phase of the restored specimen exit plane wave function, Ψ for tilts1 using the fitted aberrations in table 5.4. Figure 5.28 shows a reverse video diffractogram of Ψ and the effective CTF, $\mathfrak{T}$. The poor shape of the latter is due to the significant misalignment in the captured images. Figure 5.29 compares the modulus of the restored specimen exit plane wave function with the original axial image and figure 5.30 compares two predicted images with the original images they correspond to. The images are those at tilts $(\theta, 0)$ and $(0, \theta)$. There is only reasonable agreement between the predicted and observed images so that it is likely that the correct imaging conditions have not been found.

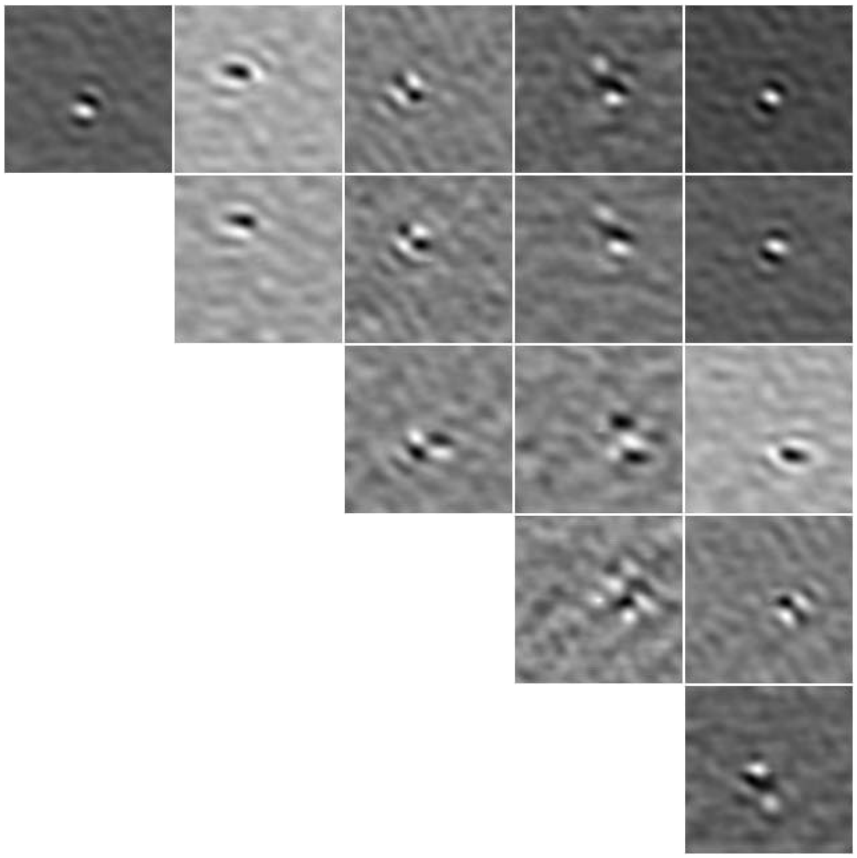


Figure 5.23: XCFs for tilt azimuth set tilts1. The XCF in row r and column c is that between image $c + 1$ and image r.

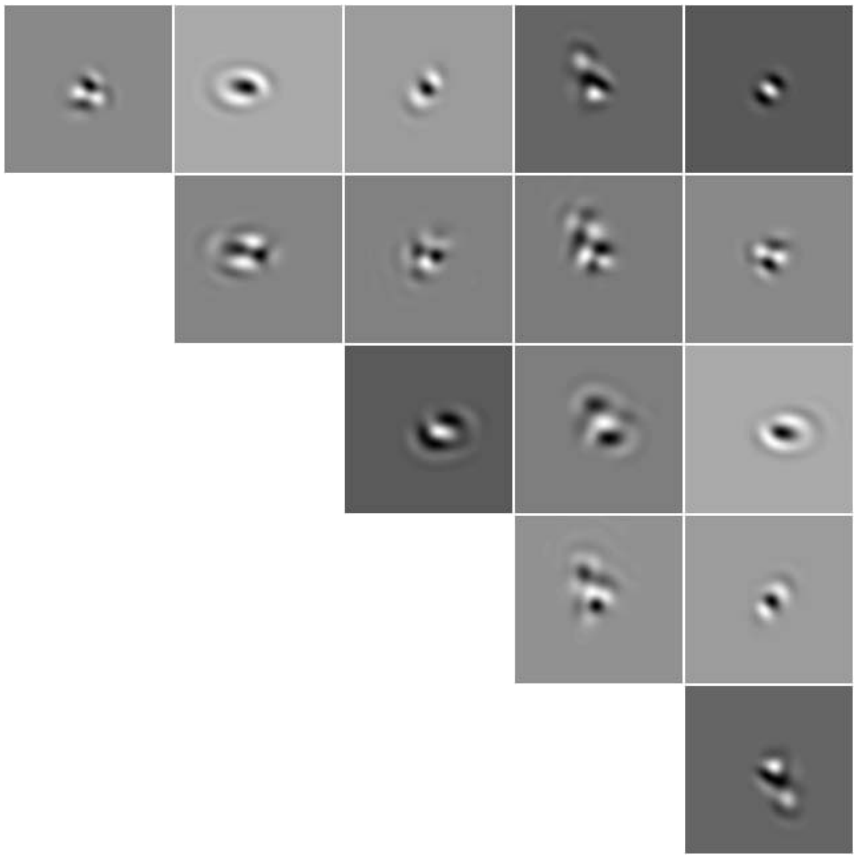


Figure 5.24: Template PSF-XCFs for tilt azimuth set tilts1

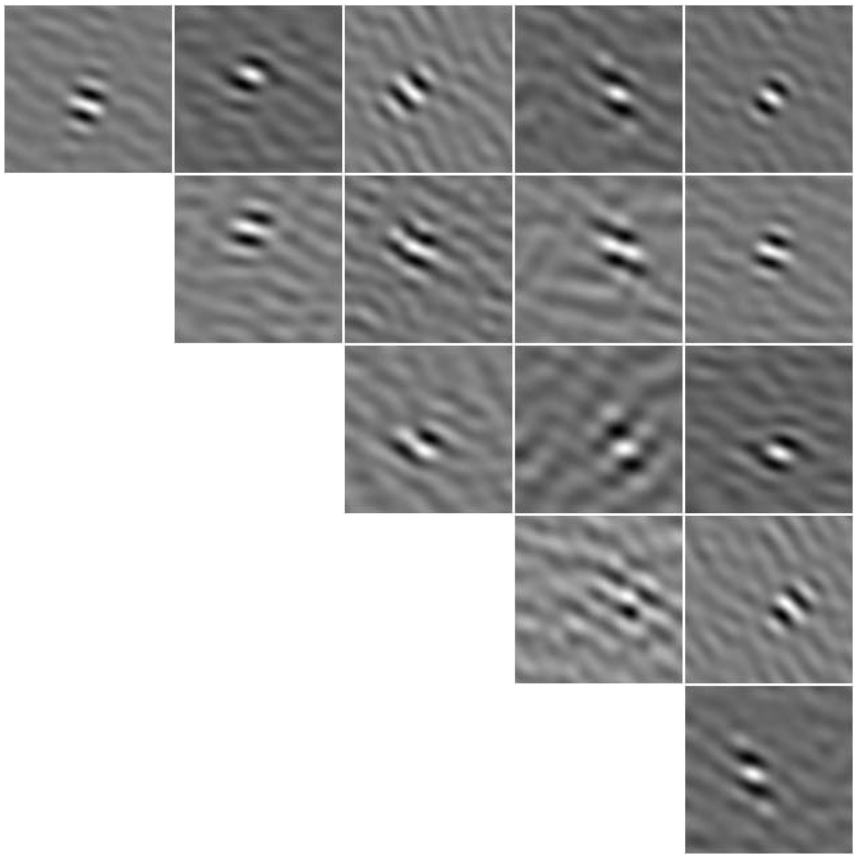


Figure 5.25: ST-XCFs for tilt azimuth set tilts1

Figure 5.26: The actual XCF between images 1 and 5 in tilt azimuth set tilts1 is shown on the left with its fitted template PSF-XCF in the middle. The image on the right is the affect of ignoring A_2 in the template PSF-XCF.

5.5 Comparison of Tilted Beam and Axial Mode Imaging

Tilted beam imaging has a number of advantages over the normal axial mode. The transfer function envelope multiplying the direct wave in the tilted beam mode is displaced relative to the primary beam so that transfer up to twice the maximum spatial frequency in the axial mode is possible in a particular direction. This means that higher resolution can be achieved without the need for a high coherence microscope.

Chau (Chau, K.-L., 1993) investigated the effect of defocus and alignment errors on tilt series reconstruction simulations and found that the reconstruction algorithm was more sensitive to errors in defocus than errors in alignment. Kirkland et al. (Kirkland, A.I. et al., 1995) showed that the tilted beam mode is more tolerant to errors in the measurement of imaging conditions than the axial mode. They additionally noted that the requirements for tilt series image registration are the same as those for focal series at the same resolution: less than 12 pm error for a 1 Å resolution reconstruction. This corresponds to roughly half a pixel in the reconstructions described above and is comfortably met in the phase compensated image registration procedure.

There are two points of comparison which are unfavourable to the tilt-azimuth approach (Kirkland, A.I. et al., 1995). The most serious is the

limitation imposed on specimen thickness by parallax problems associated with tilted beam imaging. The noise level in a tilt-azimuth reconstruction, although identical to that for focal series reconstructions for low spatial frequencies is twice as high as that for focal series reconstructions for high spatial frequencies. This is because four times fewer images contribute to any one component.

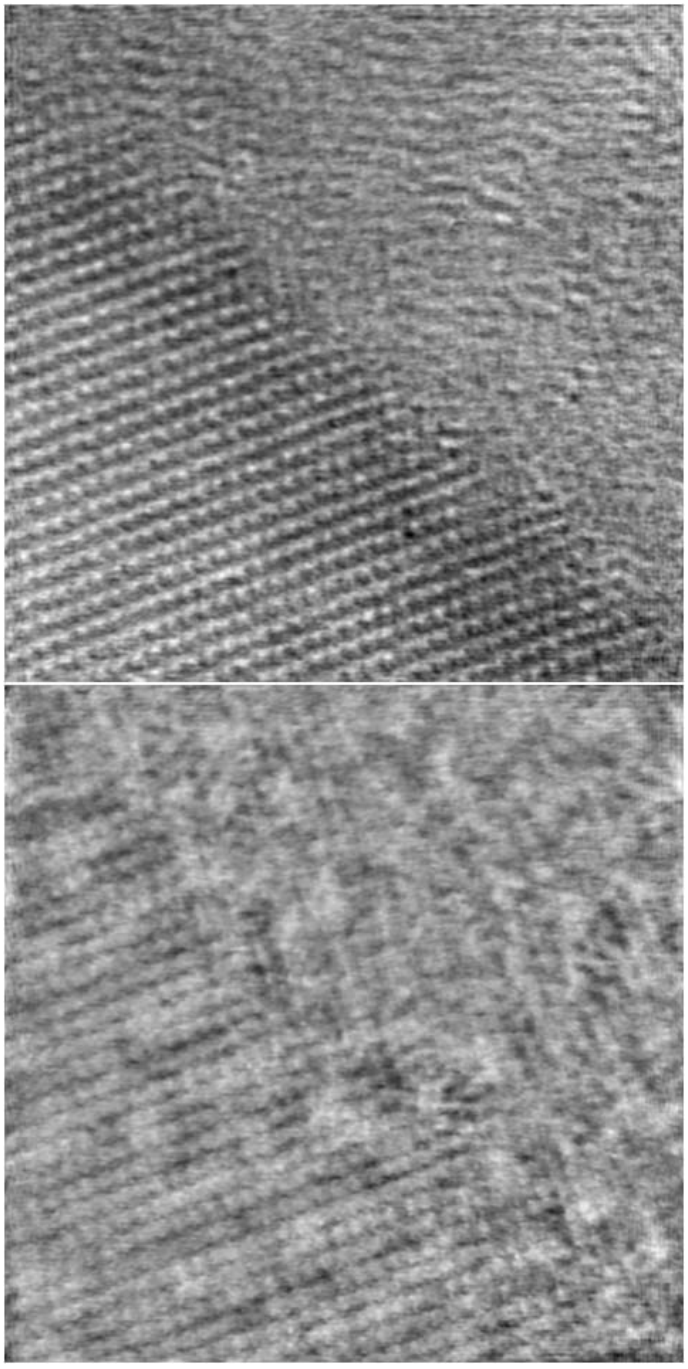


Figure 5.27: Top: modulus and bottom: phase of restored Ψ for tilt azimuth set tilts1. The field of view is 14.5 nm.

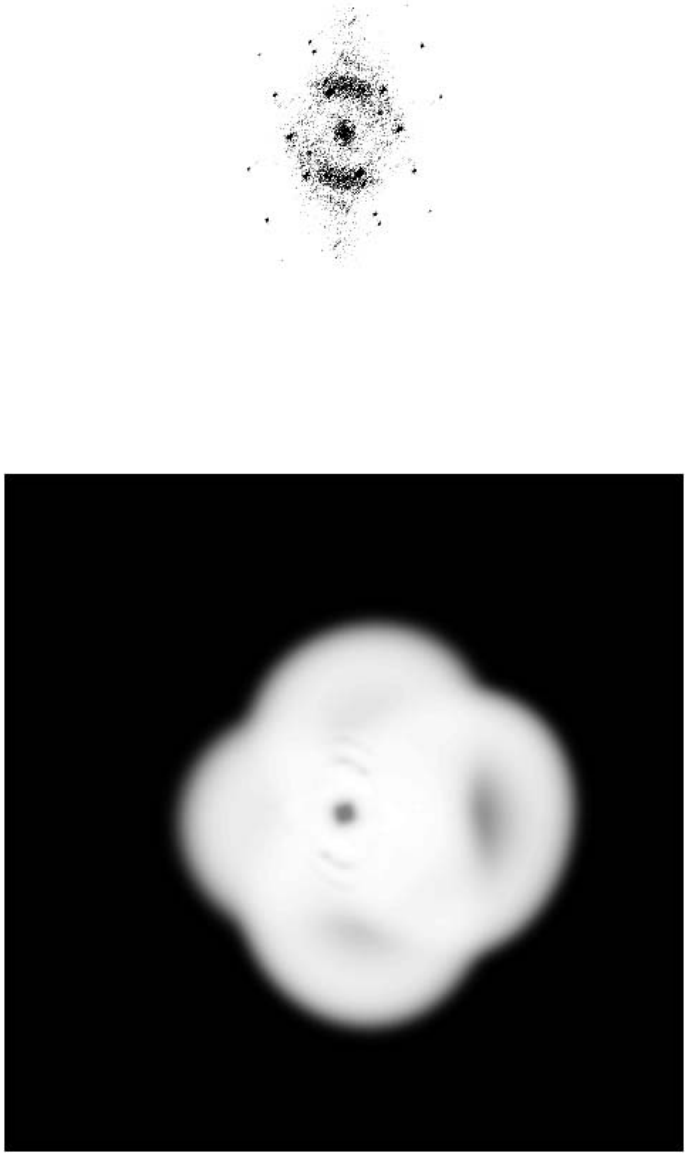


Figure 5.28: Top: diffractogram and bottom: effective CTF of restored Ψ for tilt azimuth set tilts1

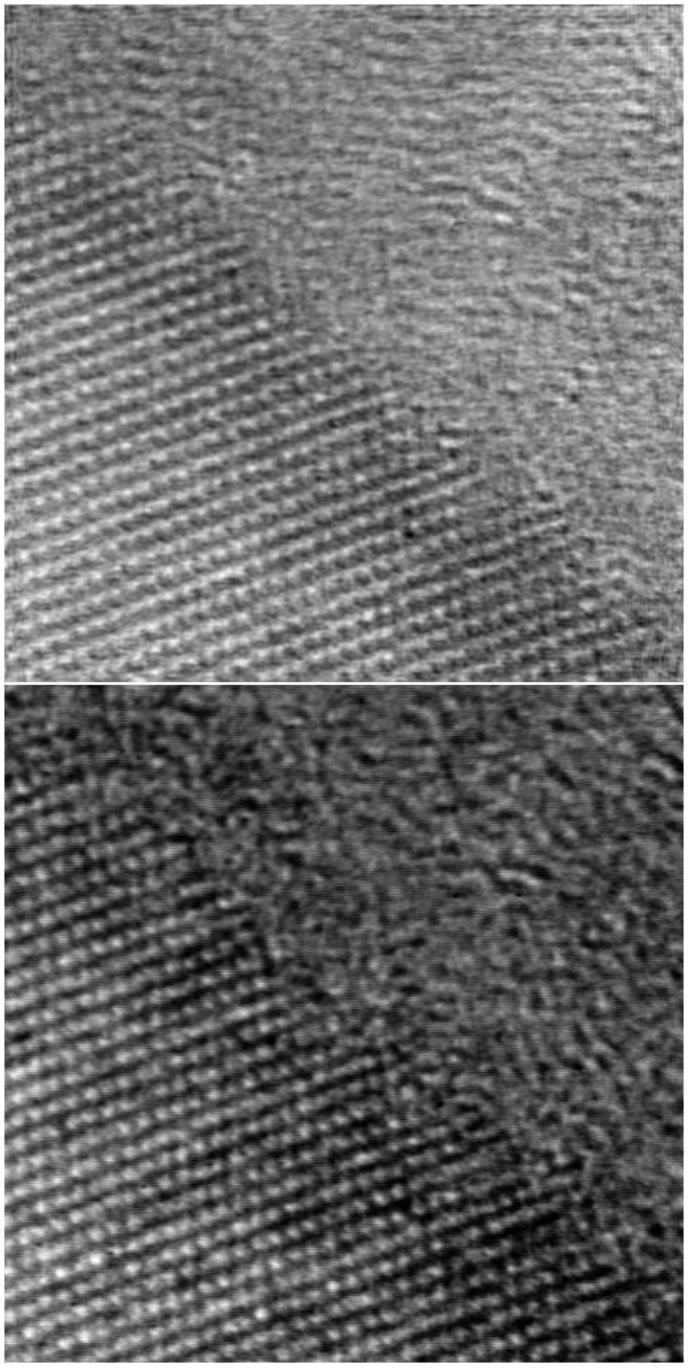


Figure 5.29: Top: modulus of Ψ and bottom: original axial image for tilt azimuth set tilts1. The field of view is 14.5 nm.

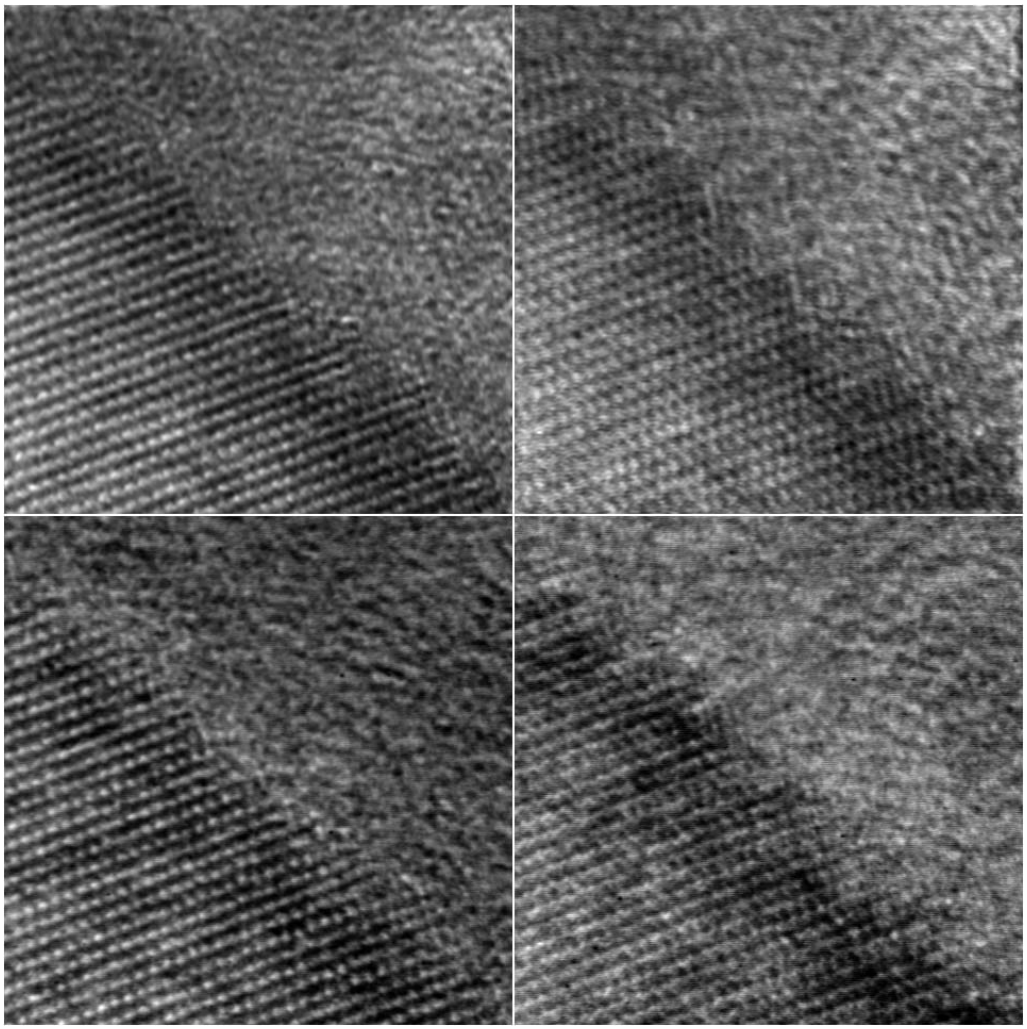


Figure 5.30: Top: predicted images and bottom: original images for two members of tilt azimuth set tilts1. The field of view is 14.5 nm.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

It has been shown that out of all the methods currently available for aberration determination with the aim of automated adjustment, displacement and diffractogram methods hold most promise. The displacement methods can deal with strong and weak objects whereas the diffractogram methods are restricted to amorphous weak phase objects. Application of the displacement method generates two instrumental parameters whereas the diffractogram method provides three. As a result, the latter approach has the advantage of requiring fewer images to be captured. Furthermore, diffractogram methods are less susceptible to specimen drift than the displacement methods. However, it is necessary to specify an assumed value for spherical aberration before diffractograms can be fitted.

The complex number based expressions for the odd and even parts of the wave aberration function have been shown to be equivalent to the corresponding vector based expressions of Koster et al. (Koster, A.J. et al., 1989).

The need for specialised XCF techniques for displacement estimation was demonstrated by examining the effect of the different PSFs in the images being cross-correlated. It was shown that generalised XCF filters do not take these different PSFs into account and are therefore unable to improve the displacement estimate based on the XCF. Phase compensating filters e.g. ST-XCFs do take the different PSFs into account and hence provide reli-

able displacement observations. Only an approximate knowledge of imaging conditions is necessary to generate such filters and the ST-XCF was chosen instead of the PC-XCF, TD-XCF and their variants because it does not depend on an estimate of the signal-to-noise ratio for the images.

The estimation of image displacement based on the linear phase ramp in the cross-power spectrum of the images introduced by Koster et al. (Koster, A.J. et al., 1988; Koster, A.J. et al., 1989) has been shown to be unreliable because of π and 2π ambiguities and the effect of coherence envelopes.

The five tilt displacement method which takes three-fold astigmatism into account is a more accurate variant of the three tilt method introduced by Koster et al. (Koster, A.J. et al., 1989). The four tilt diffractogram method using effective defocus and two-fold astigmatism measurements is a more accurate variant of the method introduced by Krivanek and Fan (Krivanek, O.L. and Fan, G.Y., 1992) and modified by Krivanek and Leber (Krivanek, O.L. and Leber, M.L., 1994). Both the five tilt displacement method and the four tilt diffractogram method have been shown to be useful as techniques for determining approximate aberration values including three-fold astigmatism without the need to do a least squares fit.

Preliminary experimental work done by Kirkland and Chau (Kirkland, A.I. and Chau, K.-L., Private Communication) was based on the linear least squares displacement method described by Chau (Chau, K.-L., 1993) and Saxton (Saxton, W.O., 1994c). It failed to provide consistent values for fitted aberrations and it is now known that this was caused by two factors: the exclusion of the tilt coil calibration matrix and three-fold astigmatism from the analysis, and an unresolved anomaly in the displacement method itself.

This dissertation has clearly shown that both the tilt coil calibration matrix and three-fold astigmatism need to be fitted in least squares displacement and diffractogram methods for accurate aberration determination. This conclusion is based on the fact that the r.m.s. error decreases significantly when these parameters are taken into account. A significantly smaller decrease in r.m.s. error is observed when four-fold astigmatism and axial star aberration are taken into account. It may be concluded that these aberrations can be safely ignored for resolutions lower than 1 Å (strictly speaking : resolution of detail smaller than 1 Å). This assumption is supported by the consistency of the fitted values of the coma-free axis position and three-fold astigmatism using both the least squares displacement and diffractogram methods.

It is disappointing that the only method which appears to be applicable to the majority of specimens suffers from behaviour which is not yet fully understood. Displacement fitting, once specimen drift has been minimised, indicates an inexplicable rotation in the tilt coil calibration matrix. This is in definite need of further investigation. Despite this drawback, the method of using template PSF-XCFs for improving peak definition in XCFs has been crucial in overcoming a major problem with displacement measurement in the past.

The tilt series reconstructions of images from low as well as high coherence microscopes show the great potential for this method now that the hurdles of imaging condition determination and image registration have been overcome. One of the main strengths of tilt series reconstruction over other methods is its ability to provide a near-doubling of resolution from microscopes not equipped with field emission guns.

There has been a successful demonstration of how image matching can determine specimen heights down to the 5 nm level. Finally, the method of matching XCFs of PSFs to actual XCFs shows promise for use in image reconstruction when the imaging conditions are difficult to determine from diffractograms.

6.2 Suggestions for Future Work

6.2.1 The displacement method

The anomaly in the displacement method may be dependent upon the model of microscope used. The experiments performed on the JEOL 4000 EXII microscope which highlighted the anomaly should therefore be carried out on different microscopes from more than one manufacturer for a range of accelerating voltages and spherical aberration values.

The displacement method already appears to work at low resolutions ($>$ 4Å) as Krivanek et al. (Krivanek, O.L. and Fan, G.Y., 1992) point out because specimen drift is less important at these resolutions. Dierksen et al. (Dierksen, K. et al., 1992) describe applications of the method as part of an automatic electron tomography procedure. Koster et al. (Koster, A.J. et al., 1992) have also applied the BID method to electron tomography and incorporate transform modulus adjustments to improve the cross-correlation peak

for displacement estimation. Template PSFs described in this dissertation could be used to improve these displacement estimates further.

If the source of the anomaly in the displacement method can be determined and corrected, specimen drift will become the factor limiting the accuracy of the technique. Kondo et al. (Kondo, Y. et al., 1994) describe the use of piezo devices to perform drift compensation by moving the specimen itself. Better characterisation of specimen drift could also improve the accuracy of the method. A simple implementation would be the capture of tilt sets of images in which every other image is an axial image. This would have the unfortunate side effect of doubling the number of images to be captured and it might be necessary to increase this number further if it was found that one axial image between tilted images was not sufficient to characterise drift adequately.

The specimen drift rate is likely to be a function of the time since the microscope session was started. It is worth investigating whether there is a time after which the magnitude of specimen drift rate is small enough for it not to have a serious effect on the displacement method.

If the displacement method can be made reliable using one of these techniques, a combination of displacement and diffractogram observations, as described in 3.6.5, would allow for estimation of C_3. This value could then be compared to that obtained via image displacements resulting from large but well-calibrated injected tilts using the technique described by Budinger and Glaeser (Budinger, T.F. and Glaeser, R.M., 1976). The elucidation of x-y anisotropy by the iteration described in section 4.10 would also be possible given reliable displacement fitting. Consistent estimates of the tilt coil calibration matrix could then be used to investigate the matrix further with the aim of better characterisation.

6.2.2 The diffractogram method

It has been shown that manual diffractogram fitting is a reliable method for 'off-line' aberration determination. It should be possible to align a microscope 'on-line' in 10-15 minutes by this method. In addition, the determination of three-fold astigmatism by diffractogram fitting would allow for its iterative compensation through the use of a suitable corrector. Although recent developments in automatic diffractogram fitting look promising (Krivanek, O.L. and Fan, G.Y., 1992), the method itself suffers from the serious

constraint on the type of specimen with which it can work.

6.2.3 Three-fold astigmatism

The coma-free axis is routinely found by the examination of diffractogram tableaus (Hosokawa, H. et al., 1993) or the texture of images taken with diametrically opposite tilts in both the x- and y-directions (Smith, D.J. et al., 1983). As these techniques fail to take three-fold astigmatism into account, the position of the coma-free axis obtained in this way is likely to be inaccurate as the least squares displacement and diffractogram fits have shown. Three-fold astigmatism can be corrected by using the existing octopole stigmator, suitably rewired as described by Typke and Dierksen (Typke, D. and Dierksen, K., 1992). Therefore, given the consistent values of this aberration from diffractogram fits, a simple iterative procedure would be involved in setting up the appropriate currents to this corrector as mentioned above.

The three-fold astigmatism value for the JEOL 4000 EXII is unfortunately the highest one measured so far (Saxton, W.O. et al., 1994; Chand, G. et al., 1995) and it is therefore important for it to be compensated for high resolution work (2 to 1 Å). The maximum phase shift caused by the typical three-fold astigmatism value of 1.9 μm in the JEOL 4000 EXII microscope is $\pm 77^\circ$ for 2 Å lattice fringes rising to $\pm 616^\circ$ for 1 Å. For comparison, a $\pm 45^\circ$ phase shift converts a $\sin\gamma$ maximum into a zero. The figure for 2 Å indicates significant image degradation even before the resolution limit (approximately 1.7 Å) so it is important that three-fold astigmatism is corrected for this microscope. For resolutions beyond 1 Å or for spherically corrected electron microscopes (Uhlemann, S. et al., 1994), four-fold astigmatism, axial star aberration and even fifth order spherical aberration may need to be controlled.

6.2.4 Reconstruction

The magnitude of the non-linear contributions to the images undergoing tilt series reconstructions should be analysed. Kirkland's (Kirkland, E.J., 1984) maximum likelihood method could be extended to tilt series reconstructions in order to aid the analysis.

The tilt series reconstruction method could be applied to images from energy filtered microscopes (Gubbens, A.J. and Krivanek, O.L., 1993). An

automated alignment procedure could also be applied to such microscopes which would enable the capture of images under optimal conditions for subsequent reconstruction.

The image matching method for height determination could be used to reconstruct regions in a crystal which are a significant distance from the amorphous region used to characterise imaging conditions. It would also provide an estimate for the angle which the crystal makes with its support film.

The use of template PSF-XCF fitting to aid imaging condition characterisation could be investigated further by using more than one XCF in the fit. It would be necessary to carry out experiments to determine the optimum number of XCFs to be fitted, bearing in mind the requirement to keep the number of calculations per iteration down.

6.2.5 Using a CCD camera

A CCD camera used in place of the current vidicon would increase the field of view and improve upon both the signal to noise ratio of the images and their dynamic range. However, there is a trade off between resolution, signal to noise ratio and frame transfer rate for such an instrument. Although video rate capture (25 frames/second) is possible for large field images (1024x1024 and higher), the poor signal to noise ratio excludes the use of such images for aberration fitting or reconstruction. Consequently, it is likely that a hybrid system will be used in practise. In this, a low resolution vidicon image would be used by the microscope operator to find an area on the specimen at video rates, and thereafter, the full CCD camera field would be used to record images for aberration fitting or reconstruction.

6.2.6 Increasing computing power

The displacement method has at its heart, the calculation of a large number of XCFs. These are more computationally expensive than the non-linear displacement fit itself. Clearly, the speed of the XCF calculation determines how fast the displacement fit can provide aberration estimates for real-time control. Using a number of powerful workstations in parallel over a network (Chand, G. and Holburn, D.M., 1994) shows promise as a means of decreasing the time taken to align the microscope and is presented in appendix C.

References

Al-Ali, L. and Frank, J. (1980). Resolution estimation in electron microscopy. *Optik*, 56(1):31–40.

Baba, N., Eisaku, O., and Kanaya, K. (1987). An algorithm for on-line digital image processing for assisting automatic focusing and astigmatism correction in electron microscopy. *Scanning microscopy*, 1(4):1507–1514.

Barry, J.C. (1992). Image-matching as a means of atomic structure evaluation in high resolution transmission electron microscopy. *Scanning microscopy supplement*, 6:209–221.

Bendat, J.S. and Piersol, A.G. (1971). *Random Data: Analysis and Measurement Procedures*. Wiley-Interscape, New York.

Budinger, T.F. and Glaeser, R.M. (1976). Measurement of focus and spherical aberration of an electron microscope objective lens. *Ultramicroscopy*, 2:31–41.

Carter, G.C. (1987). Coherence and time delay estimation. *Proceedings of the IEEE*, 75(2):236–255.

Catto, C.J.D., Smith, K.C.A., Nixon, W.C., Erasmus, S.J.E., and Smith, D.J. (1981). An image pickup and display system for the Cambridge University HREM. In Goringe, M.J., editor, *Electron Microscopy and Analysis: Cambridge, Inst. Phys. Conf. Ser. No. 61*, pages 123–126. Institute of Physics, Bristol.

Chand, G. and Holburn, D.M. (1993). Method for cross-correlating images in the presence of geometric distortion. *IEE Electronics Letters*, 29(12):1078–1079.

Chand, G. and Holburn, D.M. (1994). A real-time control application using asynchronous remote procedure calls. In Deaton, E., Oppenheim, D., Urban, J., and Berghel, H., editors, *Proceedings of the Symposium on Applied Computing 1994, Phoenix, Arizona*, pages 515–517. ACM Press, New York.

Chand, G., Saxton, W.O., and Kirkland, A.I. (1995). Aberration measurement and automated alignment of the TEM. In *Electron Microscopy and Analysis: Birmingham, Inst. Phys. Conf. Ser.* Institute of Physics, Bristol. To be submitted.

Chang, M.M.Y. (1988). *A computer-controlled system in transmission electron microscopy*. PhD dissertation, Cambridge University, Department of Engineering.

Chau, K.-L. (1993). *Automated control in high resolution electron microscopy*. PhD dissertation, Cambridge University, Department of Engineering.

Coene, W.M.J, Janssen, A.J.E.M., Op de Beek, M., van Dyck, D., van Zwet, E.J., and Zandbergen, H.W. (1993). Focus-variation image reconstruction in field emission TEM. In Bailey, G.W. and Rieder, C.L., editors, *Proceedings of the 51st Annual Meeting of the Microscopy Society of America, Cincinnati, Ohio*, pages 1070–1071. San Francisco Press, Inc., San Francisco.

Cosslett, V.W., Camps, R.A., Saxton, W.O., Smith, D.J., Nixon, W.C., Ahmed, H., Catto, C.D.J., Cleaver, J.R.A., Smith, K.C.A., Timbs, A.E., Turner, P.W., and Ross, P.M. (1979). Atomic resolution with the 600-kV electron microscope. *Nature*, 281:49–51.

de Jong, A.F. and Koster, A.J. (1992). Measurement of electron optical parameters for high-resolution electron microscopy image interpretation. *Scanning microscopy supplement*, 6:95–103.

Dierksen, K., Typke, D., Hegerl, R., Koster, A.J., and Baumeister, W. (1992). Towards automated electron tomography. *Ultramicroscopy*, 40:71–87.

Downing, K.H., Koster, A.J., and Typke, D. (1992). Overview of computer-aided electron microscopy. *Ultramicroscopy*, 46:189–197.

Erasmus, S.J. (1982). *Real-time image processing in electron microscopy*. PhD dissertation, Cambridge University, Department of Engineering.

Erasmus, S.J. and Smith, K.C.A. (1982). An automatic focusing and astigmatism correction system for the SEM and CTEM. *Journal of Microscopy*, 127:185–199.

Frank, J. (1972). A study on heavy/light atom discrimination in bright field electron microscopy using the computer. *Journal of Biophysics*, 12:484–511.

Frank, J. (1974). The envelope of electron microscope transfer functions for partially coherent illumination. *Optik*, 38(5):519–536.

Frank, J. (1975). Controlled focusing and stigmating in the conventional and scanning transmission electron microscope. *Journal of Physics E: Scientific Instruments*, 8:582–587.

Frank, J. (1980). The role of correlation techniques in computer image processing. In Hawkes, P.W., editor, *Computer Processing of Electron Microscope Images*. Springer-Verlag, Berlin.

Gubbens, A.J. and Krivanek, O.L. (1993). Applications of a postcolumn imaging filter in biology and materials science. *Ultramicroscopy*, 51:146–159.

Hanszen, K.-J. (1971). The optical transfer theory of the electron microscope: fundamental principles and applications. In Barer, R. and Cosslett, V.E., editors, *Advances in Optical and Electron Microscopy*, volume 4. Academic Press, London.

Hanszen, K.-J. and Trepte, L. (1971). The influence of voltage and current fluctuations and of a finite energy width of the electrons on contrast and resolution in electron microscopy. *Optik*, 32(6):519–538.

Hawkes, P.W. (1980). Units and conventions in electron microscopy. *Ultramicroscopy*, 5:67–70.

Hines, R.L. (1984). Practical problems in high-resolution microscopy. In Barer, R. and Cosslett, V.E., editors, *Advances in Optical and Electron Microscopy*, volume 9. Academic Press, London.

Horiuchi, S., Maramatsu, K., and Matsui, Y. (1978). The crystal structure of $4Nb_2O_5.9WO_3$ studied by 1 MV high-resolution electron microscopy. *Acta Cryst.*, 34:939–946.

Hosokawa, H., Kondo, Y., Honda, T., and Critchell, J. (1993). Online alignment system for TEM using a TV and personal computer. In Craven, A.J., editor, *Electron Microscopy and Analysis: Liverpool, Inst. Phys. Conf. Ser. No. 138*, pages 531–534. Institute of Physics, Bristol.

Ishizuka, K. and Shirota, K. (1994). Voltage-center coma-free alignment for high-resolution electron microscopy. In Bailey, G.W. and Garratt-Reed, A.J., editors, *Proceedings of the 52nd Annual Meeting of the Microscopy Society of America, New Orleans, Louisiana*, pages 410–411. San Francisco Press, Inc., San Francisco.

Jenkins, W.K. (1979). *Contrast transfer in bright field electron microscopy of amorphous specimens*. PhD dissertation, Cambridge University, Department of Physics.

Jenkins, W.K. and Wade, R.H. (1977). Contrast transfer in the electron microscope for tilted and canonical bright field illumination. In Misell, D.L., editor, *Electron Microscopy and Analysis: Glasgow, Inst. Phys. Conf. Ser. No. 36*, pages 115–118. Institute of Physics, Bristol.

Kirkland, A.I., Saxton, O., Chand, G.C., Tsuno, K., and Kawasaki, M. (1994). Experimental super resolution in the TEM: image reconstruction from a tilt azimuth series. In Jouffrey, B., Colliex, C., Chevalier, J.P., Glas, F., Hawkes, P.W., Hernandez-Verdun, D., Schrevel, J., and Thomas, D., editors, *Electron Microscopy 1994 - Proceedings of 13th ICEM, Paris*, volume 1, pages 463–464. Les éditions de physique, Les Ulis Cedex A.

Kirkland, A.I., Saxton, W.O., Chau, K.-L., Tsuno, K., and Kawasaki, M. (1995). Super-resolution by aperture synthesis: tilt series reconstruction in CTEM. *Ultramicroscopy*, 57:355–374.

Kirkland, E.J. (1984). Improved high resolution image processing of bright field electron micrographs. I. Theory. *Ultramicroscopy*, 15:151–172.

Knapp, C.H. and Carter, G.C. (1976). The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-24(4):320–326.

Kondo, Y., Hosokawa, F., Ohkura, Y., Hamochi, M., Nakagawa, Y., Kirkland, A.I., and Honda, T. (1994). Development of a specimen drift correction system using a personal computer and piezo devices. In Jouffrey, B., Colliex, C., Chevalier, J.P., Glas, F., Hawkes, P.W., Hernandez-Verdun, D., Schrevel, J., and Thomas, D., editors, *Electron Microscopy 1994 - Proceedings of 13th ICEM, Paris*, volume 1, pages 275–276. Les éditions de physique, Les Ulis Cedex A.

Koster, A.J., Chen, H., Sedat, J.W., and Agard, D.A. (1992). Automated microscopy for electron tomography. *Ultramicroscopy*, 46:207–227.

Koster, A.J. and de Ruijter, W.J. (1992). Practical autoalignment of transmission electron microscopes. *Ultramicroscopy*, 40:89–107.

Koster, A.J., de Ruijter, W.J., van den Bos, A., and van der Mast, K.D. (1989). Autotuning of a TEM using minimum electron dose. *Ultramicroscopy*, 27:251–272.

Koster, A.J., van den Bos, A., and van der Mast, K.D. (1987). An autofocus method for a TEM. *Ultramicroscopy*, 21:209–222.

Koster, A.J., van den Bos, A., and van der Mast, K.D. (1988). Signal processing for autotuning by beam tilt induced image displacement. *Scanning microscopy supplement*, 2:83–92.

Krivanek, O.L. (1976). A method for determining the coefficient of spherical aberration from a single electron micrograph. *Optik*, 45:97–101.

Krivanek, O.L. and Fan, G.Y. (1992). Application of slow-scan charge-coupled device (CCD) cameras to on-line microscope control. *Scanning microscopy supplement*, 6:105–114.

Krivanek, O.L. and Leber, M.L. (1993). Three-fold astigmatism: an important TEM aberration. In Bailey, G.W. and Rieder, C.L., editors, *Proceedings of the 51st Annual Meeting of the Microscopy Society of America, Cincinnati, Ohio*, pages 972–973. San Francisco Press, Inc., San Francisco.

Krivanek, O.L. and Leber, M.L. (1994). Autotuning for 1 Angstrom resolution. In Jouffrey, B., Colliex, C., Chevalier, J.P., Glas, F., Hawkes, P.W., Hernandez-Verdun, D., Schrevel, J., and Thomas, D., editors, *Electron Microscopy 1994 - Proceedings of 13th ICEM, Paris*, volume 1, pages 157–158. Les éditions de physique, Les Ulis Cedex A.

LePoole, J.B. (1947). A new electron microscope with continuously variable magnification. *Philips Technical Review*, 2.

Lichte, H. (1991). Electron image plane off-axis holography of atomic structures. In Mulvey, T. and Sheppard, C.J.R., editors, *Advances in Optical and Electron Microscopy*, volume 12. Academic Press, London.

McCartney, M.R. (1994). Introduction to electron holography. In Bailey, G.W. and Garratt-Reed, A.J., editors, *Proceedings of the 52nd Annual Meeting of the Microscopy Society of America, New Orleans, Louisiana*, pages 396–397. San Francisco Press, Inc., San Francisco.

McFarlane, S.C. (1975). The imaging of amorphous specimens in a tilted-beam electron microscope. *Journal of Physics C: Solid State Physics*, 8:2819–2836.

Misell, D.L. and Brown, E.B. (1987). Electron diffraction: an introduction for biologists. In Glauert, A.M., editor, *Practical Methods in Electron Microscopy*. Elsevier Science Publishers BV. (Biomedical Division) - The Netherlands.

Papoulis, A. (1989). *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill International, 2nd edition.

Pinson, L.J. (1985). *Electro-optics*. John Wiley and Sons, Inc., U.S.A.

Pitt, A.J. (1980). *Image formation and interpretation at the limits of conventional transmission electron microscopy*. PhD dissertation, Cambridge University, Department of Physics.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. (1992). *Numerical Recipes in C*. Cambridge University Press, 2nd edition.

Saxton, W.O. (1978). *Computer Techniques for Image Processing in EM*. Academic, New York.

Saxton, W.O. (1986a). Focal series restoration in HREM. In Imura, T., Suzuki, T., and Maruse, S., editors, *Proceedings of the 11th ICEM, Kyoto, Post Deadline Paper*, pages 1–4. The Japanese Society of Electron Microscopy.

Saxton, W.O. (1986b). HREM images and atom positions. In Bailey, G.W., editor, *Proceedings of the 44th Annual Meeting of the Electron Microscopy Society of America, Albuquerque, New Mexico*, pages 526–529. San Francisco Press, Inc., San Francisco.

Saxton, W.O. (1988). Accurate atom positions from focal and tilted beam series of high resolution electron micrographs. *Scanning microscopy supplement*, 2:213–224.

Saxton, W.O. (1994a). Accurate alignment of sets of images. *Journal of Microscopy*, 174:61–68.

Saxton, W.O. (1994b). Four-fold aperture synthesis, three-fold astigmatism, and two-fold resolution improvement: computer aided microscopy comes of age. In Bailey, G.W. and Garratt-Reed, A.J., editors, *Proceedings of the 52nd Annual Meeting of the Microscopy Society of America, New Orleans, Louisiana*, pages 908–909. San Francisco Press, Inc., San Francisco.

Saxton, W.O. (1994c). Tilt-shift analysis for TEM auto-adjustment: a better solution to the data fitting problem. *Journal of Computer Aided Microscopy*, 6:61–76.

Saxton, W.O. (1994d). What is the focus variation method? Is it new? Is it direct? *Ultramicroscopy*, 55:171–181.

Saxton, W.O., Chand, G., and Kirkland, A.I. (1994). Accurate determination and compensation of lens aberrations in high resolution EM. In Jouffrey, B., Colliex, C., Chevalier, J.P., Glas, F., Hawkes, P.W., Hernandez-Verdun, D., Schrevel, J., and Thomas, D., editors, *Electron Microscopy 1994 - Proceedings of 13th ICEM, Paris*, volume 1, pages 203–204. Les éditions de physique, Les Ulis Cedex A.

Saxton, W.O., Howie, A., Mistry, A., and Pitt, A. (1977). Fact and artefact in high resolution electron microscopy. In Misell, D.L., editor, *Electron Microscopy and Analysis: Glasgow, Inst. Phys. Conf. Ser. No. 36*, pages 119–122. Institute of Physics, Bristol.

Saxton, W.O. and O'Keefe, M.A. (1981). High resolution - in spite of misalignment? In Goringe, M.J., editor, *Electron Microscopy and Analysis: Cambridge, Inst. Phys. Conf. Ser. No. 61*, pages 343–346. Institute of Physics, Bristol.

Saxton, W.O., Pitt, H.J., and Horner, M. (1979). Digital image processing: the Semper system. *Ultramicroscopy*, 4:343–354.

Saxton, W.O., Smith, D.J., and Erasmus, S.J. (1983). Procedures for focusing, stigmating and alignment in high resolution electron microscopy. *Journal of Microscopy*, 130:187–201.

Schiske, P. (1968). Zur frage der bildrekonstruktion durch fokusreihen. In Bocciarelli, D.S., editor, *Proceedings of the 4th European Congress on Electron Microscopy, Rome*, volume 1, pages 145–146. Tipographia Poliglotta Vaticana.

Schiske, P. (1973). Image processing using additional statistical information about the object. In Hawkes, P.W., editor, *Image Processing and Computer Aided Design in Electron Optics*. Academic Press, London.

Slayter, E.M. and Slayter, H.S. (1992). *Light and Electron Microscopy*. Cambridge University Press.

Smith, D.J. (1989). Instrumentation and operation for high resolution electron microscopy. In Mulvey, T. and Sheppard, C.J.R., editors, *Advances in Optical and Electron Microscopy*, volume 11. Academic Press, London.

Smith, D.J., Saxton, W.O., O'Keefe, M.A., Wood, G.J., and Stobbs, W.M. (1983). The importance of beam alignment and crystal tilt in high resolution electron microscopy. *Ultramicroscopy*, 11:263–282.

Thon, F. (1966). The defocussing of phase contrast electron microscope images. *Z. Naturforsch*, 21a:476–478.

Typke, D. and Dierksen, K. (1992). Determination of image aberrations by diffractogram and cross-correlation methods. In Heydenreich, J. and Neumann, W., editors, *Proceedings of the Autumn School 1992 of the International Centre of Electron Microscopy at the Max Planck Institute of Microstructure Physics, Halle/Saale*, pages 96–107. Elbe Druckerei Wittenberg.

Uhlemann, S., Haider, M., and Rose, H. (1994). Procedures for adjusting and controlling the alignment of a spherically corrected electron microscope. In Jouffrey, B., Colliex, C., Chevalier, J.P., Glas, F., Hawkes, P.W., Hernandez-Verdun, D., Schrevel, J., and Thomas, D., editors, *Electron Microscopy 1994 - Proceedings of 13th ICEM, Paris*, volume 1, pages 193–194. Les éditions de physique, Les Ulis Cedex A.

van der Mast, K.D. (1984). Transmission electron microscopy: state of the art and future developments. In Csanády, Á., Röhlich, P., and Szabó, D., editors, *Proceedings of the 8th European Congress on Electron Microscopy, Budapest*, volume 1, pages 3–10. Programme Committee of the 8th European Congress on Electron Microscopy.

van Dyck, D. and Op de Beeck, M. (1990). New direct methods for phase and structure retrieval by HREM. In Peachey, L.D. and Williams, D.B., editors, *Proceedings of 12th ICEM, Seattle*, pages 26–27. San Francisco Press, Inc., San Francisco.

van Dyck, D. and Op de Beeck, M. (1993). Direct structure information using the focus-variation method. In Bailey, G.W. and Rieder,

C.L., editors, *Proceedings of the 51st Annual Meeting of the Microscopy Society of America, Cincinnati, Ohio*, pages 1068–1069. San Francisco Press, Inc., San Francisco.

van Heel, M., Schatz, M., and Orlova, E. (1992). Correlation functions revisited. *Ultramicroscopy*, 46:307–316.

Wade, R.H. (1992). A brief look at imaging and contrast transfer. *Ultramicroscopy*, 46:145–156.

Wade, R.H. and Jenkins, W.K. (1978). Tilted beam electron microscopy: the effective coherent aperture. *Optik*, 50(5):582–584.

Zemlin, F., Weiss, K., Schiske, P., Kunath, W., and Herrmann, K.-H. (1978). Coma-free alignment of high resolution electron microscopes with the aid of optical diffractograms. *Ultramicroscopy*, 3:49–60.

Appendix A

Description of Data sets

A.1 Introduction

This appendix details the sets analysed in this dissertation. These are a small number of those actually recorded and examined over a two year period.

A.2 JEOL 4000EXII Sets

For all JEOL 4000 EXII sets, the microscope was manually aligned to the voltage centre at an accelerating voltage of 400 kV and images were captured with a sampling interval of 0.2825 Å/pixel. Given a spherical aberration coefficient of 0.9 mm, 1 Sch is 38.46 nm and 1 Gl/Sch is 6.54 mrad.

A.2.1 7 by 7 sets

In all 7 sets containing 51 members were analysed containing images of gold particles on Germanium. The sets begin and end with an axial image and contain 49 tilted images on a 7 by 7 grid where the grid size corresponds to a tilt of 0.6 Gl/Sch. 3 of these sets were captured with a C2 aperture size of CA3 (50 μm) and 4 with a C2 aperture size of CA4 (20 μm). 3 sets were captured in one microscope session at Gaussian focus. The microscope was manually aligned to the voltage centre before capturing each set in this session. The spot size and the C2 aperture were varied : **7x7a** had a spot size of 2 and a C2 aperture size of CA3, **7x7b** had a spot size of 2 and a

C2 aperture size of CA4, **7x7c** had a spot size of 3 and a C2 aperture size of CA3. The remaining CA3 set (**7x70**) and CA4 sets were captured in a different microscope session. In this the microscope was manually aligned to the voltage centre before the first set and left there. Two sets in this session were captured at Gaussian focus: **7x70** and **7x7c0**, one at 1 Sch underfocus (**7x7c1**) and one at 2 Sch underfocus (**7x7c2**). The 10 member crosses described next were captured in the same session straight after the **7x7c2** set. Set **7x7c1** is shown in figure A.1.

Set Name	Spot size	C2 size/μm	Def/Sch	Time captured
7x7a	2	50	0	5-12-93 18:18
7x7b	2	20	0	5-12-93 18:26
7x7c	3	50	0	5-12-93 18:40
7x70	-	50	0	21-8-94 17:23
7x7c0	-	20	0	21-8-94 17:57
7x7c1	-	20	-1	21-8-94 18:06
7x7c2	-	20	-2	21-8-94 18:15
10mx4u	-	20	-4	21-8-94 18:26
10mx8u	-	20	-8	21-8-94 18:28

Table A.1: JEOL 4000 EXII sets

A.2.2 10 member crosses

2 sets containing 10 images of gold particles on Germanium were analysed. Both were captured with a C2 aperture size of CA4 (20 μm). The sets begin and end with an axial image and contain four orthogonal tilted images at tilts of 1 and 2 Gl/Sch as illustrated in figure A.2. The first set was captured at 4 Sch underfocus (**10mx4u**) and the second at 8 Sch (**10mx8u**).

A.2.3 Tilt set

1 set (**tilts1**) containing 6 images of the block oxide $4Nb_2O_5.9WO_3$ (Horiuchi, S. et al., 1978) was analysed. The set begins and ends with an axial image

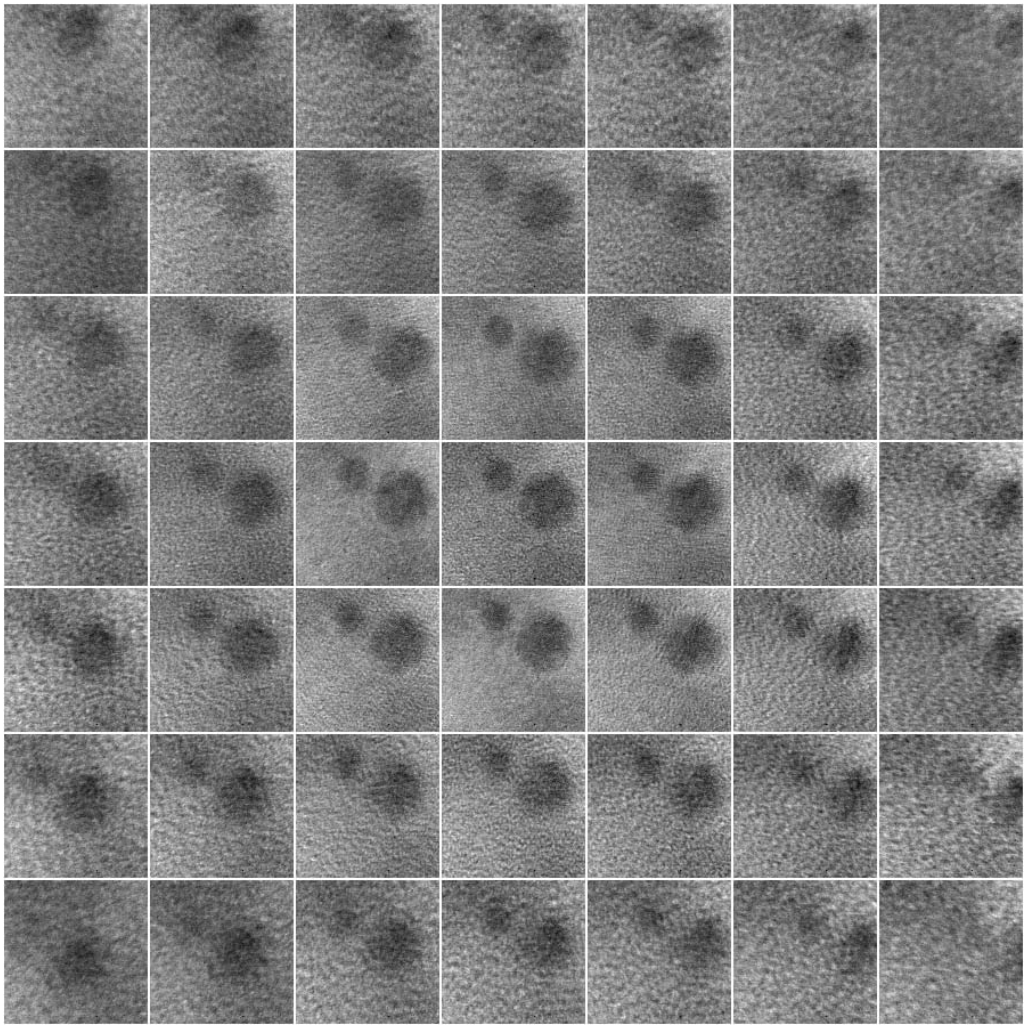


Figure A.1: Set **7x7c1** : Gold particles on an amorphous germanium film captured on the JEOL 4000 EXII microscope at 400 kV and 1 Sch underfocus by Dr. A.I. Kirkland. Cs=0.9 mm and field of view is 14.5 nm. The set was captured with tilts injected on a 7 by 7 grid where the grid size is 0.6 GI/Sch. The central image is the axial image so that the top right hand image has an injected tilt of 1.8 GI/Sch mrad in both the x- and y-directions.

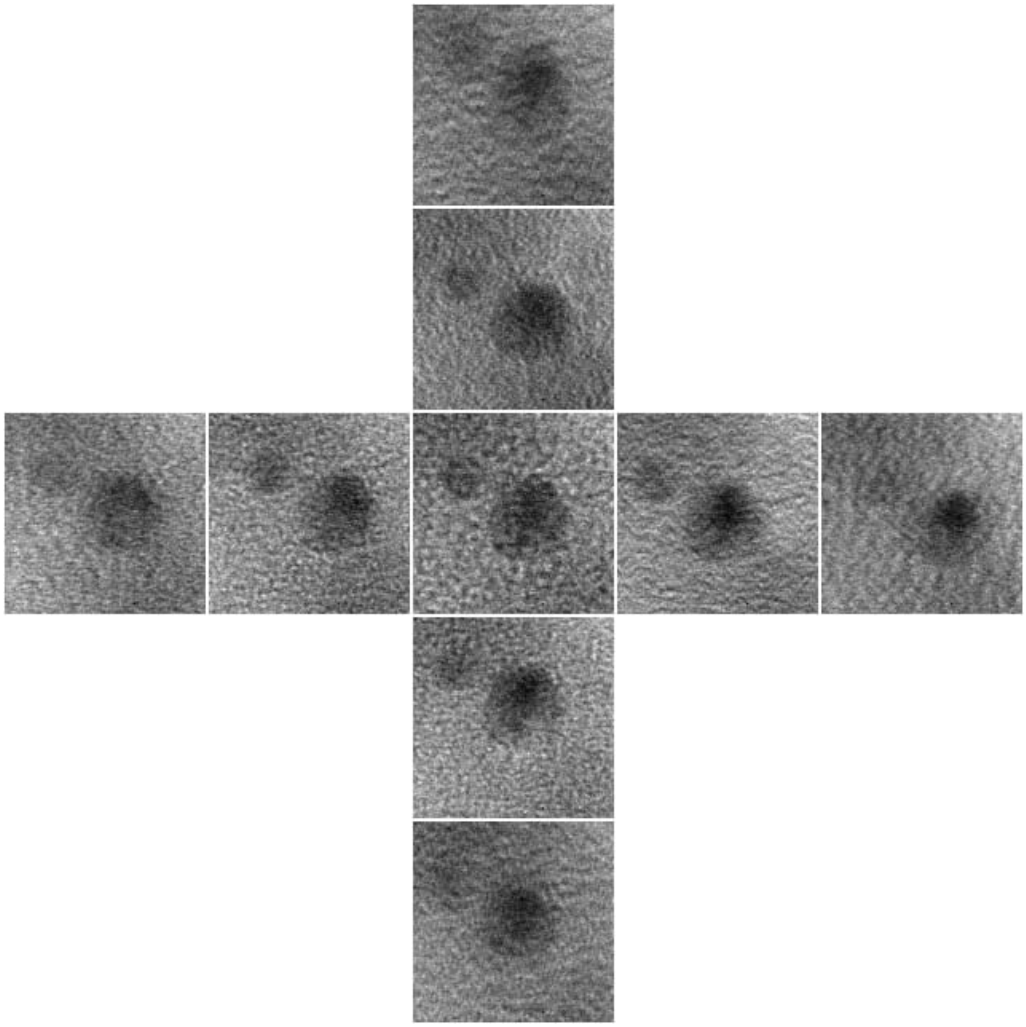


Figure A.2: Set **10mx4u** : Gold particles on an amorphous germanium film captured on the JEOL 4000 EXII microscope at 400 kV and 4 Sch underfocus by Dr. A.I. Kirkland. Cs=0.9 mm and field of view is 14.5 nm. The central image is the axial image and the injected tilts are 1 GI/Sch and 2 GI/Sch in both the x- and y-directions with azimuths as shown in the figure.

and contains 4 orthogonally tilted images. It was captured at a defocus of 1 Sch underfocus with a tilt magnitude of 1 Gl/Sch. The set was captured on 24-1-94 and is shown in figure A.3

A.3 JEOL 2010F sets

For all JEOL 2010F sets, the microscope was manually aligned to coma-free axis using four diffractograms as a guide (Hosokawa, H. et al., 1993). The accelerating voltage was 200 kV and images were digitised over 2048 pixels square at sampling interval of 0.333 Å/pixel. The spherical aberration coefficient is 1.1 mm so that 1 Sch is 52.52 nm and 1 Gl/Sch is 6.91 mrad.

A.3.1 Tilt sets

3 sets containing 6 images of gold particles on an amorphous germanium film were analysed. Each set begins and ends with an axial image and contains 4 orthogonally tilted images. The first set was captured at 2 Sch underfocus with a tilt magnitude of roughly 1.4 Gl/Sch (**tilts2**), the second at 1 Sch underfocus and 1 Gl/Sch (**tilts3**) and the third at 0.25 Sch underfocus and 0.5 Gl/Sch (**tilts4**). Set **tilts2** is shown in figure A.4.

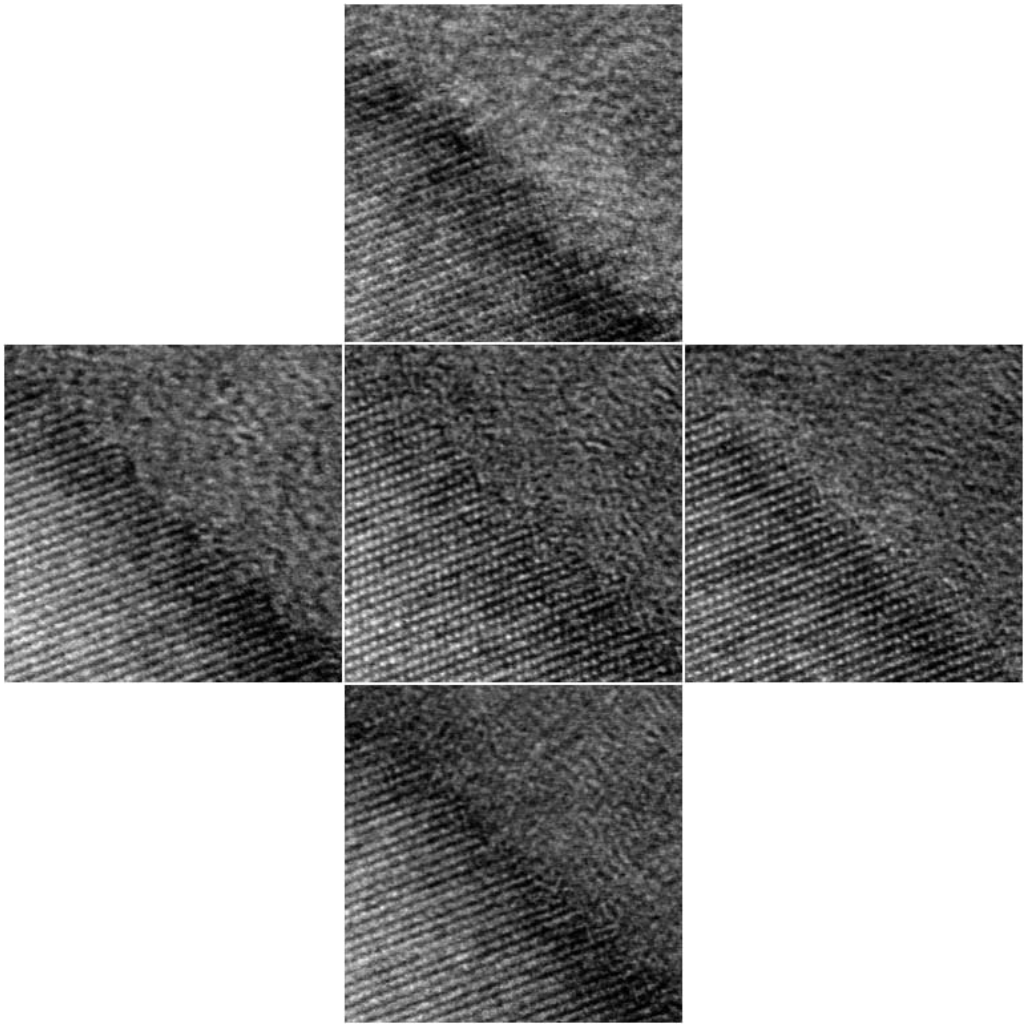


Figure A.3: Set **tilts1** : Tilt azimuth data set of the block oxide, $4\mathrm{Nb_2O_5}.9\mathrm{WO_3}$, captured on the JEOL 4000 EXII microscope at 400 kV and 1 Sch underfocus by Dr. A.I. Kirkland. Cs=0.9 mm and the field of view is 14.5 nm. The central image is the axial image and the tilt magnitudes are 1 GI/Sch with azimuths as shown in the figure.

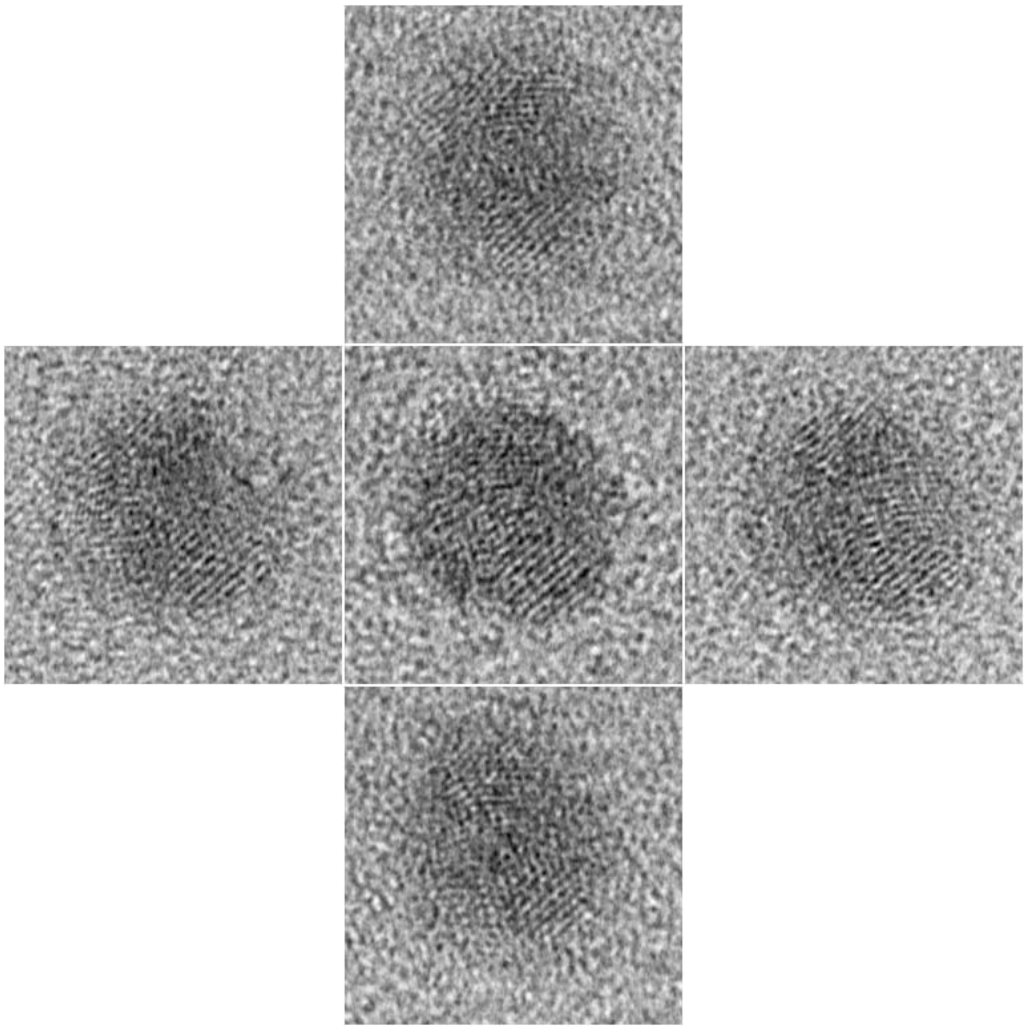


Figure A.4: Set **tilts2** : Tilt azimuth data set of a gold particle on an amorphous germanium film captured on the JEOL 2010F microscope at 200kV and 2 Sch underfocus by Dr. A.I. Kirkland in Japan with the assistance of Dr. K. Tsuno and Mr. M. Kawasaki of JEOL. Cs=1.1 mm and the field of view is 8.5 nm. The central image is the axial image and the tilt magnitudes are roughly 1.4 Gl/Sch with azimuths as shown in the figure.

Appendix B

Reconstruction Code

The following is a Semper run file listing for the tilt series reconstruction of set tilts2 (see Section 5.4.1.3). The `tctf` command is a comprehensive version of `ctf` which takes more aberrations into account.

```
! Reconstruction of set tilts2
!
! Assign tilts2 disk
assign name 'tilts2' wp noverify
ts=n
!
! size of images
sz=512
!
! number of images used
npics=6
!
! first and last images
img=100
n1=img+1
n2=img+npics
!
! Reconstruction, effective CTF, modulus and phase
R =n2+1
eT=n2+2
```

```
mR=n2+3
pR=n2+4
!
! Fourier transform of result, CTFs and predicted images
fR=n2+5
Tis=img+20
prd=img+30
!
! aberration values and imaging conditions
physical=yes
np=0.1
dd=0
!
kv=200
cs=1.1
step=0.0333
swidth=0.2
ewidth=4
vibration=0.0
!
a1r=-2.02 a1i=2.10
a2r=-0.150 a2i=-0.278
a3r=0 a3i=0
b3r=0 b3i=0
c3=cs
txx=1 txy=0 tyx=0 tyy=1
!
! x-tilt      y-tilt     defocus      x-pos      y-pos
tx1=-0.370  ty1= 0.358  d1=-110.09  mx1= 0.00  my1= 0.00
tx2=-9.946  ty2=-0.940  d2=-110.37  mx2= 4.00  my2= 2.00
tx3= 9.206  ty3= 1.655  d3=-110.64  mx3=17.33  my3=-1.67
tx4= 0.899  ty4=-9.400  d4=-110.92  mx4= 0.00  my4= 2.33
tx5=-1.639  ty5=10.115  d5=-111.19  mx5= 6.17  my5=-1.17
tx6=-0.370  ty6= 0.358  d6=-111.47  mx6=-7.50  my6=-3.00
!
! Extract images from tilt fisk with a common origin
for i 1,npics
```

```
  extract ts:i to (img+i) size sz position mx#i,my#i
  origin reset (img+i)
loop
deassign device ts
!
Ti=991 Tmi=992 S=993 Smi=994 CCC=995 Ii=996
!
create S size sz fp fourier value 0
copy to Smi
create CCC size sz complex fourier val 0
!
! S, Smi and CCC
for i=1,npics
  c1=dd+d#i mx=tx#i my=ty#i
  tctf wave to Ti siz sz
  extract to Tmi angle pi
  calculate :S+msq(:Ti)
  calculate :CCC+:Ti*:Tmi
loop
extract S to Smi angle pi
!
! Reconstruction
create R size sz complex fourier value 0
for i=1,npics
  c1=dd+d#i mx=tx#i my=ty#i
  tctf wave to Ti siz sz
  extract to Tmi angle pi
  copy Ti to Tis+i
  fit subtract divide n1+i-1 to Ii fp
  fourier
  fullplane
  calculate :R+((:Smi+np)*cc(:Ti)-cc(:CCC)*:Tmi)*:Ii
loop
calculate :R/((:Smi+np)*(:S+np)-msq(:CCC))
pixel 0,0=sz*sz
copy to fR
image R
```

```
!
! Modulus and phase of reconstruction
calculate modulus(:R) to mR
calculate phase(:R) to pR
!
! Effective CTF
calculate ((:Smi+np)*:S-msq(:CCC)) to eT
calculate :eT/((:Smi+np)*(:S+np)-msq(:CCC))
calculate modulus(:select)
!
! Predicted images
for i=1,npics
  calculate :fR*(:(Tis+1)) to :prd+i
  image
  calculate msq(:select)
loop
```

Appendix C

Parallel Programming

Most of the computation time in the least squares fit displacement method (see sections 3.6.2.2 and 4.6) is spent on the calculation of cross-correlation functions. The time taken by a typical 486 DX2-66 PC to perform a 256x256 XCF is approximately 11 seconds. When this is compared with the fraction of a second the HP 735 takes it is clear that the calculations are best done on a workstation. However, the price of framegrabbers and framestores for a workstation approaches that for a PC fitted with them so a flexible and cost-effective solution would be to use the PC to perform the image acquisition and storage and leave the computation to the workstation.

This appendix reprints a copy of a paper published in the proceedings of Symposium on Applied Computing 1994 (SAC '94) which describes the mechanisms involved in allowing a PC to capture images and a number of workstations to perform XCFs on them. Two strategies are covered: synchronous and asynchronous. In the synchronous method, the PC captures images and transfers them over the network to a single workstation. The PC waits until the workstation has performed the XCF and returned a peak height location before sending the next image. The asynchronous method allows the PC to capture images continuously and send them to a number of workstations which performs XCFs in parallel.

The asynchronous method allows for a maximum N-times speedup for N workstations although this is unlikely to be achieved in practice due to program execution and network overheads. The maximum number of workstations which can be used in parallel is related to the ratio of the time taken to calculate XCFs to the time taken to transfer the images over the

network. This ratio is proportional to the logarithm of image size so that more workstations can be used in parallel as the image size increases. As computer networks increase in bandwidth, the ratio will correspondingly increase. This implies that the method will become more attractive in the future when larger images are cross-correlated by workstations on a faster network.

A copy of my paper as published in the SAC '94 proceedings follows.

Appendix D

The Nested Correlation Function

Before the development of phase compensated XCF techniques, the Nested Correlation Function (NCF) was investigated as a means to improve the accuracy of electron microscope XCFs. This appendix reprints a copy of a paper published in Electronics Letters describing the function. The NCF is an *ad hoc* technique which is simple to implement: it involves weighting the original XCF transform with that of the reference image ACF. This weighting acts to accentuate features whose frequencies are present in the reference signal which makes the technique attractive in situations where geometrical distortion is present in one of the images being cross-correlated. This is the case in TEM images thanks to the frequency dependent image distortion indicated by equation (3.29).

The performance of the NCF is impressive in simulations of electron microscope images when compared with the XCF. However, when it comes to real image data, the results are disappointing. Investigation of the displacement method since the NCF was first tested points to an anomaly in image displacements, described in section 4.8, as a possible cause of the failure of the NCF to deliver accurate displacements.

A copy of my paper as published in Electronics Letters follows.

Appendix E

Reprints

This appendix reprints two 1994 Paris ICEM papers which I have co-authored (Saxton, W.O. et al., 1994; Kirkland, A.I. et al., 1994).