Spacing measurements of lattice fringes in HRTEM images using digital darkfield decomposition

Martin Rose

A Thesis Submitted to The Graduate School at
the University of Missouri - St. Louis in partial fulfillment of the
requirements for the degree Master of Science in Physics

July 2006

Advisory Committee:

P. Fraundorf, Ph.D.
Chairperson

B. Feldman, Ph.D.

D. Leopold, Ph.D.

Abstract

In this thesis I present the progress I made in understanding, improving and applying a digital method for lattice fringe spacing measurements in HRTEM images. The method uses filtering in Fourier space to form complex valued dark field images. These images are bright where the image contains the selected periodicity and dark otherwise. The amplitude and phase information in these dark field images can be used to analyse the distribution of periodicities and the change in lattice spacings. It turns out that the phase gradient can be interpreted as a position dependent wave vector. The method has been analyzed in detail in order to quantify the systematic error and the error introduced by noise. Windowing and Wiener filtering techniques have been applied to improve the measurement. A new approach to reduce the effect of noise uses multiple periodicities to determine the base vectors in every unit cell which leads to a reduction of noise. The method has been applied to simulated images with known noise levels to measure the error. Finally it has been applied to experimental images of heterostructures.

Contents

1 Introduction **3**

1.1 Motivation . . . 3
1.2 Crystals and transmission electron microscopy . . . 4
1.3 The basic assumption for image processing . . . 6
1.4 Different methods for spacing measurements . . . 7
1.5 The DF method . . . 9

2 Mathematical foundation of the DF method **11**

2.1 The 2D DFT . . . 11
2.2 Formation of a darkfield image . . . 12
2.3 Measurement of spacings from the DF image . . . 12
2.4 Simulation of TEM images . . . 14
2.5 Data windowing . . . 21
2.6 Error by changing wave vector . . . 24

3 The nature of noise **29**

3.1 A model for noise in a HRTEM image . . . 30
3.2 Estimation of the error in the phase of the DF image . . . 34
3.3 Measurement of the error in spacing measurements . . . 36
3.4 Removal of noise . . . 40
3.5 Multiple spot measurement . . . 42

4 Application of the DF method **44**

4.1 computer generated images . . . 44
4.2 experimental images . . . 46

5 Summary, conclusions and outlook **60**

6 Acknowledgment **62**

A Implementation of the DF method **63**

List of Figures **64**

References **65**

1 Introduction

This thesis presents the results of the work I have done at the Center for Molecular Electronics at the University of Missouri - St.Louis between November 2005 and June 2006 in the group of Prof. Dr. P. Fraundorf. I worked on a method called *digital darkfield imaging* or simply *DF method* which is used to analyse digitized high resolution transmission electron microscopy (HRTEM) images. The method uses Fourier space filtering to form so called darkfield images in direct space. The brightness and the phase of these complex valued images can be used to analyse periodicities and changes therein. The selection of frequencies in Fourier space used in this digital method is analogous to the selection of a beam of scattered electrons in a TEM in order to form darkfield images - therefore the name.

1.1 Motivation

Transmission electron microscopy provides us with high resolution TEM (HRTEM) images that show structures at the atomic scale. That means we are able to identify features on the size of 2Å $(2 \cdot 10^{-10}m)$. We can use this tool to take a look at the arrangement of atoms in a crystal. As the structure of crystals is known this provides no new information at first glance, but it gets very interesting when we examine interfaces between crystals or defects in a crystal.

An interface can be created for example by epitaxial growth where one substance is deposited ontop of another, layer by layer. The two substances must be of the same lattice type and have similar lattice spacings in order to fit. This layerwise deposition will result in a crystalline structure. Far away from the interface both substances will have the same properties, i.e. lattice parameters, as they have in a perfect crystal. But at the interface these parameters have to change, as it connects two similar but not identical lattices. The result is a deformation of the lattice called strain. Forms of strain are for example compression, expansion and shear. We also expect these changes to be very small so that they can not be seen in a digitized HRTEM image with the naked eye. Let's say we want to examine changes of about two percent in a 2.5Å spacing, which corresponds to 5 picometers (*pm*). This is way below the resolution of any microscope available! The heterostructures I mentioned before are the basic material in device manufacturing from LEDs and transistors to very large scale integrated circuits. As the properties of a semiconductor depend strongly on lattice parameters it is crucial to know them exactly.

Another important topic in material science are defects. They often occur at interfaces and change the electric properties locally. In some applications they are very usefull, for example they scatter electrons in a crystal and allow an electrical current to flow but in others they are just a disturbing side effect of the manufacturing process. In solid state physics defects can be treated as perturbations, so again the parameters to describe a defect are very important.

So how do we obtain the spacings at an atomic scale and more important how can we measure tiny changes of spacings if we are not working with a perfect

single crystal? One theoretical approach is elastic theory which calculates atom position from balancing forces. This method is not further considered in this thesis. An experimental approach is to extract this information from digitized HRTEM images. The characterization, improvement and application of the DF method is the main topic of this thesis. I will first describe the DF method mathematically, after that the effect of noise and other limiting factors will be considered. Finally the method will be applied to simulated and experimental HRTEM images of defect free structures, as the DF method turns out to be most precise in this case. I will focus on measuring changes in spacings of about $1 - 5\%$ in heterostructures.

1.2 Crystals and transmission electron microscopy

At this point I want to review the basics of crystallography, transmission electron microscopy and the process of image formation in a TEM. An introduction to crystallography and diffraction can be found in [2]. I found [9] to give the best introduction to transmission electron microscopy.

crystallography

The main property of crystals is the periodic, three dimensional arrangement of the atoms they are made of. Crystals are highly ordered compared to other states of matter like gases or fluids which contain only local order or none at all. Atoms in a crystal are located at points that can be described as a three dimensional grid called the Bravais lattice. This grid in turn can be constructed from three basis vectors. Due to periodic arrangement, some physical properties are not isotropic anymore, i.e. they depend on the orientation of the crystal. We therefore need a way to identify directions and planes in a crystal.

A group of three atoms that are not on one line define a plane. All parallel planes form a set of planes, defined by three numbers, the Miller indices. These indices are determined in the following way: The origin of the lattice must be put in a neighboring plane. The points where the plane intersects the coordinate axes must be determined. The inverse of these points must be multiplied in order to obtain integer numbers. If a plane does not intersect an axis the corresponding Miller index is zero.

Directions in a crystal are defined simply by the components of a vector pointing in the specific direction. The basis of this vector is the basis of the crystal.

diffraction of electrons

Bragg discovered in 1913 that x-rays scattered by crystals create characteristic patterns called diffraction patterns. This can be explained by interference of rays scattered by neighboring planes. Constructive interference occurs when the Bragg condition is fullfilled:

$$n \cdot \lambda = 2 \cdot d \cdot sin(\theta)$$

where λ is the wavelength of the x-rays, d the distance between the planes and θ the angle of incidence. n is an integer describing the order. The same experiment can be performed with electrons due to their wave nature. Usually, when a crystal is examined in a TEM the sample is rotated so that the Bragg condition is fullfilled so that bright and dark field techniques can be used.

harmonic analysis

The analysis of periodic structures can be summarized as harmonic analysis. Bragg's law connects scattering angles and spacings of lattice planes in a crystal. Therefore a diffraction pattern can be used to determine lattice plane spacings. This is an optical, analog method which uses frequency space data. Periodicities can also be analyzed starting from direct space data for example from a HRTEM image which is a projection of the crystal. A Fourier transform can be used to calculate the PS. This PS is very similar to the diffraction pattern as it contains spots in the same direction for all periodicities that are visible as fringes. We can also analyse periodicities locally using selected area electron diffraction (SAED). A future method intermediate between direct and frequency space might take advantage of wavelets. The windowed Fourier transform used in the DF method can be regarded as a simple wavelet. Using more advanced window functions it should be possible to analyse and locate periodicities more accurately. [8] focuses on this topic while [5] shows applications of wavelets in signal processing.

TEM and high resolution imaging

In a TEM a very thin specimen is irradiated by an electron beam. Magnetic lenses, solenoids, are used to manipulate the electron beam. The electrons travel along the symmetry axis of the lenses where an axial and transversal magnetic field is present. The axial component makes the electrons circle around the axis. This circular motion and the transversal component of the magnetic field are responsible for a force pointing towards the center of the lens. The point of focus can be controlled by the field strength, i.e by the electric current. The ability to focus electrons in a tiny area is the key to electron microscopy.

To understand the different imaging modes it is necessary to understand the imaging system of a TEM. It is assumed that the specimen is irradiated by an electron beam. The first lens below the specimen is called *objective lens*. This lens forms an intermediate image with a magnification of 20 to 50 times in it's *image plane*. Rays that leave the specimen at the same angle are collected in the same spot in the *back focal plane* of the objective lens. Either of these planes can be projected on the viewing screen by a second lens. To form a real image the image plane is projected, to examine the diffraction pattern the back focal plane is projected. The back focal plane is where the *objective aperture*, a small hole in a metallic plate, is located. This aperture allows single spots in the back focal plane to be selected for image formation. When the unscattered beam is selected the resulting image is called a bright field image, while a dark field image is formed by scattered electrons. When no object

aperture is used all electrons, scattered and unscattered, are used to form the image. These images usually have little contrast compared to dark or bright field images if non crystalline specimens are examined. In this case the contrast originates from inelastic scattering of the electrons by atoms in the specimen. This contrast mechanism is referred to as *mass-thickness* or simply Z contrast. When crystalline specimens are imaged the DF and BF images described above can be used because crystals scatter electrons as described by Bragg's law.

The back focal plane represents the Fourier transform of the distribution of atoms in the specimen, multiplied by the contrast transfer function. In general smooth functions have only few Fourier coefficients in a small range, while rapidly varying functions have Fourier components in a much wider range. The distribution of atoms is definitely a rapidly varying function. The small aperture used in DF or BF mode truncates the Fourier transform so that only coefficients in a small range can contribute to the image. Therefore these conventional imaging modes can not produce high resolution images. Even if no aperture would be used in order to include all Fourier coefficients the spacial resolution is limited by the phase distortion $W(k)$ which increases with k. The contrast transfer function mentioned before is defined by

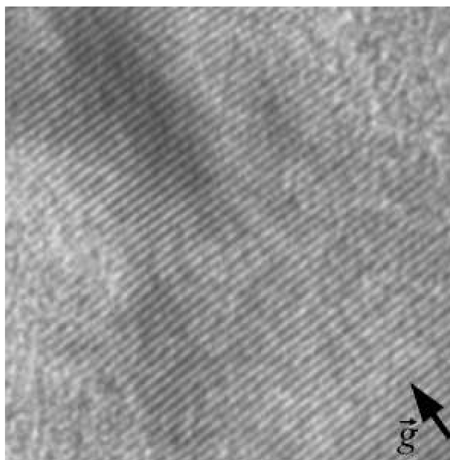
$$exp(iW(k))$$

As $W(k)$ increases rapidly with k high frequencies can not be used for image formation. This limits the resolution of a standard TEM to about 2Å.

To obtain high resolution images containing lattice fringes, the interference of the unscattered and the scattered beam is used. The electrons can be described as a wavefront. The phase of this wavefront is shifted when the electrons pass the specimen. The phase shift depends on the potential in the specimen. A quantum mechanical derivation of the image formation in the high resolution mode can be found in chapter 10 of [9].

1.3 The basic assumption for image processing

Before I will explain different methods to measure spacings between lattice fringes, I have to explain why spacings on an atomic scale can be deduced from HRTEM images. The image to the right is an example of the images we want to analyse. The bright stripes are called lattice fringes, they can appear in one or more directions. If the specimen is very thin, these fringes can be interpreted as the projection of tunnels between columns of atoms, while the dark lines are the atoms themselves.



This interpretation is true when the first zero in the contrast transfer function is at higher frequencies than the periodicity examined. When the specimen gets thicker a phase shift might occur and the meaning of dark and bright fringes is interchanged. This phase shift occurs when the first zero in the transfer function is at lower frequencies than the examined frequency. Article [11] analyses the extent to which the spacings between fringes in a HRTEM image correspond to atomic spacings. Lens transfer theory is used to give some practical rules for the application of the geometric phase method, which is similar to the DF method:

- The thickness of the specimen and the defocus should be chosen so that the lattice fringe contrast is maximal. This leads to lower error levels when the spacings are measured.
- Regions where fringe contrast changes rapidly should not be used.
- The analysis of spacings or displacements should be confirmed using different reflections, see multi spot method in chapter 3.5.
- The analysis should be carried out at different defocus values.

From this point on it is assumed that the fringes represent columns between atoms so that the spacings between lattice fringes correspond to atomic spacings. The unit for spacings it the pixel from now on. The scale factor between pixel spacings and spacings in nanometers has to be determined by the user. This is not that easy because the magnification is not known with high accuracy, but there are various methods to do that:

- The first possibility is to use spacing measurements in an undisturbed region of the image as reference, assuming the material and therefore the lattice spacings are known at this position.
- Another possibility is to deposit known structures on top of the specimen that can serve as a reference. For example a thin layer of polycrystaline aluminum can be used. Because of the random orientation of the single crystals this layer will create rings in the power spectrum corresponding to the spacings in the aluminum crystals. These rings define a scale from which spacings of other spots can be deduced.
- If the magnification is known within calibrated limits the scale factor can be calculated from the resolution of the scanner and the magnification. This method often yields larger absolute spacing uncertainties.

1.4 Different methods for spacing measurements

Now that the connection between the images and the real crystal has been made we can focus on methods to measure the spacings between lattice fringes.

An early reference - a usermanual of an image processing unit - that describes a method involving Fourier transforms goes back to 1990 [17]. In general all

methods for spacing measurements can be divided in two groups: those working in direct space only and those interpreting Fourier space data.

The first direct space method to measure the spacings between lattice fringes that comes to mind is simply using some kind of ruler to measure spacings between adjacent fringes. In the image above the spacings are about 6 pixels. With this simple method the spacings can be determined with an accuracy of about 0.5 pixel. This can be improved when the measurement is averaged over more fringes, but at the same time this averaging process limits the spacial resolution so that spacing changes close to a sharp interface can not be resolved.

A more sophisticated direct space method has been described in [4]. This method uses the occurrence of Moiré structures when two patterns are superimposed. First of all a reference lattice is calculated from an undisturbed area of the image. This pattern describes the position of bright spots in this area. Then this reference pattern is extrapolated over the whole image. Now the separation between the reference lattice points and the closest experimental point is measured. This method has been implemented in a software package called LADIA, see [7]. These displacement measurements are the basis for chemical mapping and strain analysis.

Fourier space methods analyse the power spectrum and the Fourier coefficients of a HRTEM image. A perfect unstrained lattice gives rise to very sharply peaked spots in the PS. The spacing and angle of lattice fringes can be calculated from the position of the corresponding spot in the power spectrum, as shown in [6]. In this article the spacings between fringes in TiO_2 and TiN particles in a catalyst are measured. As the spacing is constant within a particle the spot in the PS has very little structure, it's nearly a point. By interpolation the exact center of a spot can be found. This method can be used to measure constant spacings with an accuracy of 0.001-0.05Å depending on the specimen and imaging conditions. However this method can not spatially resolve tiny changes in spacings. Changes in spacings and the shape of the region containing periodicities change the profile of the corresponding peak in the PS, it gets wider and smoother.

Another well established Fourier space method is the *geometric phase technique*, described in [12] and applied in [15] and [10]. It uses filtering in Fourier space to calculate a DF image. The phase $P_g(\vec{r})$ of the DF image is related to the displacement field $\vec{u}$ by:

$$P_g(\vec{r}) = 2\pi \cdot \vec{g} \cdot \vec{u}$$

The phases $P_{g1}(\vec{r})$ and $P_{g2}(\vec{r})$ of DF images corresponding to two different periodicities at position $\vec{g}_1$ and $\vec{g}_2$ are used to calculate the displacement field:

$$\vec{u} = -\frac{1}{2\pi}\left(P_{g1}(\vec{r})\vec{a}_1 + P_{g2}(\vec{r})\vec{a}_2\right)$$

where $\vec{a}_1$ and $\vec{a}_2$ are the corresponding direct space vectors to $\vec{g}_1$ and $\vec{g}_2$. This method can be used to measure the fringe position with an accuracy of 1% of the spacing.

1.5 The DF method

The DF method is similar to the geometric phase technique as it interprets the phase of a DF image. The DF image is intermediate between direct and reciprocal space as it contains localized information about periodicities, see [8]. Using the phase gradient it measures the local spacing between lattice fringes. The amplitude of the DF image can be used to locate areas containing periodicities. From these fundamental measurements other quantities like the displacement field and the strain tensor can be derived. To do that, more than one set of fringes is required of course. Figure 1 shows the steps of the DF method in detail.

First of all the input image is loaded, shifted and scaled. The shifting is done to decrease the magnitude of the DC peak which leeds to a better visibility of the periodicities in the PS. Scaling is not necessary but it increases the contrast when the input is saved again. These operations have no influence on the measurements, see chapter 3. A window function is applied to the input in order to decrease the effect of leaking, see chapter 2.5. Now the Fourier transform is calculated using the FFT algorithm. The power spectrum can be examined to determine the SNR which is a measure for the quality of the image.

For a detailed description of the noise in a TEM image and the effect it has on the spacing measurements see chapter 3. Once the noise model has been established, a Wiener filter can be applied in order to reduce the noise level. Noise can also be added when simulated images are used in order to understand the effect of noise.

A circular aperture is then centered around the brightest point of a spot in order to select a periodicity i.e. a set of lattice fringes. The effect of the choice of center and radius is discussed in chapter 4. Another window is applied to the aperture. In a method previously used the Fourier coefficients were shifted so that the center of the aperture become the DC peak. This step is not used here. Now the DF image is calculated using the inverse Fourier transform. From the DF image the spacings between fringes can be measured by the phase gradient, see chapter 2.3. The amplitude of the DF image can be used to locate areas containing the periodicity selected by the aperture. Finally measurements from multiple DF images can be combined in order to increase the accuracy of the measurements without loosing spacial resolution. This multi spot method is described in chapter 3.5.

Direct space

Fourier space

Load image – size NxN
- shift and scale data for better visualisation
- apply window

compute power spectrum unsing FFT

- Select aperture – center and radius
- add or remove noise for simulation or to improve SNR
- analyse azimuthal intensity in PS
- measure signal and noise levels
- apply window to aperture

compute dark field image using iFFT

Analyse dark field image
- calculate phase gradient for spacing measuremtns
- analyse DF brightness to localize periodicities
- combine multiple DF images to improve accuracy

Figure 1: single steps of the DF method

2 Mathematical foundation of the DF method

This chapter outlines the mathematics behind the DF method and the simulation of HRTEM images. The unit for spacial coordinates is one *pixel*. All pixel values must be interpreted to connect them to physical distances which is done when the method is applied in chapter 4. There are some parameters used to define the input and to control special functions. Some of these parameters have already been mentioned in the description of the method and will be defined in this mathematical context. Appendix A lists all available parameters and the functions they control. First of all an image of size $N \times N$ containing lattice fringes is required as input. The first parameter is therefore N, a power of 2, usually 512, 1024 or 2048 in order to take advantage of the FFT algorithm.

2.1 The 2D DFT

The two dimensional discrete Fourier transform

$$\hat{f}(h_x, h_y) = \sum_{x=0,y=0}^{N-1} exp\left(2\pi i \cdot \frac{(h_x x + h_y y)}{N}\right) \cdot f(x, y)$$

and its inverse

$$f(x, y) = \frac{1}{N^2} \sum_{h_x=0,h_y=0}^{N-1} exp\left(-2\pi i \cdot \frac{(h_x x + h_y y)}{N}\right) \cdot \hat{f}(h_x, h_y)$$

as defined in [16] is used to translate between direct and reciprocal space. The unit of the Fourier coefficients $\hat{F}(h_x, h_y)$ is $pixel^{-1}$. With the help of these two quantities

$$\vec{k} \equiv 2\pi \cdot \begin{pmatrix} h_x \\ h_y \end{pmatrix} \qquad \text{and} \qquad \vec{r} \equiv \frac{1}{N} \cdot \begin{pmatrix} x \\ y \end{pmatrix}$$

the inverse transform can be rewritten:

$$f(\vec{r}) = \frac{1}{N^2} \sum_{\vec{k}} exp\left(-i \cdot \vec{k}\vec{r}\right) \cdot \hat{f}(\vec{k})$$

where $\vec{k}$ is the discrete wave vector and each $\hat{f}(\vec{k})$ corresponds to one $\hat{f}(h_x, h_y)$. The components of $\vec{r}$ (r_x and r_y) vary between 0 and 1. This form of the inverse transform makes it obvious that the image in direct space can be represented as a sum of complex valued plane waves.

As the input is real the Fourier coefficients must have conjugate symmetry when the Fourier transform (FT) of the input is calculated: $\hat{f}(\vec{k}) = \hat{f}^*(-\vec{k})$. This symmetry is the reason for the symmetry of the power spectrum (PS) with respect to the DC peak. The value of the PS at a given point is defined in a not so common way as the absolute value of the Fourier coefficient at this point.

2.2 Formation of a darkfield image

In transmission electron microscopy a darkfield image (DF) is formed by allowing only electrons that were scattered at a certain angle to form the image according to chapter 2.3 in [9]. A circular mechanical aperture made out of metal called the objective aperture is positioned in the back focal plane to select electrons that were scattered into the same angle. When scattered electrons are allowed to pass the aperture we call the resulting image a dark-field image. The analogous process in the DF method is filtering in Fourier space where the DC peak corresponds to the unscattered electrons and the spots correspond to diffraction spots i.e. the scattered electrons. This is just an analogy, the PS of a HRTEM image should not be confused with the diffraction pattern we can observe in the TEM! To apply an aperture one point in Fourier space is chosen as the center of the aperture. All Fourier coefficients outside a circle with radius r_{AP} are set to zero. Inside the aperture a Hamm window is applied. The choice of the center point and the effect of the radius will be discussed in 4.1. This filtering process in Fourier space destroys the conjugate symmetry of the Fourier coefficients. The result of the inverse transform of these asymmetric coefficients is called the darkfield in close analogy to optical darkfield images. Regions of the specimen that contain no periodicities will look dark if the DC peak is outside the aperture while regions containing periodicities light up. The value of the DF image $DF(\vec{r})$ at each point is given by the superposition of the plane waves corresponding to the Fourier coefficients inside the aperture as described in the previous section.

$$DF(\vec{r}) = \frac{1}{N^2} \sum_{\vec{k}} exp\left(-i \cdot \vec{k}\vec{r}\right) \cdot \hat{f}(\vec{k}) = A_{Res}(\vec{r}) \cdot exp\left(-i \cdot \Phi_{Res}(\vec{r})\right)$$

where $A_{Res}(\vec{r})$ is the resulting amplitude and $\Phi_{Res}(\vec{r})$ is the resulting phase at each point.

2.3 Measurement of spacings from the DF image

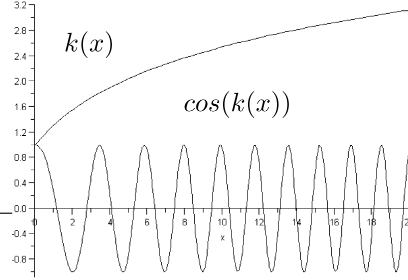
Let's assume for a moment that we would also use the Fourier coefficients at $-\vec{k}$ for each $\vec{k}$ inside the aperture to form the DF image. In this case it would be a real image showing the periodicities chosen by the aperture:

$$DF_R(\vec{r}) = A_{Res}(\vec{r}) \cdot cos\left(\Phi_{Res}(\vec{r})\right)$$

$\Phi_{Res}(\vec{r})$ can be written in the following way:

$$\Phi_{Res}(\vec{r}) = \vec{k}_{Res}(\vec{r}) \cdot \vec{r}$$

So the DF image can be interpreted as a plane wave with position dependent wave vector $\vec{k}_{Res}(\vec{r})$. The figure to the right shows the function $cos(k(x) \cdot x)$. The upper graph is $k(x)$, the wave vector which increases logarithmically with x. The effect of the position dependent wave vector can be seen best at the left side of the plot.



Now it is assumed that the wave vector varies only slightly within one period so that the spacing d between two maxima is given by

$$d(\vec{r}) = \frac{2\pi}{\left|\vec{k}_{Res}(\vec{r})\right|}$$

The distance d is about six pixels in the simulated images so a measurement of $\vec{k}$ within this distance is assumed to give the same value for d.

However the phase $\Phi_{Res}(\vec{r})$ can not be measured in the real DF image with high accuracy, because it's difficult to determine the intensity maxima with high accuracy. The same problem sets limits for the simple ruler method described in chapter 1.4. Therefore the complex DF image described in the section before is used:

$$DF_C(\vec{r}) = A_{Res}(\vec{r}) \cdot \{cos\left(\Phi_{Res}(\vec{r})\right) + i \cdot sin\left(\Phi_{Res}(\vec{r})\right)\}$$

Now the phase is given at each point by

$$\Phi_{Res}(\vec{r}) = \vec{k}_{Res}(\vec{r}) \cdot \vec{r} \equiv arctan \frac{\Im(DF_C(\vec{r}))}{\Re(DF_C(\vec{r}))}$$

From this the gradient of $\Phi_{Res}(\vec{r})$ can be calculated

$$\begin{aligned} \vec{\nabla}\left(\vec{k} \cdot \vec{r}\right) &= \left(\vec{k} \cdot \vec{\nabla}\right)\vec{r} + \left(\vec{r} \cdot \vec{\nabla}\right)\vec{k} \qquad \text{where} \qquad \vec{k} \equiv \vec{k}_{Res}(\vec{r}) \\ &= (k_x\partial_x + k_y\partial_y)\vec{r} + (r_x\partial_x + r_y\partial_y)\vec{k} \\ &= k_x\hat{x} + k_y\hat{y} + (r_x\partial_x + r_y\partial_y) \cdot (k_x\hat{x} + k_y\hat{y}) \\ &= \vec{k} + \hat{x} \cdot (r_x\partial_x k_x + r_y\partial_y k_x) + \hat{y} \cdot (r_x\partial_x k_y + r_y\partial_y k_y) \end{aligned}$$

Making the same assumption as before, namely that $\vec{k}_{Res}(\vec{r})$ varies only slightly with position the gradient of the phase is approximately $\vec{k}_{Res}(\vec{r})$. Therefore the local spacing d can be calculated as a first approximation from the phase gradient by

$$d(\vec{r}) = \frac{2\pi}{\left|\vec{\nabla}\Phi_{Res}(\vec{r})\right|}$$

This assumption will be checked and the error in the spacing measurements due to the change of $\vec{k}$ will be examined in section 2.6.

The magnitude of a value in the DF image, $|DF_C(\vec{r})|$, can be interpreted as well in close analogy to the DF image in transmission electron microscopy. In a TEM DF image areas of the specimen which scatter electrons appear bright while those areas which do not, appear dark. For example in a polycrystaline sample this contrast mechanism would reveal crystals in the same orientation. In the DF method this value is used as a binary mask, because the value is high in areas with low noise levels and clear periodicities while it tends to zero in areas where the specimen is amorphous and no periodicities are present.

The DF method is a tool for measuring spacings between lattice fringes. The measurement of distances is the most basic operation in geometry. From distances between atoms more complicated quantities like the displacement field $\vec{u}$ and the strain tensor, which is composed of the partial derivatives of $\vec{u}$, can be calculated. The stress tensor is connected to the strain tensor by material relations. I will therefore describe a way to obtain the displacement field from the spacing measurements, as it makes the connection to other methods like the one used by Hytch et al. (see section 1.4).

Lets assume we want to determine the displacement field for a structure as shown in figure 21 where the dark dots correspond to columns between atoms. The DF methods provides the wave vectors at some spot that can be choosen arbitrarily, for example in an area without distortions. From that we can compute the positions of the neigbouring atoms or tunnels respectively. The next step is to measure the wave vectors at the new positions in order to repeat the previous step. This procedure yields the positions $\vec{a_i}$ of all the spots in the HRTEM image. As the displacement field describes the distance from some reference position we need to introduce a reference lattice. This reference lattice can be calculated for example from the positions of spots in an undistorted area.

2.4 Simulation of TEM images

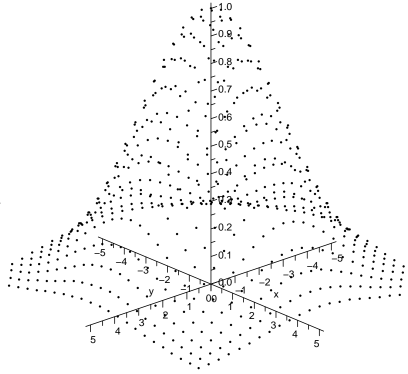
A basic method in crystallography and material science is to compare experimental HRTEM images to computer simulated images in order to check the atomic models. A major part of my work was to apply the DF method to simulated images in order to see how accurate it works under perfect conditions and to determine the error in the measurements obtained by the DF method. Chapter 10.4 in [9] gives an introduction to a simulation method called multislice method. The input to any simulation is the position of atoms inside the region of interest. From this the electron wavefunction at the specimen exit surface, at the back focal plane and at the image plane can be calculated. The second step is to modify the calculated wave function at the various positions in order to simulate the characteristics and deficiencies of the microscope e.g. spherical aberration, astigmatism and defocus.

I used a different approach to generate test images similar to high resolution TEM images. In the introduction of [18] it is mentioned that the tunnels between columns of atoms in a periodic structure appear as nearly gaussian shaped dots in a HRTEM image. The dots are either black or white depending on the

thickness of the specimen and the focus. The distance between the centers of these dots corresponds to the lattice spacing as described in chapter 1.3. The profile of a lattice fringe can be examined using the 3D visualisation plugin in ImageJ and the shape of a gaussian is revealed. Therefore Gauss profiles at positions given by a 2D Bravais lattice are used to simulate HRTEM images. Images like these can be obtained from a microscope with Gaussian point spread but no spherical aberration. A newly implemented technique of aberration correction might make such microscopes available, see [3]. This simulation does not take deficiencies of the microscope into account. The brightness at a position x/y in the micrograph due to a tunnel at position p/q is therefore given by

$$exp\left(-c \cdot ((x-p)^2 + (y-q)^2)\right)$$

The image to the right shows a Gauss profile. The constant c is chosen to be 0.5. As the image is made out of pixels at discrete integer positions it is necessary to integrate the intensity that contributes to one pixel over the area of the pixel. The pixels are aligned to the micrograph in such a way that the center of a pixel corresponds to an integer coordinate in the micrograph. The brightness B_{xy} of a pixel is then obtained by integrating the Gauss profile over the area of one pixel.



$$\begin{aligned} B_{xy} &= \int_{x-0.5}^{x+0.5} \int_{y-0.5}^{y+0.5} exp\left(-c \cdot ((x-p)^2 + (y-q)^2)\right) \\ &= \frac{\pi}{4c} \cdot \left\{ \Gamma\left(\sqrt{c} \cdot (x - 0.5 - p)\right) - \Gamma\left(-\sqrt{c} \cdot (x + 0.5 - p)\right) \right\} \\ &\quad \cdot \left\{ \Gamma\left(\sqrt{c} \cdot (y - 0.5 - q)\right) - \Gamma\left(-\sqrt{c} \cdot (y + 0.5 - q)\right) \right\} \end{aligned}$$

where $$\Gamma = \frac{2}{\sqrt{\pi}} \int_0^x \exp\left(-\frac{t^2}{2}\right) dt$$

Compared to a method previously used to simulate images this approach allows us to put tunnels at arbitrary positions. The grain size of the particles on the film which detect the electrons is very small and they are randomly distributed. Therefore it is assumed that the film can detect electrons continously in space without introducing any grid. The brightness is proportional to the number of electrons hitting the film at a certain position. The integration over the area of a pixel in the simulated micrograph corresponds to the process of digitalization in the scanner. As the same profile is used for all tunnels the integration ensures that the brightness of all spots is equal. In the microscope this means that the same number of electrons was detected in one spot.

Figure 2 shows the basic geometry of the simulated images. The Bravais lattice with the basis vectors $\vec{a}$ and $\vec{b}$ defines the position where the Gauss profiles

will be centered. This lattice contains a number of periodicities running in different directions as indicated by the various pairs of parallel lines. The spacing marked by d will be examined when the DF method is applied to simulated images. The following basis is used:

$$\vec{a} = \frac{6}{109}\begin{pmatrix} 10 \\ 3 \end{pmatrix} \qquad \text{and} \qquad \vec{b} = \frac{7}{37}\begin{pmatrix} 1 \\ 6 \end{pmatrix}$$

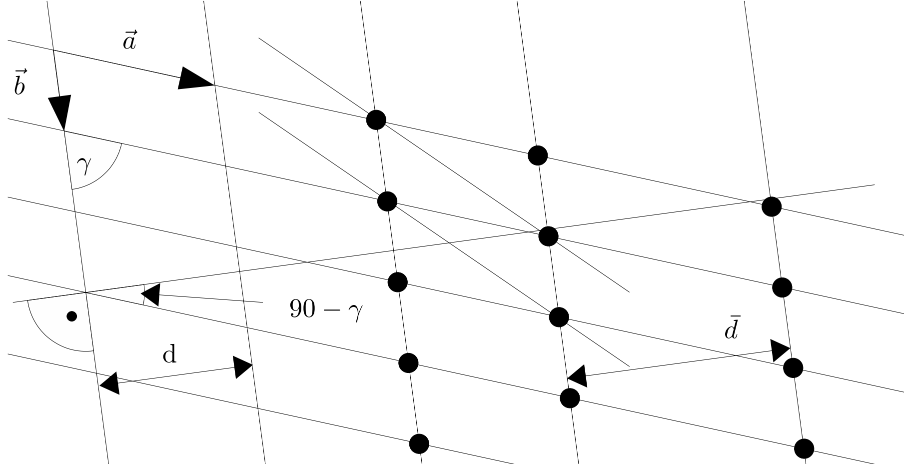


Figure 2: Bravais lattice and the examined spacings

The spacing d is calculated as follows:

$$d = a \cdot \cos(90 - \gamma) = 5.385326 \qquad \text{and} \qquad \gamma = arcos\left(\frac{\vec{a}\vec{b}}{|\vec{a}| \cdot \left|\vec{b}\right|}\right)$$

To introduce a change in this spacing the base vector $\vec{a}$ is multiplied by a scale factor f, see for example $\bar{d}$.

The simulated images do not obey the periodic boundary condition assumed by the DFT. The image can be regarded as a square window function multiplied by an infinite image. The PS is therefore a convolutution of the FTs of the 2D square window and the image. Lines running perpendicular to the boundary of the image appear in the PS. Because of this noise in the PS the examined periodicity should not run perpendicular to a boundary.

There are two structures which will be used to test the DF method. Both structures start from the lattice described above. They differ in the way the spacings are changed.

Figure 3 and 4 show the spacing between a certain set of parallel rows of Gaussian peaks. The spacing is defined as the distance between two neighboring rows. The question that arises is where to measure a spacing, or which point

is the spacing assigned to. By definition the spacing does not change between two peaks. Therefore the spacing is assigned to the center of the peak as it is located in the middle of columns of atoms. The spacings are an output of the program which simulates the images, and they will be used in chapter 4.1 as a reference for the spacing measurements.

Sharp step Although the abrupt change of the spacing in this geometry is somehow unphysical, it is very useful to understand the effect of noise which is often abundant in experimental images.

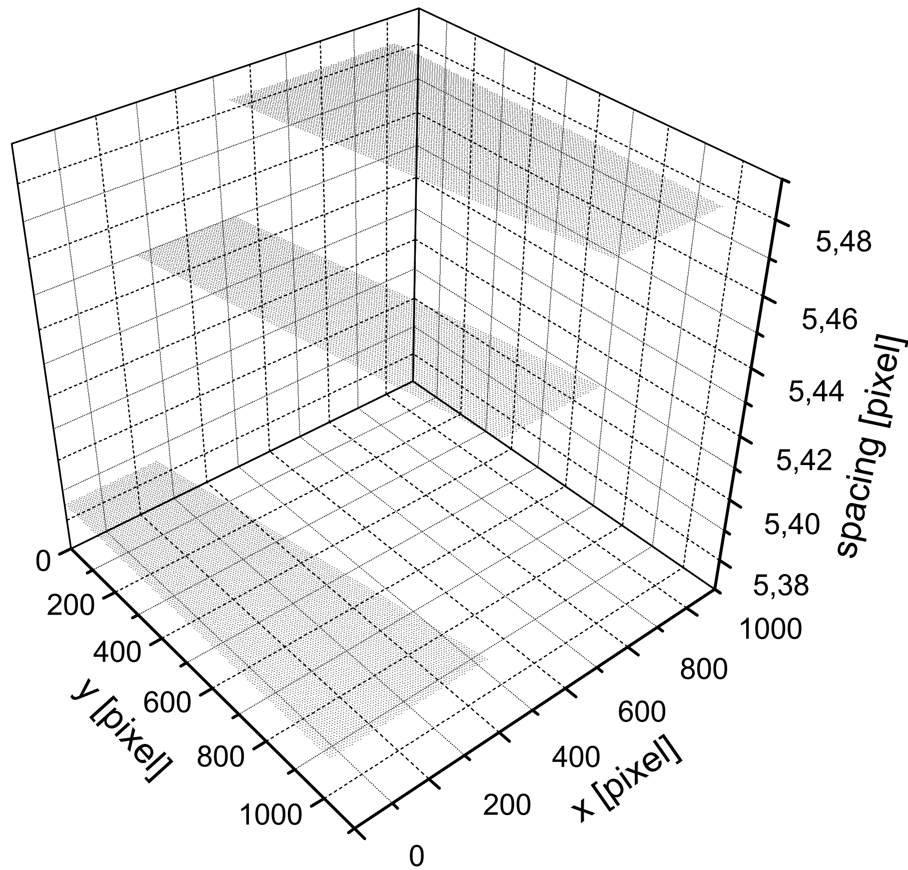


Figure 3: local spacings for the sharp step geometry

The geometry is divided into three regions of equal spacings: in the lower left the spacing is that of the starting geometry $d_0 = 5.385326$. At the first step the spacing is multiplied by 1.012 leading to $d_1 = 5.44995$. At the second step the spacing is multiplied by 1.006 resulting in a spacing of $d_3 = 5.48265$. The lines which divide areas of constant spacing intersect the x-axis at $x_1 = 259$ and

$x_2 = 535$.

Wide step - interface This geometry is a model for a homogeneous interface - a change in spacings over a wider range.

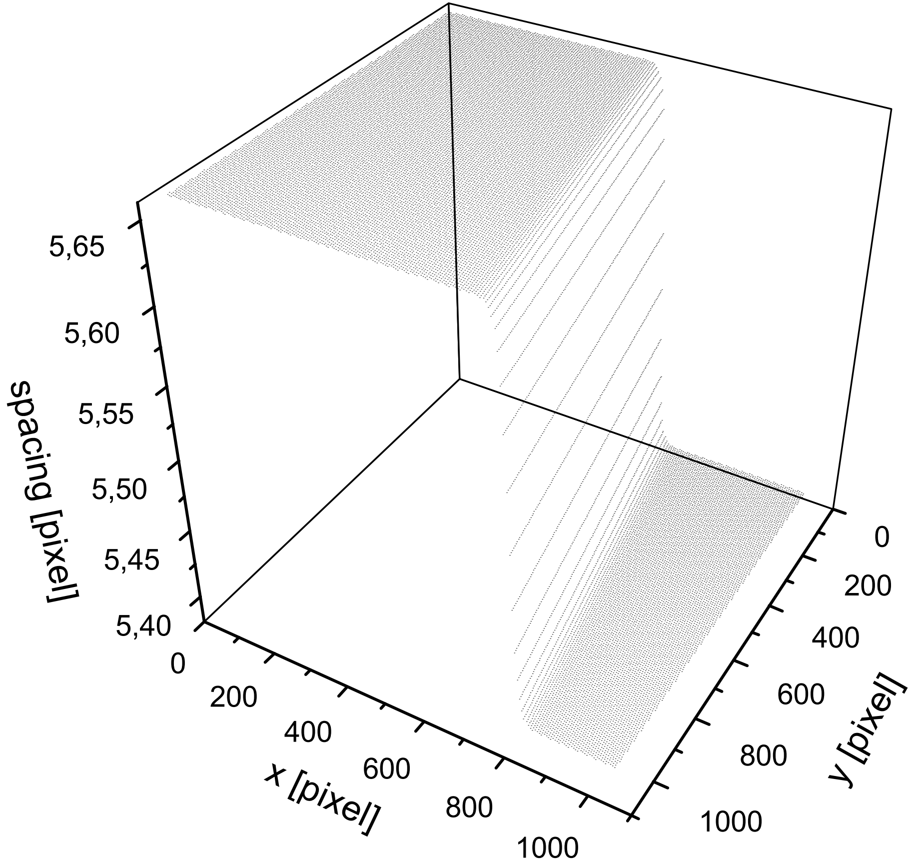


Figure 4: local spacings for the wide step geometry

The scale factor is a steady function similar to the Fermi Dirac distribution:

$$f(x) = 1 + \frac{a}{1 + \exp(\frac{x - x_0}{w})}$$

where a defines the height of the step and w determines the characteristic length on which the spacing changes. The point where the spacing has changed by $a/2$ is set by x_0. The width w_s on which 90% of the spacing change happens can be calculated to be

$$w_s = 2 \cdot ln(19) \cdot w$$

This enables us to connect the parameter w to a width w_s in pixels. The maximum slope is $-a/4w$ at $x = x_0$.

With this geometry it will be possible to measure the error introduced by a changing wave vector, as described in chapter 2.6. We will be able to set limits for the accuracy of the method depending on the geometry, in other words we can tell how abrupt and how big the change in a spacing can be, in order to be measured with a certain accuracy. This way the two major sources of error - noise and a changing wave vector - can be compared.

The power spectrum As both geometries are very similar they have nearly equal power spectra. The left of figure 5 shows the power spectrum of the step geometry in linear scale, where the color has been inverted for better visualisation. The right shows the same PS in logarithmic scale. The PS is symmetric to the DC peak, therefore the part of the PS that is shown contains all the information.

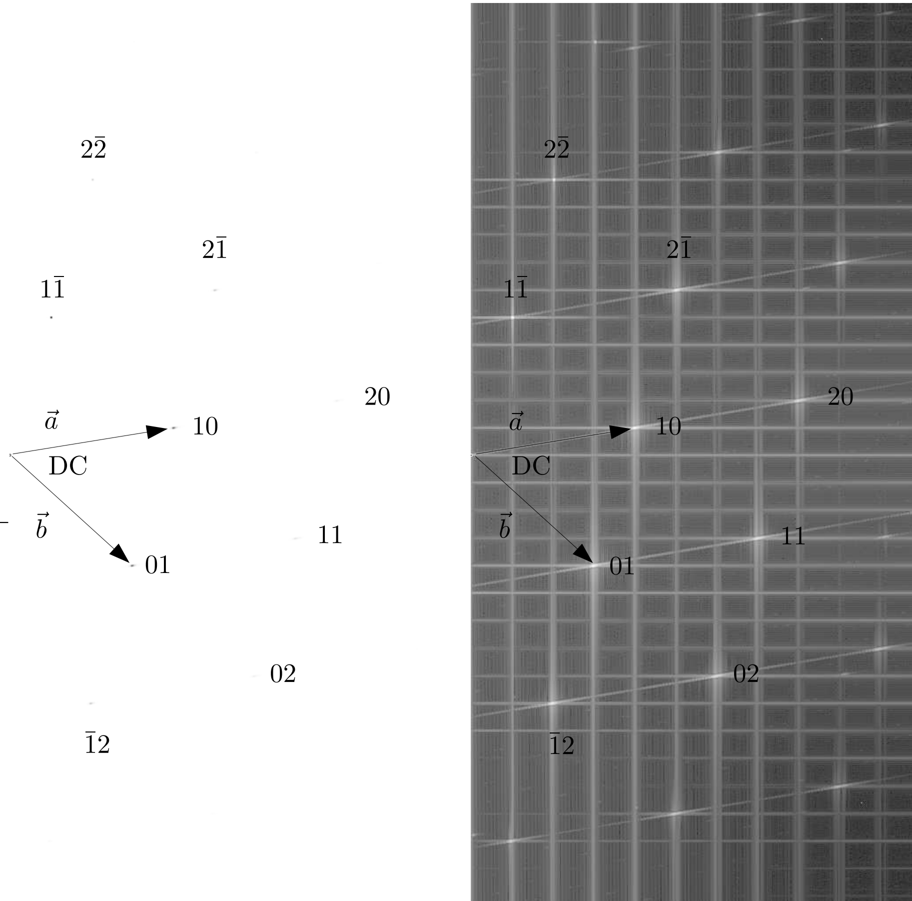


Figure 5: PS of the step geometry, right log scale, left lin scale

The spots in the PS appear very faint in the linearly scaled image, it is usefull to identify the strongest periodicities. In the logarithmically scaled image the weaker periodicities but also the artifacts running perpendicular to the edges of the image become visible. The spots have been indexed using the arbitrarily chosen base vectors $\vec{a}$ and $\vec{b}$. The spots that are hardly visible in the left image can be found easily in the right image.

2.5 Data windowing

Data windowing is a mathematical method used to improve the accuracy of the DFT. Any signal is allways sampled on a limited interval for example in time or space. The process of measuring can be thought of as taking a look at the signal through a window which is opened when we start to measure and closed when the measurement is finished. In mathematical terms this means that the measured data is equal to the signal multiplied with a window function which varies between zero and one while the measurement is in progress. Windowing is allways present even if the user of the DFT is not aware of it: The square windows multiplies every data point by one, it's the simplest window and gives the poorest accuracy of the PS. The accuracy of the PS is limited by a process called leakage which is described in detail in chapter 13.4 in [16]. Leakage is the influence on a Fourier coefficient by nearby coefficients. High leakage means coefficients far away in the spectrum contribute to a coefficient, while low leakage means that only coefficients close by have an influence. Leakage can be minimized by data windowing i.e. by multiplying the input with a window function as described in [16] where some 1D window functions are defined. The 2D window function at position (x,y) is the product of two 1D functions, one evaluated at x the other one at y. The effect of the following window functions on the x component of the phase gradient has been compared:

rectangle $w_{xy} = w_x \cdot w_y = 1$

Bartlett $w_{xy} = \left(1 - \left|\left(x - \frac{N}{2}\right)\frac{2}{N}\right|\right) \cdot \left(1 - \left|\left(y - \frac{N}{2}\right)\frac{2}{N}\right|\right)$

Welch $w_{xy} = \left(1 - \left(\left(x - \frac{N}{2}\right)\frac{2}{N}\right)^2\right) \cdot \left(1 - \left(\left(y - \frac{N}{2}\right)\frac{2}{N}\right)^2\right)$

Hann $w_{xy} = \left(0.5 - 0.5 \cdot cos\left(\frac{2\pi x}{N}\right)\right) \cdot \left(0.5 - 0.5 \cdot cos\left(\frac{2\pi y}{N}\right)\right)$

Hamming $w_{xy} = \left(0.54 - 0.46 \cdot cos\left(\frac{2\pi x}{N}\right)\right) \cdot \left(0.54 - 0.46 \cdot cos\left(\frac{2\pi y}{N}\right)\right)$

The indices x and y run from 0 to $N - 1$. Figure 6 shows the different window functions.

Figure 7 and Figure 8 show a step in the x-component of the phase gradient with different windows applied to the input and to the aperture. The naming convention in the legend is the following: the window applied to the input in direct space is on the left while the window applied to the aperture in Fourier space is on the right.

Figure 7 clearly shows the impact of windowing on the phase gradient. Using rectangular windows the phase gradient is not very smooth. The windows that are shown in this figure have a very similar effect on the phase gradient. The main difference is at the border of the image where different windows produce different oscillations. This region is not shown here as it is not important for the

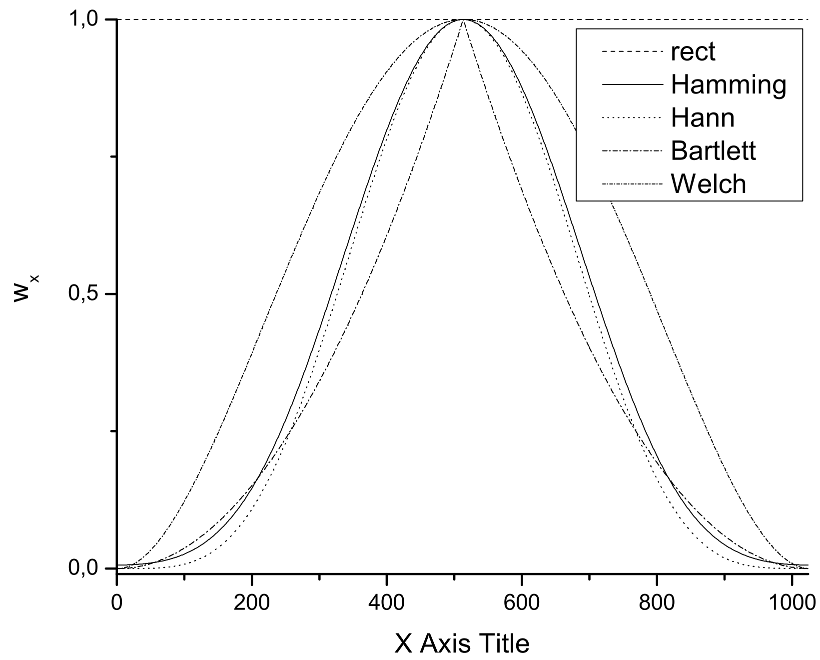


Figure 6: the various 1D window functions. (N=1024)

measurements. With the information in this figure I decide to use a Hamming window as the standard window for the input. Figure 8 illustrates the changes when a window is applied to the aperture. As a combination of Hamming window for the input and a Hann window for the aperture results in the smoothest curve and the lowest oszillation at the border, this combination of windows is used by default. Also note how the slope of the graph gets smaller when using windows. This effect can be compensated by a bigger aperture.

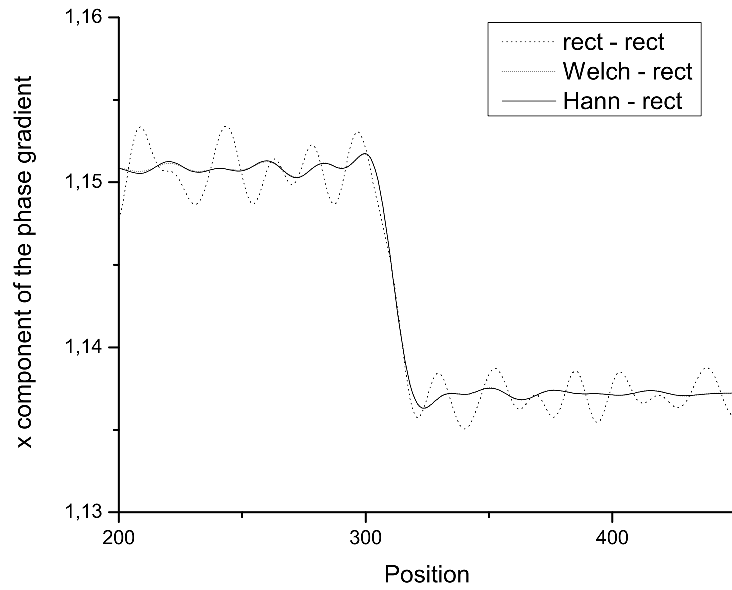


Figure 7: different window functions for the input, no window on aperture. (N=1024)

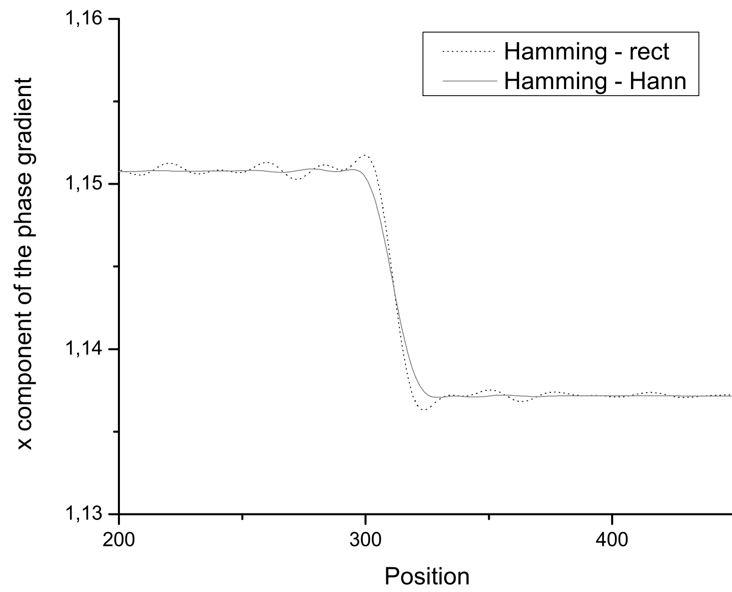


Figure 8: different windows applied to the aperture, Hamming window for the input. (N=1024)

2.6 Error by changing wave vector

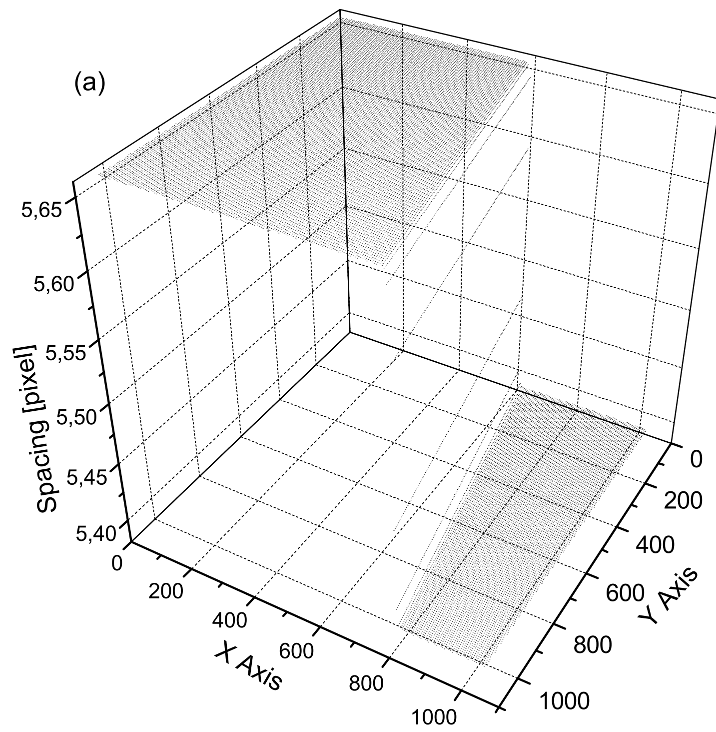
Now that all the parts of the DF method are understood the error inherent in the method will be examined. In section 2.3 we assumed that the phase gradient is the wave vector if the wave vector changes only slightly with position. We will check this assumption and quantify the introduced error by applying the DF method to computer generated images and comparing the obtained spacings to the correct values. The wide step geometry described in 2.4 is used because it allows us to generate smoothly changing spacings. The geometry is described by two parameters, the height of the step and its width. Figure 9 shows two geometries with extreme parameters, both at the same scale: (a) is a narrow ($w = 3$) 5% step while (b) is a wide ($w = 100$) 1% step.

The maximum of the relative error will be measured for different combinations of the parameters. The height of the step varies between 1 and 5 percent while w varies between 60 and 1. Figure 10 and 11 show examples of such measurements, the maximum error occurs where the spacing changes most. The figures in this section show the projection of the datapoints onto the x-z plane to make it easier for the reader to see the 3 dimensional distribution.

Table 1 shows the results of the error measurements. Each row contains the relative error measured in ‰ for a certain step height. Different columns represent different values of w, the parameter for the width.

	60	40	20	10	5	3	1
1%	0.16	0.16	0.42	1.06	2.09	3.01	4.08
2%	0.24	0.37	0.87	2.17	4.42	5.67	8.55
3%	0.37	0.57	1.33	3.36	6.53	9.24	12.08
4%	0.51	0.78	1.81	4.40	8.82	11.82	16.56
5%	0.64	0.98	2.27	5.53	11.23	15.27	22.25

Table 1: relative error [‰] in spacing measurements in simulated images



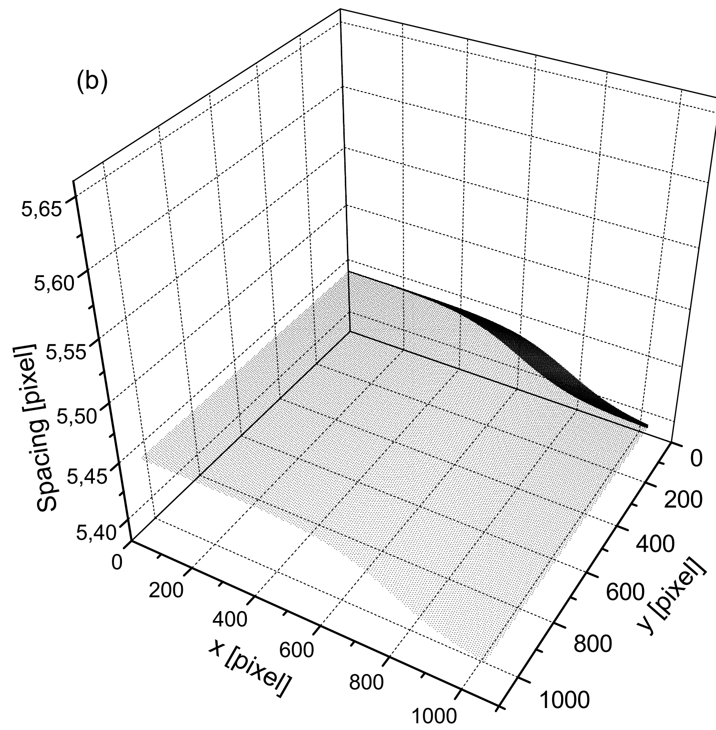


Figure 9: (a) 5% sharp ($w = 3$) step (b) 1% wide ($w = 100$) step

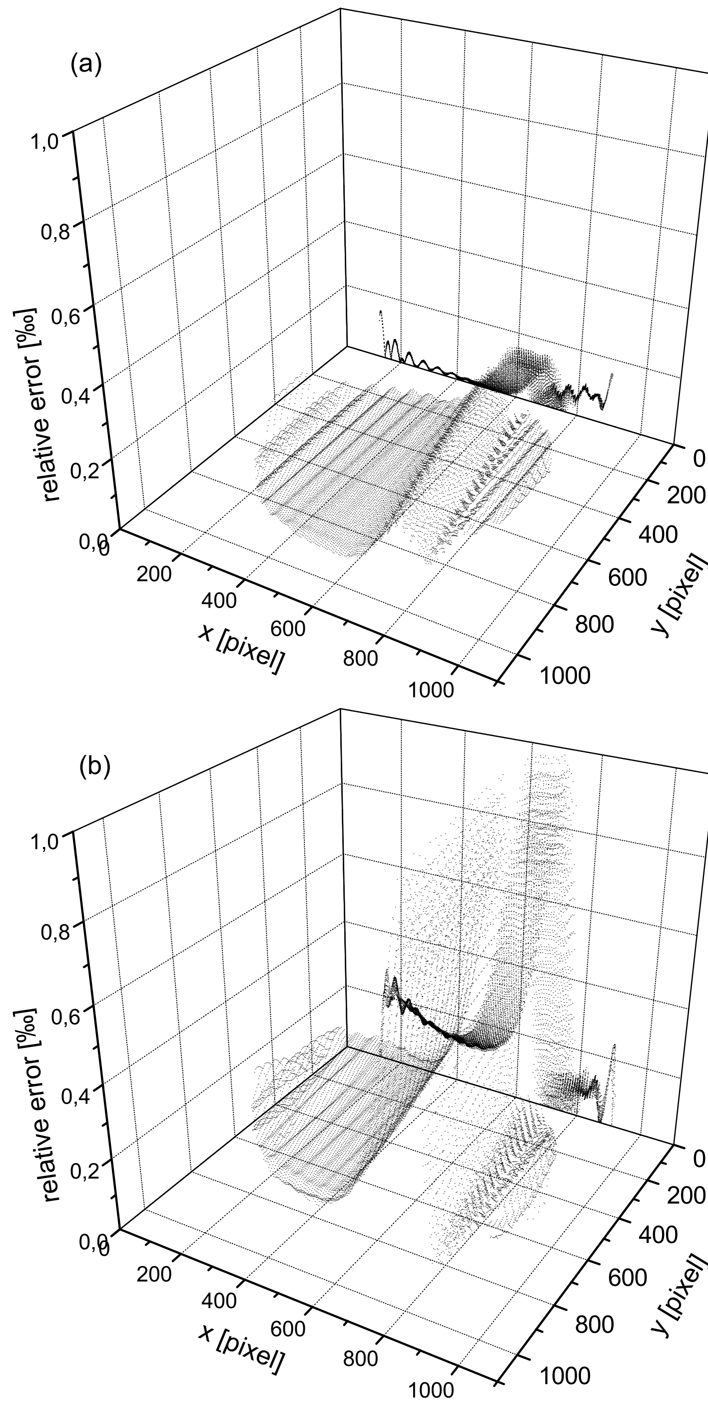


Figure 10: width: 40 (a) 1% step (b) 5% step

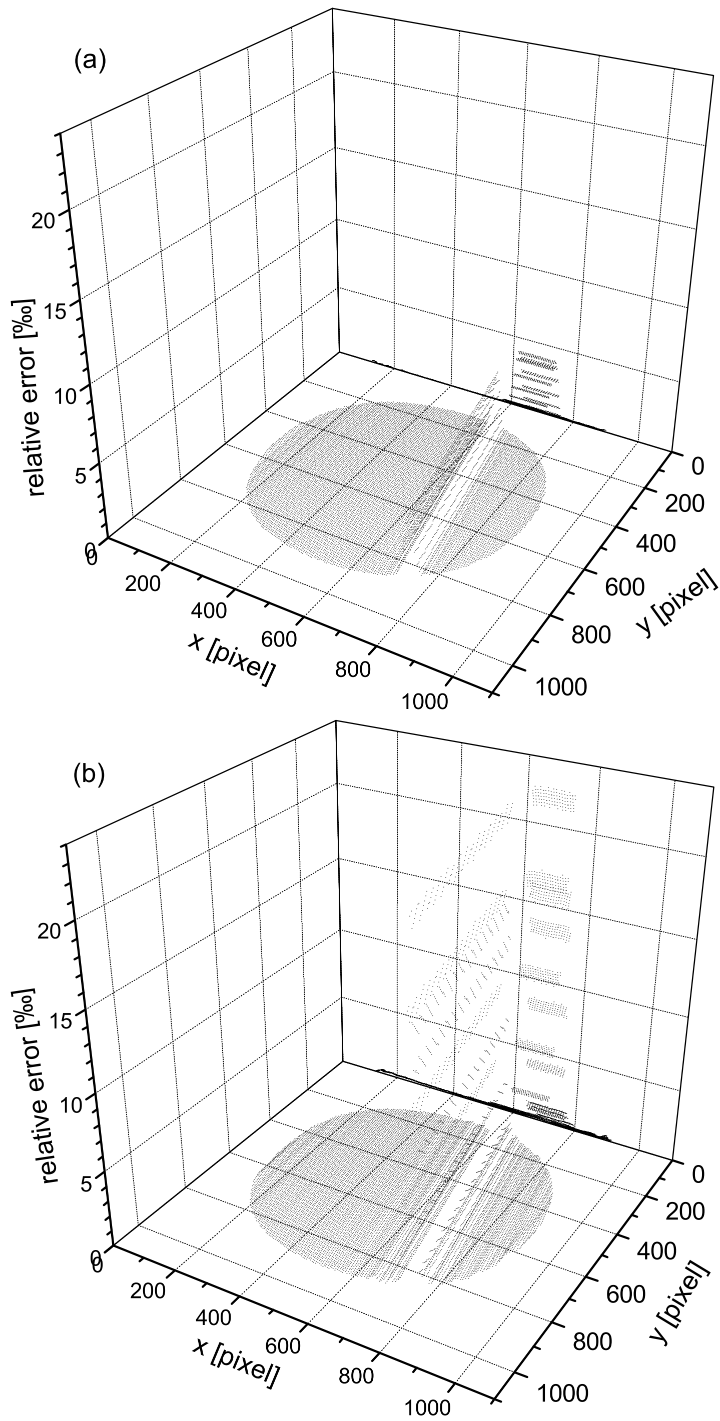


Figure 11: width: 1 (a) 1% step (b) 5% step

From the error levels we found we can draw the following conclusions:

- Besides the noise level the geometry itself must be considered as a reasonable source of error when the method is applied to experimental images.
- A sharp step that can be described as a 5% change in spacings within about 3 fringes can be detected with an accuracy of about 2.2%. This corresponds to an error of about 6*pm* in a 2.5Å spacing.
- Even sharper or higher steps can be measured with less accuracy, so other methods for spacing measurements should be considered for these geometries.

3 The nature of noise

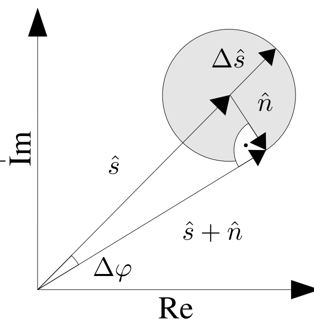
Noise is a fundamental phenomenon present in every physical measurement. It is defined in chapter 13.3 of [16] as a random signal $n(t)$ that is added to the original signal $s(t)$. It is most convenient to discuss the nature of noise and its effects in reciprocal (frequency) space. Any signal $s(t)$ can be represented by its complex valued Fourier components $\hat{s}(\nu)$ which are obtained by the FT of $s(t)$. Chapter 12 of [16] gives an overview of the properties of the FT. A very important property of the FT which follows directly from the definition is the additivity:

$$\widehat{s+n} = \hat{s} + \hat{n}$$

While the amplitudes of the signal $s(t)$ and the noise $n(t)$ are added at every instant in direct space, the corresponding Fourier components $\hat{s}(\nu)$ and $\hat{n}(\nu)$ are added in reciprocal space. This will allow us later to find a noise model in order to reduce the effect of noise!
A more precise definition of noise can be given by the properties of its Fourier components $\hat{n}(\nu)$:

- they have random phases without any correlation among them
- their absolute value changes randomly but it can usually be approximated by a steady function $n(\nu)$

The image to the right illustrates the additivity of the Fourier coefficients: The arrow pointing to the center of the circle represents the signal $\hat{s}$ while the circle represents the random Fourier components $\hat{n}$ with a constant absolute value that are added to $\hat{s}$. $\Delta\hat{s}(u,v)$ is simply $\hat{n}$, the amplitude of the noise coefficient, while $\Delta\varphi_{\hat{s}}(u,v) = \arcsin\left(\frac{\hat{n}}{\hat{s}}\right)$.



As the coefficients are added the arrows pointing to the edge of the circle represent the Fourier coefficients of the measured signal $\hat{s} + \hat{n}$. Note however that the error in the phase of the coefficient is about $\pm\frac{\pi}{10}$ while the error in the absolute value is about 24% in this example. The error in the phase is small if the amplitude of the noise is smaller than the amplitude of the original signal but the error increases with increasing noise level. The ratio of the signal and the noise amplitudes determines the quality of measured data. A common way to quantify the quality is the *signal to noise ratio* (SNR) defined in this context as:

$$SNR = 20 \cdot \log \frac{|\hat{s}|}{|\hat{n}|}$$

where $|\hat{s}|$ and $|\hat{n}|$ are the absolute values of the Fourier components. Usually the SNR is used in direct space but it is as useful in reciprocal space.

Does image processing change the SNR? No, it can be shown that shifting and scaling the input image in direct space does not change the SNR in reciprocal space. A common method is contrast enhancement which is a combination of shifting and scaling.

- Shifting in direct space adds a constant a to the input image. This constant has an effect on the DC coefficient only, thus the SNR in reciprocal space is not affected.
- Scaling the input in direct space is a multiplication with a constant c. The Fourier coefficients are multiplied by this constant, so they cancel out in the SNR.

Nevertheless attention must be paid when the input image is converted to or from another image format. The current implementation of the DF method requires 16 bit TIFF images as input. The TIFF format offers lossless LZW compression. JPEG images, for example, can not be used because the JPEG algorithm achieves compression by not saving high frequency Fourier components!

The SNR will be used later to measure the level of simulated noise and as a quantity from which the error of spacing measurements can be determined in experimental images.

3.1 A model for noise in a HRTEM image

In the previous chapter the SNR was defined. Now the power spectrum of a region in an experimental HRTEM image will be examined in order to obtain values for $|\hat{s}|$ and $|\hat{n}|$. Figure 12 shows the PS of an experimental image containing lattice fringes. In comparison figure 13 shows the PS of a simulated image. The 3D visualisation plugin of ImageJ was used to create these graphics.

In figure 12 the noise background and the peak which corresponds to the examined periodicity can be seen clearly. The plugin allows to rotate the graphic. Doing so, one notices that the PS is azimuthally symmetric except for the signal of course. This is true for images if the astigmatism has been corrected. If astigmatism were present the contour lines would be ellipses instead of circles. Because of the symmetry the azimuthal average of the PS with the DC peak as center will be calculated. This has the advantage that a one dimensional noise model can be made which is much simpler than a two dimensional model. Figure 14 and 15 show examples of azimuthally averaged power spectra.

Now the signal and the noise levels can be measured from these graphs. The signal level $|\hat{s}|$ is defined as the maximum value inside the aperture. This value and it's position ν_0 is measured by the program. To measure $|\hat{n}|$ the following noise model is made: We assume that the intensity of the noise background decreases linearly to the left of the peak in this logarithmic graph. Figure 14 is a nice example for that linear dependence. If a signal is present - as in figure 15

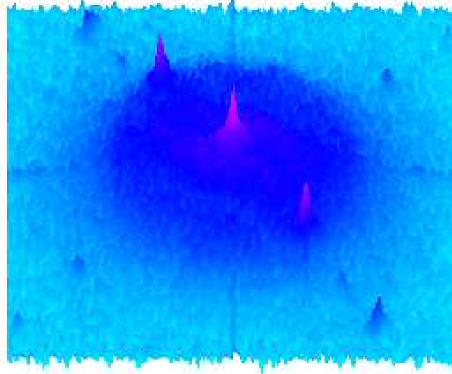


Figure 12: experimental PS

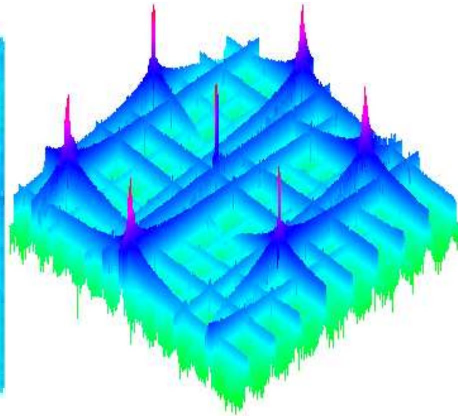


Figure 13: theoretical PS

- this linear relation is not valid anymore. Therefore we approximate the region to the left of the signal where it's effect is negligible by linear interpolation. When the parameters for this line have been found we define it's value at ν_0 as the noise level $|\hat{n}|$. From here the SNR can be calculated.

Table 2 shows the parameters for the interpolation of the power spectra shown in figure 15 and the resulting signal to noise ratios. $|\hat{n}|$ is approximated by $|\hat{n}(\nu)| = m \cdot \nu + c$.

negative	m [int/pix]	c [int]	$\lvert\hat{s}\rvert$ [int] @ ν_0	ν_0 [1/pix]	SNR [dB] @ ν_0
5001	-5.26	7.069	8.81	0.1346	49
5001	-3.35	6.576	8.65	0.1338	50
5209	-3.93	6.795	8.1938	0.1641	41
5209	-3.68	7.07	8.3212	0.1638	37

Table 2: SNR's of experimental images

Now that we know the quality of our experimental images we need to examine the effect the noise has on the spacing measurement.

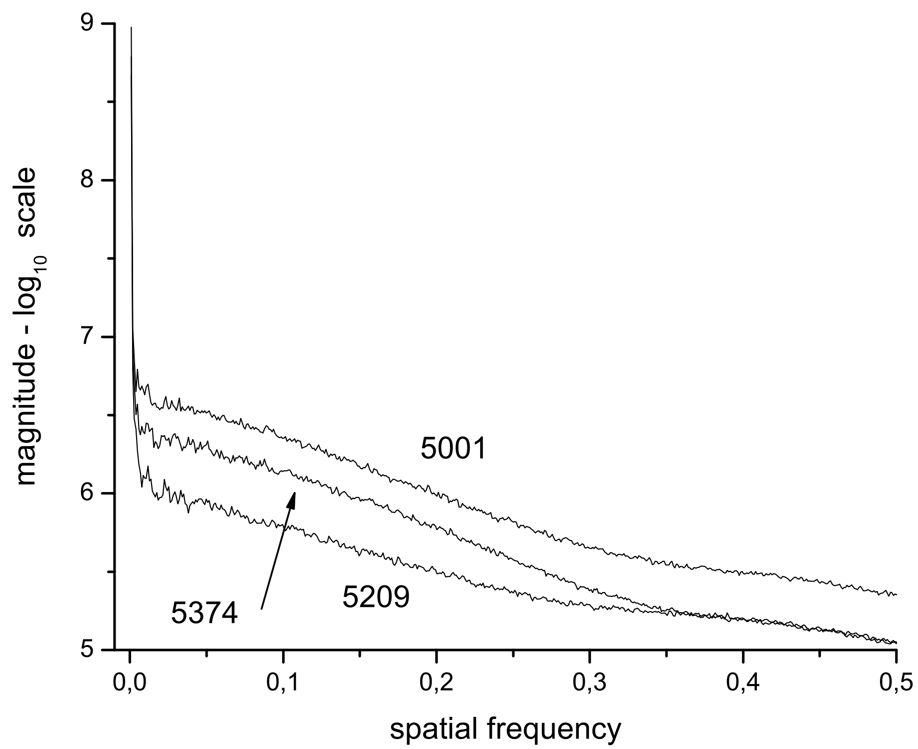


Figure 14: azimuthally averaged PS of images showing a hole in the specimen

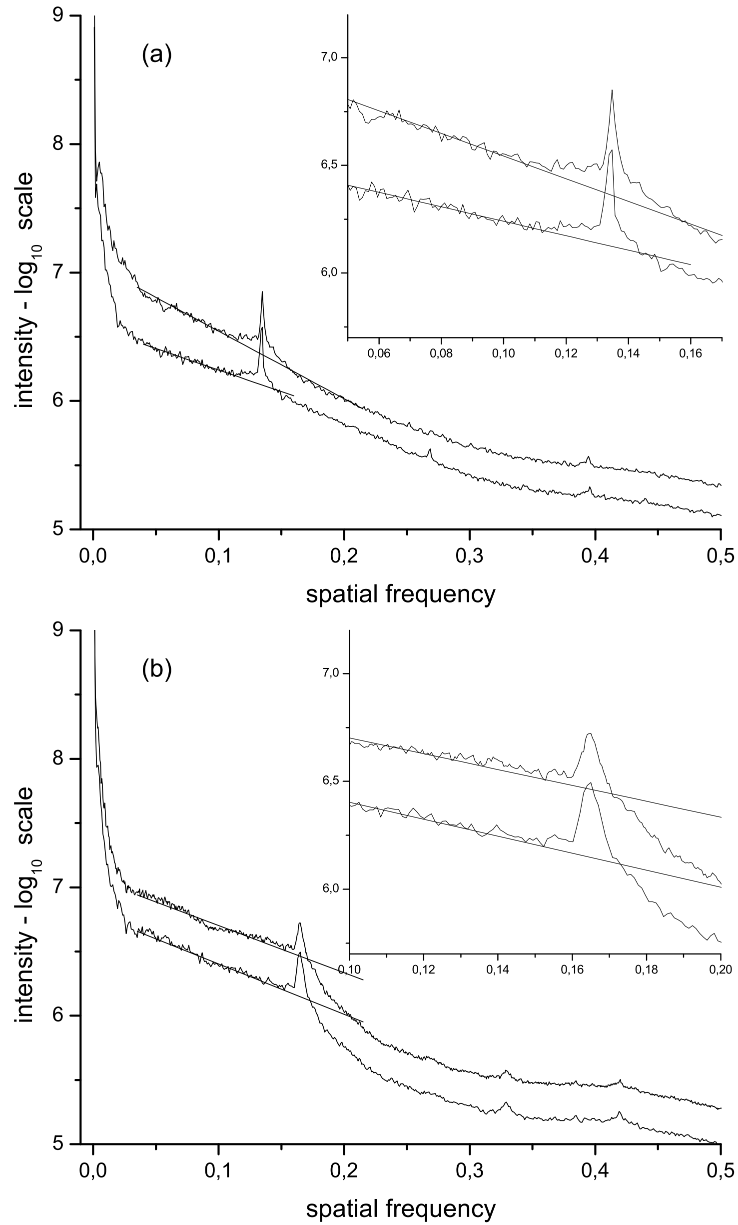


Figure 15: (a) negative 5001, fringes (b) negative 5029, fringes and amorphous material

3.2 Estimation of the error in the phase of the DF image

The error in the phase gradient will be approximated using Gauss's error propagation. Each point $s(x,y) \equiv S(x,y) \cdot \exp(i\varphi_s(u,v))$ in the DF image is given by the inverse FT of the Fourier coefficients $\hat{s}(u,v) \equiv \hat{S}(u,v) \cdot \exp(i\varphi_{\hat{S}}(u,v))$ by

$$\begin{aligned} s(x,y) &= \frac{1}{N^2} \sum_{u,v=0}^{u,v=N-1} \exp\left(i2\pi \frac{ux+vy}{N}\right) \cdot \hat{s}(u,v) \\ &= \frac{1}{N^2} \sum_{u,v=0}^{u,v=N-1} \hat{S}(u,v) \cdot \exp\left(i2\pi \frac{ux+vy}{N} + i\varphi_{\hat{S}}(u,v)\right) \end{aligned}$$

Capital letters denote the absolute value of a complex number while lowercase letters denote complex numbers, the hat denotes quantities in Fourier space. With Euler's formula we get

$$\begin{aligned} s(x,y) &= \frac{1}{N^2} \sum_{u,v} \hat{S}(u,v) \cdot \left\{ \cos\left(2\pi \frac{ux+vy}{N} + \varphi_{\hat{S}}(u,v)\right) + i \cdot \sin\left(2\pi \frac{ux+vy}{N} + \varphi_{\hat{S}}(u,v)\right)\right\} \\ &= \frac{1}{N^2} \sum_{u,v} \hat{S} \cdot \cos\left(2\pi \frac{ux+vy}{N} + \varphi_{\hat{S}}\right) + i \cdot \sum_{u,v} \hat{S} \cdot \sin\left(2\pi \frac{ux+vy}{N} + \varphi_{\hat{S}}\right) \\ &\equiv \frac{1}{N^2} \left(p(x,y) + i \cdot q(x,y)\right) \end{aligned}$$

where $p(x,y)$ is the first sum in the equation above and $q(x,y)$ is the second. Now the phase φ_s can be calculated

$$\varphi_s(x,y) = \arg(f(x,y)) = \arctan\left(\frac{q(x,y)}{p(x,y)}\right)$$

Gauss's error estimation is used to quantify the error in φ_s. The (x,y) dependency is not written.

$$\Delta\varphi_s^2 = \left(\frac{\partial\varphi_s}{\partial p}\right)^2 \cdot (\Delta p)^2 + \left(\frac{\partial\varphi_s}{\partial q}\right)^2 \cdot (\Delta q)^2$$

With the definition of φ_s follows

$$\Delta\varphi_s^2 = \frac{q^2}{(p^2+q^2)^2} \cdot (\Delta p)^2 + \frac{p^2}{(p^2+q^2)^2} \cdot (\Delta q)^2 = \frac{q^2(\Delta p)^2 + p^2(\Delta q)^2}{(p^2+q^2)^2}$$

The next step is to find Δp and Δq. Again Gauss's estimation is applied to p and q.

$$\begin{aligned} \Delta p^2 &= \sum_{u,v} \left(\frac{\partial p}{\partial \hat{S}(u,v)}\right)^2 \cdot (\Delta\hat{S}(u,v))^2 + \left(\frac{\partial p}{\partial \varphi_{\hat{S}}(u,v)}\right)^2 \cdot (\Delta\varphi_{\hat{S}}(u,v))^2 \\ \Delta q^2 &= \sum_{u,v} \left(\frac{\partial q}{\partial \hat{S}(u,v)}\right)^2 \cdot (\Delta\hat{S}(u,v))^2 + \left(\frac{\partial q}{\partial \varphi_{\hat{S}}(u,v)}\right)^2 \cdot (\Delta\varphi_{\hat{S}}(u,v))^2 \end{aligned}$$

From this point on the following simplifications and assumptions are made:

- $\Delta\hat{S}(u,v) \equiv \hat{N}$ where $\hat{N}$ does not depend on u and v which means that white noise is assumed here. This is a reasonable assumption as the radius of the aperture is small compared to the frequency range on which the noise level changes. This assumption is used only here, it does not apply to the noise model used for Wiener filtering. $\hat{N}$ has to be measured in the image by the method described in the previous section.
- $\hat{N} = 0$ outside the aperture so the number of values contributing to the sum is the number of pixels inside the aperture: $\#_{AP}$
- The Taylor expansion of $\arcsin\left(\frac{\hat{n}}{\hat{s}}\right)$ is used to approximate $\Delta\varphi_{\hat{S}} \approx \frac{\hat{N}}{\hat{S}} + O\left(\frac{\hat{N}}{\hat{S}}\right)^2$. This is a very good approximation for an experimental image, because the typical SNR is about 40dB which corresponds to $\frac{\hat{N}}{\hat{S}} = 10^{-2}$.

Now Δp^2 and Δq^2 take the following form

$$\begin{aligned}\Delta p^2 &= \sum_{u,v}\left\{\cos^2\left(2\pi\frac{ux+vy}{N} + \varphi_{\hat{S}}\right)\cdot\hat{N}^2 + \hat{S}^2\cdot\sin^2\left(2\pi\frac{ux+vy}{N} + \varphi_{\hat{S}}\right)\cdot\left(\frac{\hat{N}^2}{\hat{S}^2}\right)\right\}\\ &= \hat{N}^2\cdot\sum_{u,v}1 = \hat{N}^2\cdot\#_{AP} \qquad = \qquad \Delta q^2\end{aligned}$$

So $\Delta p^2 = \Delta q^2 = \hat{N}^2\cdot\#_{AP} \equiv \Delta^2$ and $p^2 + q^2 = N^4\cdot S^2$ leads to

$$\Delta\varphi_s^2 = \frac{2\Delta^2}{p^2+q^2} = 2\cdot\#_{AP}\frac{\hat{N}^2}{N^4\cdot S^2}$$

$\Delta\varphi_s$ depends on N, the number of points of the FFT and on the number of points in the aperture. However S is measured in direct space while $\hat{N}$ is measured in Fourier space! So a connection between S and $\hat{S}$ must be made.

The maximum signal S is obtained when all waves inside the aperture are added in phase. The upper limit for S is the sum of all Fourier coefficients inside the aperture. This sum is written as a multiple of $\hat{S}$: $S = \frac{\hat{S}\cdot m}{N^2}$ where m is a unknown constant. The final expression for the error in the phase is therefore:

$$\Delta\varphi_s = \frac{\sqrt{2\cdot\#_{AP}}}{m}\cdot\frac{\hat{N}}{\hat{S}} \qquad \text{where} \qquad \#_{AP} = r_{AP}^2\cdot\pi$$

The spacing d is the quantity we are interested in. It is calculated from the phase gradient which is obtained by subtracting nearby phases, depending on which numerical expression is used. The error in the phase gradient is $c\cdot\Delta\varphi_s$ while c depends on the numerical expression. $c = \sqrt{2}$ for the forward or backward derivative and $c = \sqrt{1/2}$ for the central derivative which is more accurate. Using the definition of the spacing and Gauss's error propagation the relative error in a spacing measurement is found to be

$$\frac{\Delta d}{d} = r_{AP}\cdot\frac{\hat{N}}{\hat{S}}\cdot\frac{c\cdot\sqrt{2\pi}}{m\cdot k}$$

It increases linearly with the radius of the aperture r_{AP} and the noise to signal ratio $\vec{N}/\vec{S}$. As $d = 2\pi/k$ the error in k has a bigger impact on d if k is small. Therefore the relative error increases with the spacing d. The constant m is known approximately only.

3.3 Measurement of the error in spacing measurements

In this section the error in spacing measurements will be measured in order to check the dependencies on aperture radius and noise level. To do so the spacings must be known of course, so the spacings in computer generated images will be measured. Even if we had a perfect sample and a perfect micrograph we would still not know the spacings between lattice fringes in the micrograph with high spacial resolution because the magnification of the microscope is not know with high accuracy.

As indicated before these measurements will be used to estimate the error in the measurement of a spacing in an experimental image. A relation between the error and the quality of the image is required to do that. This relation has been derived in the previous section but the constant m is just approximately known. To measure this relation a certain amount of white noise is added to the image and the spacing measurement is performed. At this point it is assumed that the simulated images contain no noise. This is a reasonable assumption as the periodicity is not running parallel to an edge of the image, so the artifacts due to non periodic boundaries do not overlap with the peak which corresponds to the periodicity. The only noise present in the whole PS is due to round off errors that occur during the FFT, but we expect them to be negligible.

The noise will be added directly in Fourier space so we do not need to calculate the image of the noise in direct space. A noise level $\hat{N}$ is set and each Fourier coefficient inside the aperture is modified in the following way:

$$\hat{f}(u,v) = \hat{f}(u,v) + \hat{N} \cdot (1+r) \cdot \exp(i \cdot \varphi_r)$$

where r is a random number between -0.1 and 0.1 and φ_r is a random phase between 0 and 2π. r creates amplitude noie, while φ_r creates phase noise. When the noise has been added the measurement is performed as usually. 500 measurements with the same noise level are made at a certain location in order to have enough values for a statistical analysis. As the noise coefficients themselves are different each time the result of the measurement is different every time.

To examine the distribution of the measurements the number of measurements that lie in a certain interval is counted. The *frequency count* feature of OriginPro [1] is used to do that. The obtained distribution can be interpolated by a gaussian profile. The FWHM of this gaussian profile is defined as the error in the measurement. Figure 16 (a) shows an example of such distributions. The spacing has been measured at 500/500 with an aperture of 20 in the sharp step geometry described in 2.4. The regions where the spacing has been measured have a constant spacing. Therefore the error in the measurements is exclusively due to the noise. The effect of the change in $\vec{k}$ will be examined in chapter 4.1.

Figure 16 (b) shows the linear increase of the error with increasing N/S. Each point in this figure has been obtained from a distribution of 500 measurements. The y axis shows the relative error in the measured spacing.

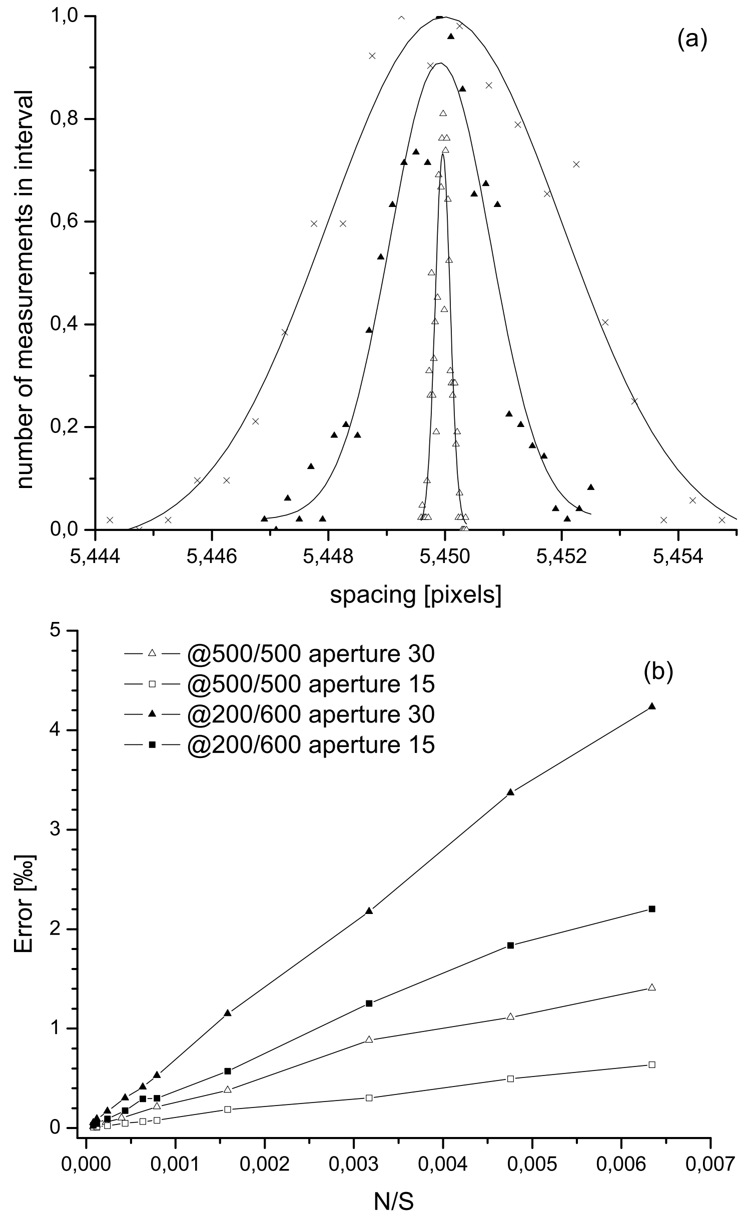


Figure 16: (a) statistical distribution of measurements (b) linear increase of error with N/S

Now the influence of the aperture on the error will be examined. Figure 17 shows the relative error in spacing measurements and the center of the distribution of measurements for different radii of the aperture. The noise level is the same for all measurements. Again 500 measurements were made at each point. The same analysis of the distribution as described before is used. The linear increase of the error with increasing aperture radius r_{AP} can be seen clearly. We can also see that the distribution of measurements is not centered around the correct value (horizontal line) for small radii.

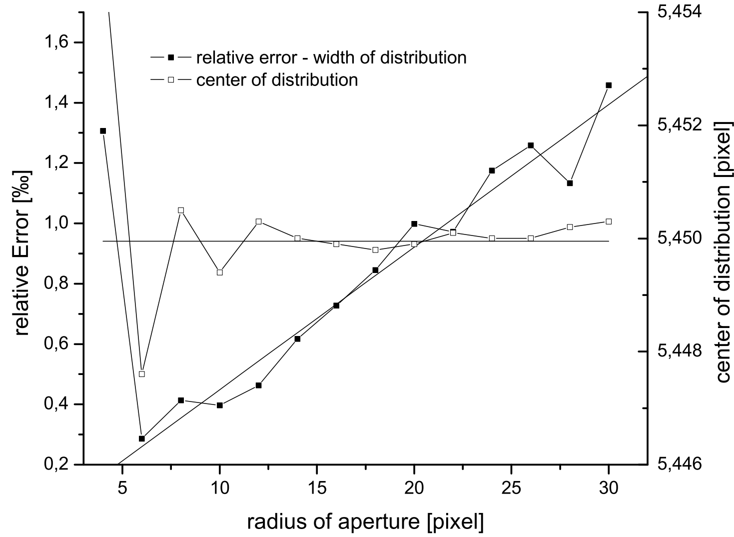


Figure 17: linear increase of error with aperture, $N/S \approx 0.006$ SNR=44bB

We can conclude that there must be an ideal radius for the aperture. The error decreases as long as decreasing the aperture excludes points containing mostly noise. If the aperture gets too small points containing the signal will be excluded. The error resulting from that is hard to predict. As the aperture gets smaller the position of the center of the aperture becomes more important. These effects will be studied more accurately in chapter 4.1.

From the linear interpolations in figure 16 b and 17 values for m can be determined. To do this, the error is linearly interpolated:

$$\frac{\Delta d}{d} = B \cdot \frac{\hat{N}}{\hat{S}} \qquad \text{or} \qquad \frac{\Delta d}{d} = B \cdot r_{AP}$$

depending on which quantity is varied. B is the slope of the linear interpolation.

m is simply given by

$$m = r_{Ap} \cdot \frac{c \cdot \sqrt{2\pi}}{k \cdot B} \qquad \text{or} \qquad m = \frac{\hat{N}}{\hat{S}} \cdot \frac{c \cdot \sqrt{2\pi}}{k \cdot B}$$

where r_{AP} (figure 16 (b)) and $\hat{N}/\hat{S}$ (figure 17) is constant respectively. In the case where r_{AP} is kept fixed and $\hat{N}/\hat{S}$ is varied the following values for m are found:

B	r_{Ap}	Position	k	m
0.681	30	200/600	1,1529	66
0.361	15	200/600	1,1529	63
0.228	30	500/500	1,1668	201
0.100	15	500/500	1,1668	228

Table 3: m obtained from figure 16 (b)

In the case where r_{AP} is varied we find a slope of $B = 0.05$ which leads to $m = 185$ where the $\hat{N}/\hat{S}$ ratio is 0.006.

The values that have been determined for m are much bigger than expected. Therefore the error needs to be measured when both sources of error are present. This will be done in section 4.1.

3.4 Removal of noise

Now that a noise model for experimental and simulated images has been established and the effect of noise on the Fourier coefficients is understood, it suggests itself to remove the noise or at least to reduce its effect. The noise models described above model the amplitude of the Fourier coefficients of the noise, not the phase! A model for the phase is way more complicated or even impossible to come up with, because the processes creating the noise are not known in detail. For the simulated white noise it's definitely impossible because the phase is random.

The simplest way to reduce the effect of noise is to substract the absolute value of the noise coefficient from the absolute value of the signal coefficient. This would correct the error if both coefficients had the same phase, which is in general not the case.

A more sophisticated method is *Wiener filtering* which is derived in chapter 13.3 of [16]. This technique requires the noisy signal and a noise model to calculate a filter factor Φ. Each noisy Fourier coefficient is multiplied by the corresponding factor $\Phi(\nu)$ which yields a coefficient $\hat{U}'(\nu)$ that is as close as possible to the uncorrupted coefficient $\hat{U}(\nu)$. That means the obtained coefficients minimize the sum of the squared deviations:

$$\int_{-\infty}^{\infty} \left| \hat{U}'(\nu) - \hat{U}(\nu) \right|^2 \Rightarrow min$$

Φ is found to be:

$$\Phi(\nu) = \frac{\left|\hat{S}(\nu)\right|^2}{\left|\hat{S}(\nu)\right|^2 + \left|\hat{N}(\nu)\right|^2}$$

$|\hat{S}(\nu)|$ and $|\hat{N}(\nu)|$ are the Fourier coefficients of the signal and the noise at frequency ν, both can be determined from the PS. $|\hat{S}(\nu)|$ is simply given by the noisy Fourier coefficient minus $|\hat{N}(\nu)|$, see figure 3.4 (a) for an example of a noise model and (b) for the extracted signal. $\Phi(\nu)$ is close to 1 where the noise is negligible and 0 where the noise dominates. Figure 3.4 (c) shows the factor Φ and (d) shows the filtered signal. Because the factor $\Phi(\nu)$ results from a minimization problem the obtained result $\hat{U}'(\nu)$ differs from $\hat{U}(\nu)$ by an amount that is second order in the precision to which $\Phi(\nu)$ has been determined. This means that the noise model we established for experimental images can be used for Wiener filtering, although it is just a first order approximation.

In chapter 4 this filtering technique will be applied to simulated images after white noise has been added and to experimental images using the (in log_{10} scale) linear noise model.

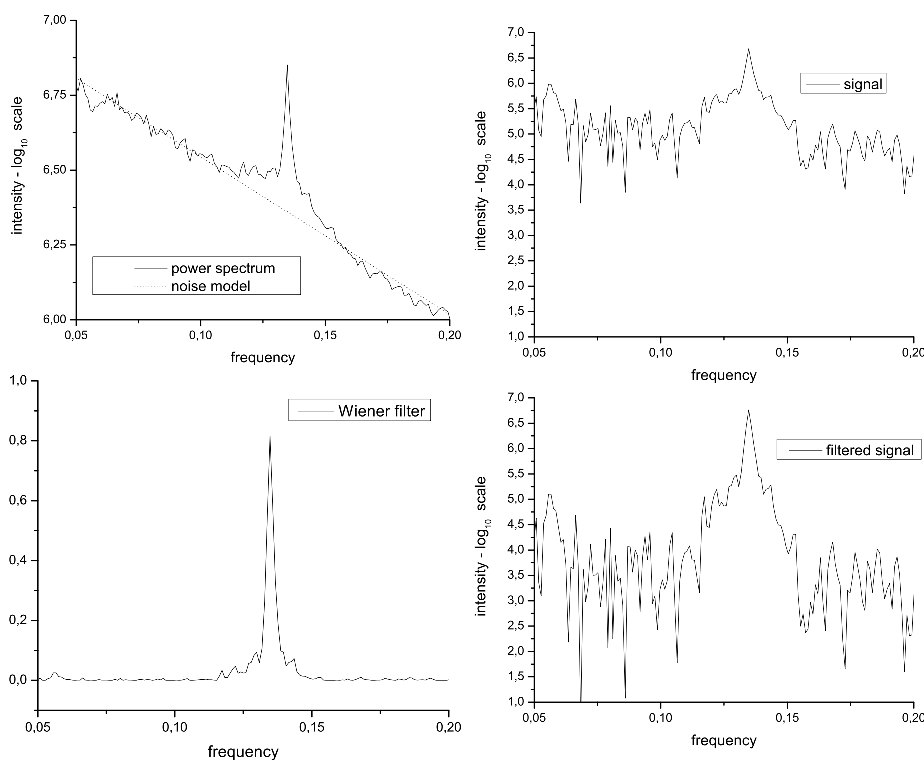


Figure 18: (a) azimuthally averaged PS and noise model (b) extracted signal (c) filter factor $\Phi(\nu)$ (d) filtered signal

3.5 Multiple spot measurement

In section 3.3 the statistical distribution of spacing measurements has been examined. If a measurement of a noisy quantity is performed it is natural to repeat the measurement in order to obtain more data and to find the final answer by some kind of averaging over all data points. Figure 16 (a) shows a distribution of measurements. The maximum of the distribution is at the average value if a gaussian distribution is assumed. The average value is very close to the exact value, as we expected!

How can we make use of the idea to average multiple measurements in the DF method? Well, there are basically two different possibilities: The first one is to collect measurements from different images taken in sequence. This would be a good solution if the images could be obtained easily from a digital camera connected to the microscope as we could assume then, that the noise level and the corresponding noise model is identical in every image. We use an analog camera to record the image on a micrograph which is developed and finally digitized.

We do not expect to have the same noise model in two different micrographs because there are too many processes involved in obtaining the image. Therefore we use a method that collects different measurements from the same micrograph. In theory we can measure a spacing at a given position in the image for every spot that appears in the PS. In figure 5 a basis was introduced and the spots were indexed using this basis, so that it is possible to relate the measured effective wave vector $\vec{k_{eff}} = \vec{k_i}$ to the basis vectors. For each spot we obtain a vector $\vec{k_i}$ and a linear equation:

$$\vec{k_i} = c_i \cdot \vec{a} + d_i \cdot \vec{b}$$

where c_i and d_i are the indices assigned to the spot i. We end up with a set of N equations containing wave vectors and indices from N different spots. This linear system needs to be solved for the basis vectors from which all spacings at a certain position can be calculated. N can be larger than 2 so there can be more equations than variables which is assumed to be the case. For example the set of equations for the x components of the basis vectors can be written like this:

$$\begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \\ ... & ... \\ ... & ... \end{pmatrix} \cdot \begin{pmatrix} a_x \\ b_x \end{pmatrix} = \begin{pmatrix} k_{1x} \\ k_{2x} \\ ... \\ ... \end{pmatrix} \equiv \mathbf{A} \cdot \vec{x} = \vec{b}$$

We are looking for the solution $\vec{x}$ which minimizes the norm

$$\left| \mathbf{A} \cdot \vec{x} - \vec{b} \right|$$

With the help of the singular value decomposition the matrix $\mathbf{A}$ can be written as $\mathbf{A} = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V^T}$, as described in chapter 2.6 of [16]. $\mathbf{U}$ and $\mathbf{V}$ are orthogonal and $\mathbf{S}$ is diagonal containing the singular values of $\mathbf{A}$. The solution is given by

$$\vec{x} = \mathbf{A^{-1}} \cdot \vec{b} \qquad \text{where} \qquad \mathbf{A^{-1}} \equiv \mathbf{V} \cdot \mathbf{S^{-1}} \cdot \mathbf{U^T}$$

is called the pseudo inverse of $\mathbf{A}$.

The solution $\vec{x}$ is determined from multiple points. This minimization process is not to be confused with spatial averaging. The spatial resolution of the DF method is not reduced by the multi spot method because each spot provides an additional data point. It is expected that this method reduces the effect of noise because noise tends to cancel in an averaging process, however it can not reduce the error introduced by the change in $\vec{k}$.

The noise model introduced in section 2 of this chapter assumes a linear (in log scale) decrease of the noise amplitude with frequency, so using spots at higher frequencies automatically has the advantage of lower noise levels! However the signal amplitude is weaker for higher frequencies as well. The SNR of each spot can therefore be used to weigth the values obtained from different spots. Measurements obtained from a spot with a higher SNR are to be trusted more than measurements obtained from noiser spots.

If the images were taken by a digital camera, both averaging techniques described above could be used in parallel.

4 Application of the DF method

In this chapter the DF method will be applied to simulated and experimental images. First it will be applied to simulated images in order to learn more about the error when both sources of error, namely noise and changing spacings are present. We also want to understand the influence of the center and the radius of the aperture. Examples for noise removal and multispot measurements in simulated and experimental images will be shown.

4.1 computer generated images

unstrained geometry with noise

The first image that will be examined here is a unstrained geometry with a noise level of 40bB. We use the basic geometry described in 2.4 and add noise. The maximum in the aperture is $10^{9.148}$ which is the signal strength. We add noise with an amplitude of 10^7 which corresponds to a SNR of 43bB. The table below shows the results for various settings.

Aperture center	r_{Ap}	Error [‰]
697/481	25	10
697/481	20	6
697/481	15	2
699/480	10	1.2

We see that even without noise removal the spacings can be measured easily with an accuracy of 1 ‰ in the unstrained geometry. This is not too suprising though because the spot in the PS has little structure, so the aperture can be very small.

wide step geometry with noise

The most interesting simulated image is probably the wide step geometry with extreme parameters as shown in figure 9 (a) with noise. Again we add noise with an amplitude of 10^7, the signal level is the same as before. The parameters for the geometry are $w = 1$ and $a = 5\%$. The table below shows the results of the measurements.

Aperture center	r_{Ap}	Error [‰]	comment
690/482	25	25	
690/482	20	22	
690/482	20	20	no noise
690/482	10	-	step is dark in DF image
695/482	10	22	
695/482	10	32	SNR 23dB

The third measurement shows an interesting effect: The radius of the aperture has been decreased and one side of the step appears dark in the DF image, see figure 19 middle. This means some important Fourier coefficients were not included. When the aperture is shifted to the right by 5 pixels, both sides of the step appear bright again, figure 19 left. This effect tells us to position the aperture so that the whole spot is inside the aperture. In an experimental image this is harder to do because the spot is spread out more, but the DF image can be used to see whether enough data was included in the aperture for the region of interest. The right image in figure 19 shows a DF image with high noise levels.

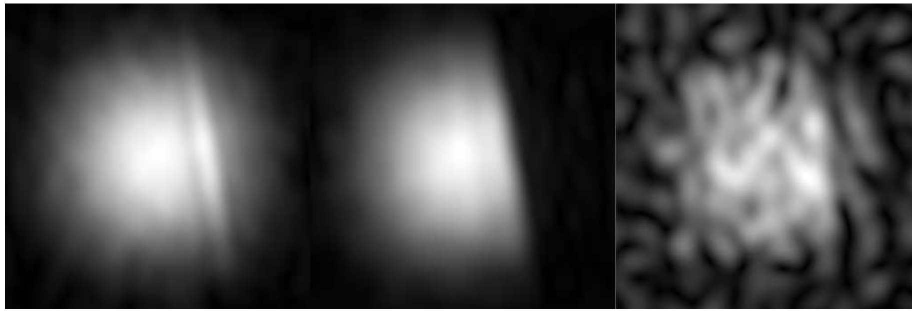
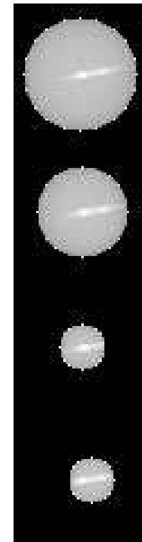


Figure 19: DF images for different settings and noise levels



The image to the right shows the position, the size and the content of the aperture for the different measurements. The brightness changes because of different scale factors, this and the white dots are due to visualization.
In the image on the top the radius is 25 pixels, the whole spot is inside the aperture. We can recognize two bright areas connected by a dimmer area. In the second image the radius is decreased to 20 pixels and the aperture still contains both bright areas. The number of points containing noise only has been reduced. The third image explains why the side of the step which contains the smaller spacings appears dark in the DF image: The DC peak is to the lower left in these images, so the bright area further away from the DC peak (to the top right) corresponds to higher frequencies or smaller spacings respectively. As this area is not included in the aperture the DF image is dark in regions containing the smaller spacing. The fourth image shows the aperture shifted a little to the right. The whole spot is inside the aperture again, the DF image is bright in both areas.

We realize that result of the measurement is not very sensitive to the choice of the center of the aperture. When placing the aperture care should be taken so that the whole spot is inside the aperture. Then the radius should be minimized in order to reduce the number of points containing noise only.

4.2 experimental images

3 buried quantum dots

In this section the DF method will be applied to a GaInP/InP heterostructure grown on GaAs (001) substrate. As described in [14] the sample was grown by molecular beam epitaxy which allows to deposit single layers of atoms on a substrate. First of all a 200nm buffer GaAs layer was grown, followed by a 45nm GaInP layer. The lattice constant of this layer was matched to GaAs. This matching is done by changing the Ga/In ratio, here $Ga_{0.52}In_{0.48}P$ is used. Now three monolayers of InP are deposited followed by a GaInP spacing layer. This InP/GaInP sequence is repeated three times with spacing layer widths between 2 and 16nm. Finally the structure was capped by a 45nm GaInP layer. The lattice mismatch between GaInP and InP is 3.7% which makes it an ideal test structure because the error inherent in the method is expected to be smaller than 2%. Due to this lattice mismatch the atoms have to rearrange close to every InP/GaInP interface which leads to so called buried quantum dots. The electrical and optical properties of these quantum dots depend on their size and their shape which in turn can be derived from spacing measurements.

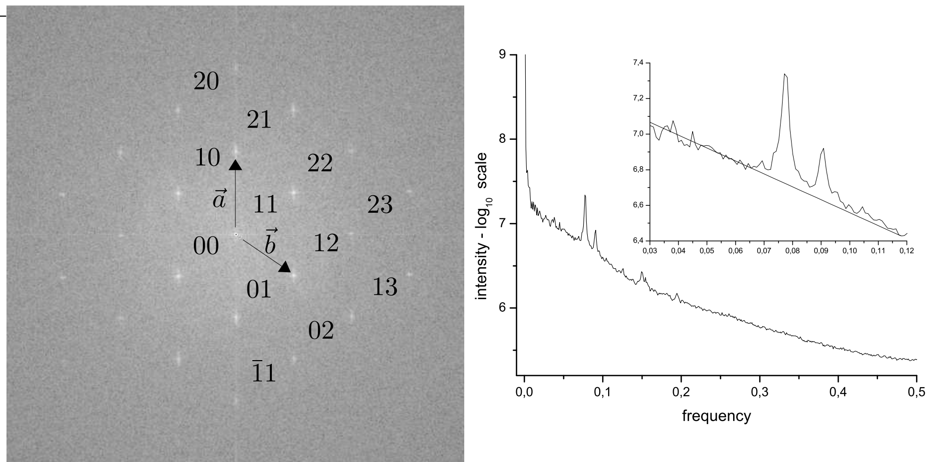


Figure 20: PS and azimuthal average of PS

Figure 21 shows a HRTEM image of the region containing the structure described above, this image and image 25 was kindly provided by Dr. N.Y. Jin-Phillipp, Max-Planck-Institut für Metallforschung at Stuttgart. Figure 20 shows the power spectrum and its azimuthal average. The PS shows many high frequency spots which indicates that only slight changes in spacings are present. The spots have been indexed using the basis vectors $\vec{a}$ and $\vec{b}$.

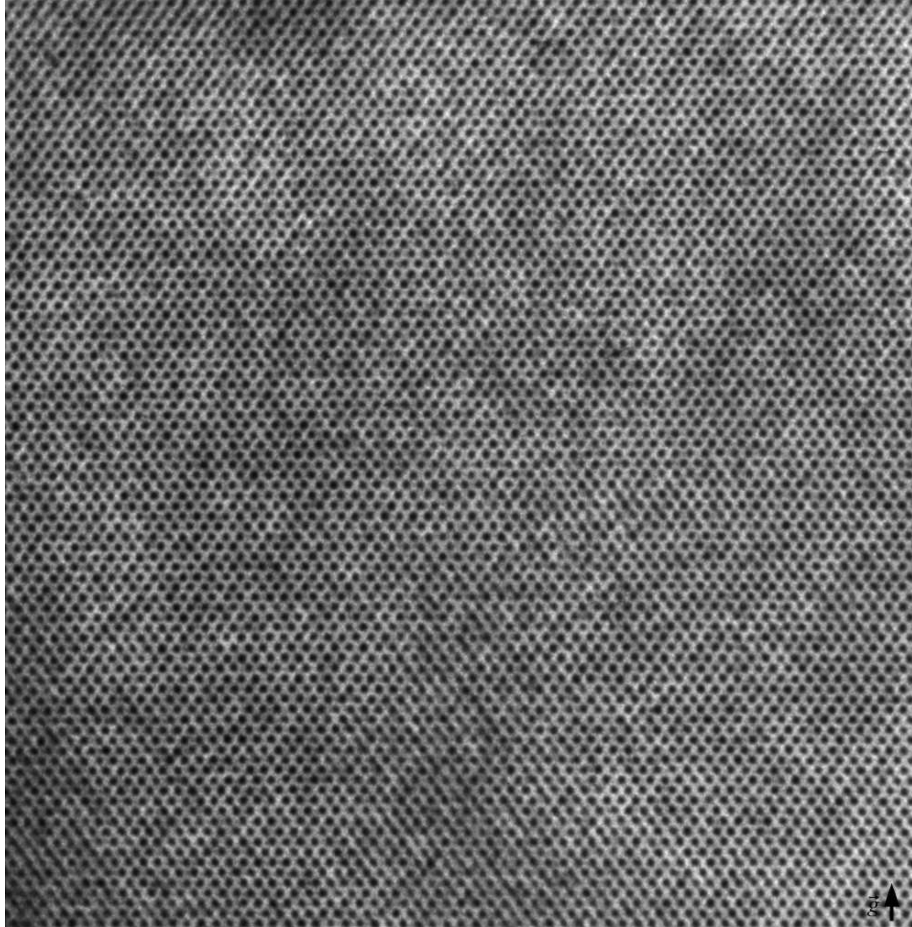


Figure 21: HRTEM image of 3 buried InP quantum dots

First of all the DF method will be applied using the 10 spot, as we expect a change in the corresponding spacing due to the lattice mismatch described before. An aperture of radius 11 is drawn around the brightest point of the 10 spot. Figure 22 shows the result of the spacing measurement. We can clearly identify the three quantum dots that appear as areas with larger spacings, separated by areas of smaller spacings. However, the quantum dots were expected to appear on the top of the image, as inferred in [14]. The top left corner of the input corresponds to the origin in figure 22. This problem can be solved by comparing the orientation of the TEM image with the orientation of the specimen. At this point we will accept this difference as it is not a systematical error. As the lattice mismatch is 3.7% we expect the error to be 2% at most.

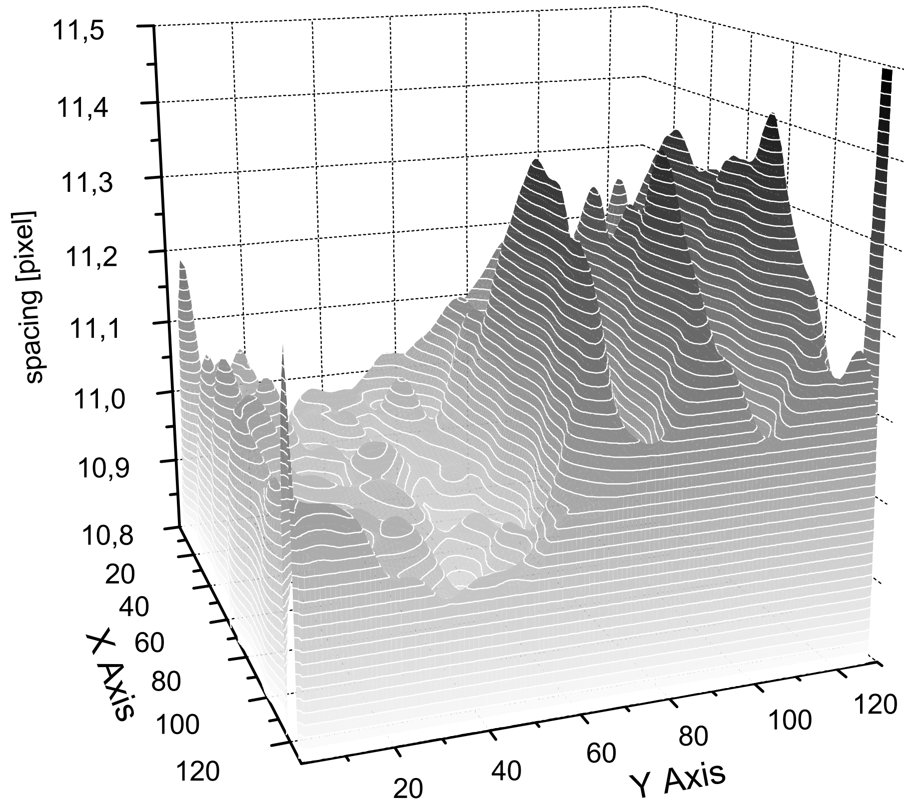


Figure 22: spacing measurements in GaInP/InP heterostructure

Now multi spot measurements of the same spacing will be performed. The first multi spot measurement will use these spots: 10, 01 and 11. The second one uses spots 12 and 21 additionally, while the third measurement uses all previous spots plus the 20 and 02 spot. The results from all measurements are weighted equally for simplicity. The table below shows the noise levels for the used spots.

Spot	SNR [dB]
10	41
01	44
11	45

Spot	SNR [dB]
12	30
21	34

Spot	SNR [dB]
20	23
02	29

Figure 23 shows the results of the multi spot measurements: the top left is the single spot measurement, top right: 3 spots, bottom left 5 spots, bottom right 7 spots. If we compare the two images on the top we can see that the structures are a little smoother in the 3 spot measurement. On the bottom we can see that the result does not improve further. The black dots in the middle of a white region are points where the result is out of scale. This is because the amplitude of the DF image has not been taken into account here. We know that the spots corresponding to higher frequencies have a lower SNR therefore they should not be weighted equally. The amplitude in the DF image can be used as a quantity for local weighting of measurements obtained from one spot, while the SNR could be used as a global weighting factor for all measurements obtained from one spot.

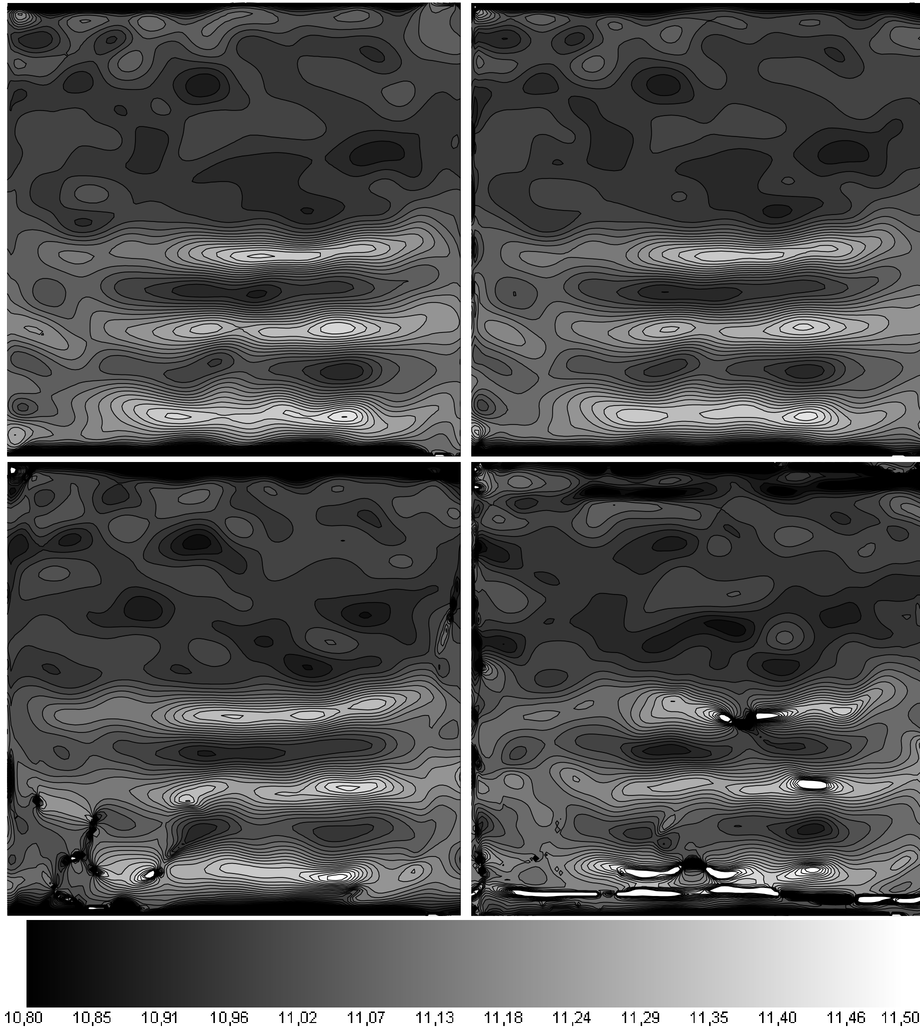


Figure 23: multi spot measurements: top left 1 spot, right 3 spots; bottom left 5 spots, right 7 spots

Figure 24 shows an overlay of results from the 3 spot measurement and the input image. This map makes it easier to interpret the result and to connect it to the specimen. In the top area we find spacings between 10.91 and 11.01

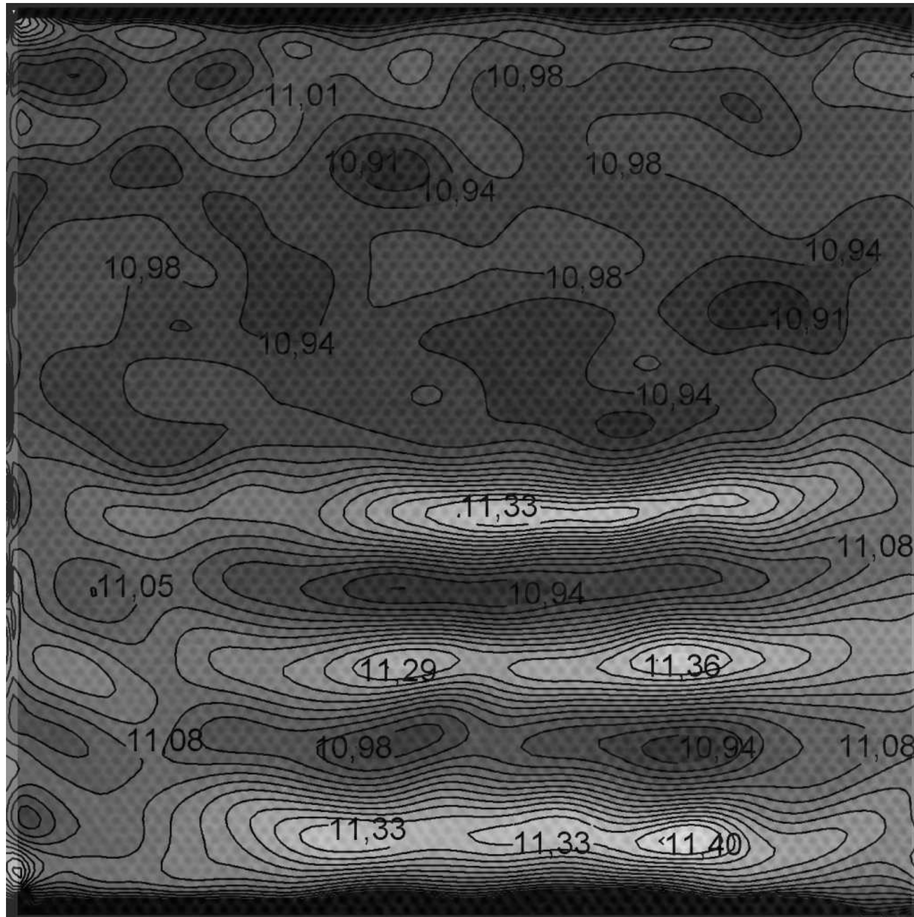


Figure 24: spacing measurements in GaInP/InP heterostructure

pixels. We can also say the average spacing is $10.95 \pm 1\%$ in this area. In the areas of the quantum dots the spacing increases to $11.33 \pm 2\%$ pixels; between the dots it decreases to $10.94 \pm 2\%$ again. In the input image the different layers are not visible, therefore the DF method can be used to detect hidden structures! This map was created using the GNU Image Manipulation Program (www.gimp.org).

free standing quantum dot

In this section the DF method will be applied to measure the latice spacings in a free standing quantum dot, shown in figure 25. Again the substrate is GaAs while the quantum dots are made of $In_{0.6}Ga_{0.4}As$ as described in [13]. The PS contains the same spots as in the previous example because the same material is used and the sample is oriented in the same way. Therefore the same indexing as in figure 20 is used to label the spots.

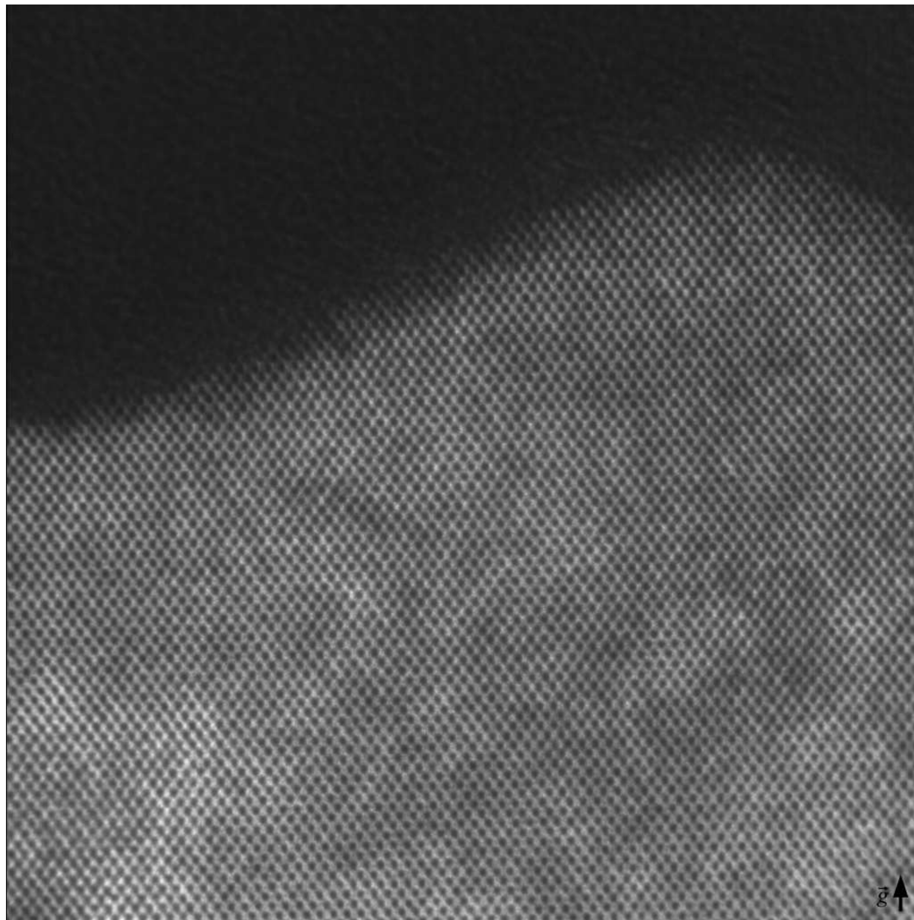


Figure 25: HRTEM image of a free standing quantum dot

The table below shows the SNRs for the various spots:

spot	10	01	11	12
SNR [dB]	33	43	44	33

From the results of the previous example we expect a change in the spacing which corresponds to the 10 spot. But we just found that the SNR for the 10 spot is about 10 dB smaller compared to the 01 spot for example. This is a perfect example for the multi spot method because we can get a second measurement for the 10 spacing to increase the accuracy, simply by using the 01 and 11 spots.

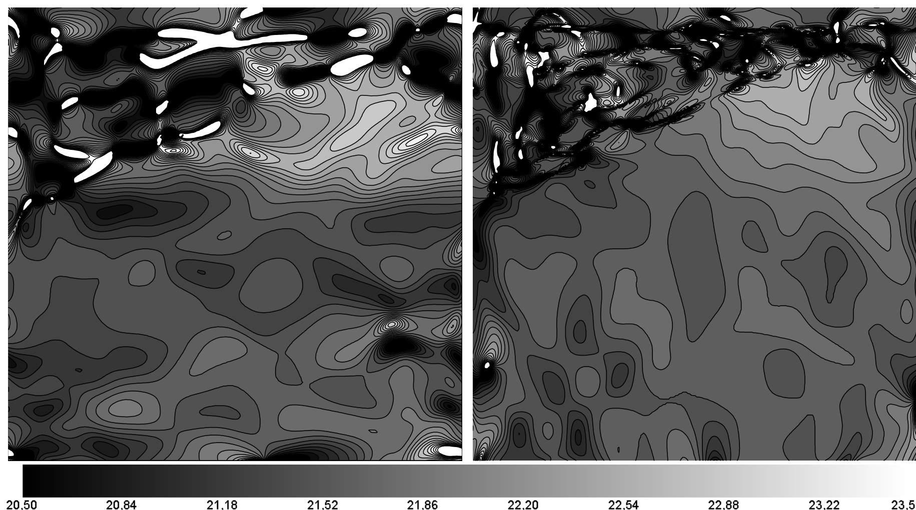


Figure 26: left: 10 spot, right: 10, 01 and 11 spot

The measurements have been performed with an aperture of radius 11, centered around the brightest point of each spot. Again all spots were equally weighted. The left image in figure 26 shows the results of the spacing measurements for the 10 spot. We can clearly see the region to the upper left and top where dark and bright spots are close together. This area is dark in the DF image, as it contains no periodicities but amorphous material. We recognize many noisy areas as well as bright and dark peaks in the area containing the periodicities. The right image shows the result of the multi spot measurement. A Wiener filter with a noise level of $7.9 \cdot 10^6$ has been applied to all spots. The edge between amorphous and periodic regions is sharper compared to the 10 spot measurement. In the periodic region the noisy spots disappeared or got washed out. The distribution of spacings appears much smoother although the multi spot method does no spacial averaging.

The width of the interface is about 30 pixels which corresponds to $w = 10$ if we want to approximate the structure with the wide step geometry. The spacings

increases by 5% between the darker area and the brighter area. Therefore we expect the error to be less than 1%. Figure 27 shows the overlay of the multi spot spacing measurements and the input in order to connect the result to the specimen. In the area of constant spacing the spacing is $15.2 \pm 1\%$ pixel and it increases to $15.9 \pm 1\%$ in the top of the quantum dot.

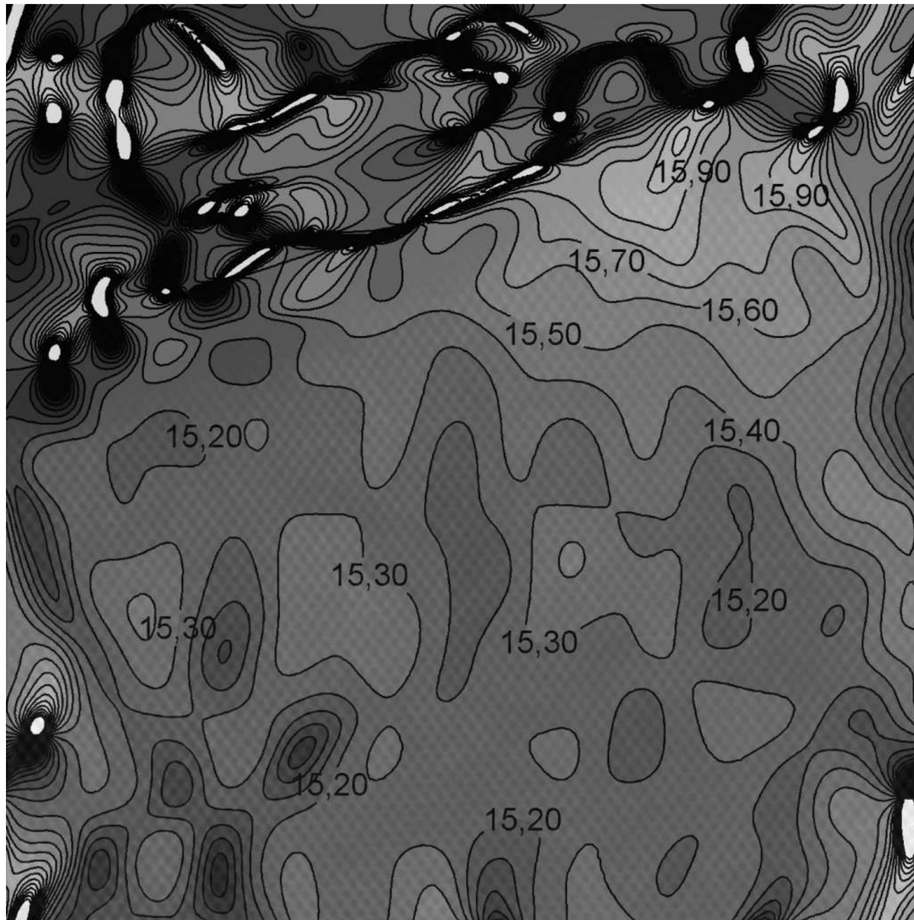


Figure 27: overlay of the input and the spacing map

heterostructure on sapphire

In this section the DF method will be applied to a HRTEM image of a heterostructure grown by Dr. D. Leopold. The different layers were grown on a sapphire substrate. First of all a 54*nm* GaN layer was deposited followed by a 68*nm* AlGaN layer, a 95*nm* GaN layer and a 284*nm* InGaN layer. The thickness of these layers was determined by Sam from a TEM micrograph. The image contains one set of lattice fringes running parallel to the interface along the diagonal from the lower left to the upper right. Figure 28 shows a part of the micrograph that has been digitized and saved as a 7000x7000 pixel image.

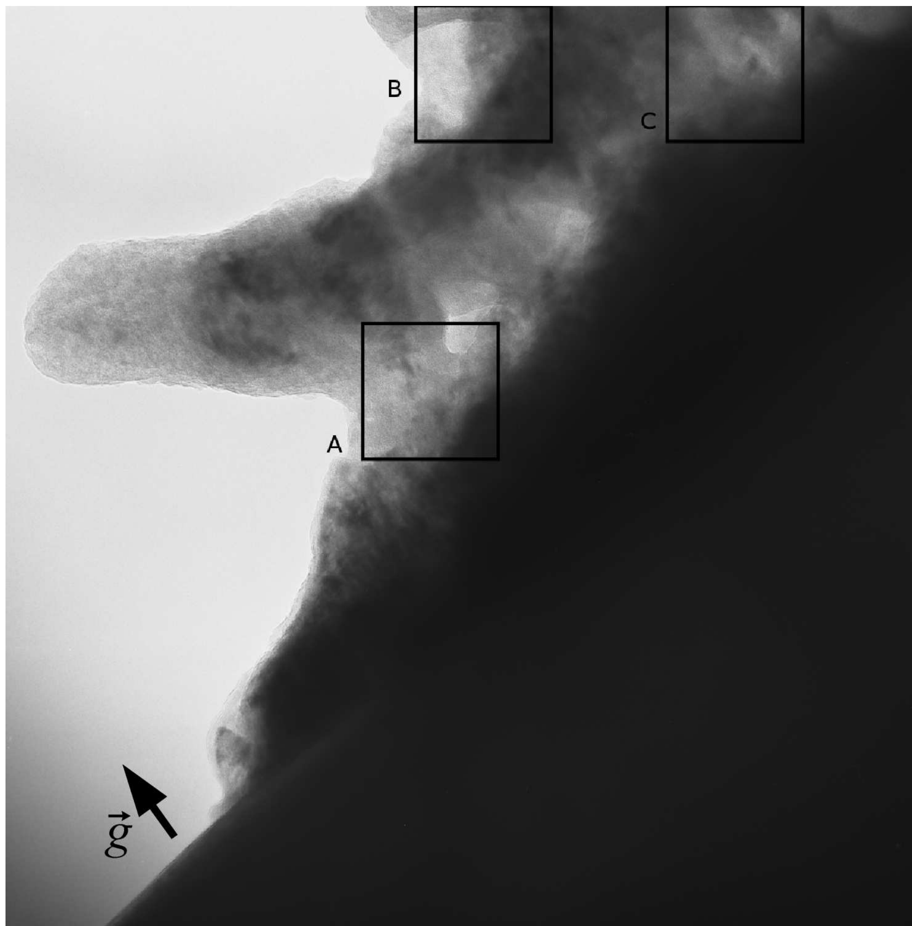


Figure 28: HRTEM image of a heterostructure

The dark area to the lower right contains no fringes, neither does the bright area to the upper left. The areas containing periodicities are spread over the image. The image is large compared to the images used before which leads to several disadvantages: The memory required increases by a factor of 4 every time N is multiplied by 2. Therefore we not only reach the limit of the machine used but also that of 32 bit machines in general. Computation time increases from about 3 seconds for a $N = 1024$ image to about 80 seconds for a $N = 4096$ image. Yet another problem is noise: All the areas containing no fringes contribute to the noise level but no to the signal, we therefore expect a poor SNR if a large input is used. The solution to some of these problems is to analyse the image in smaller regions and combine the results. The regions are marked by the boxes (A,B,C), their size is 1024x1024 pixels. This splitting and combining is reasonable because we use one single micrograph. All the regions that will be examined have been digitized at the same time, so we do not expect any systematic errors like different noise levels or errors from misaligning the micrograph.

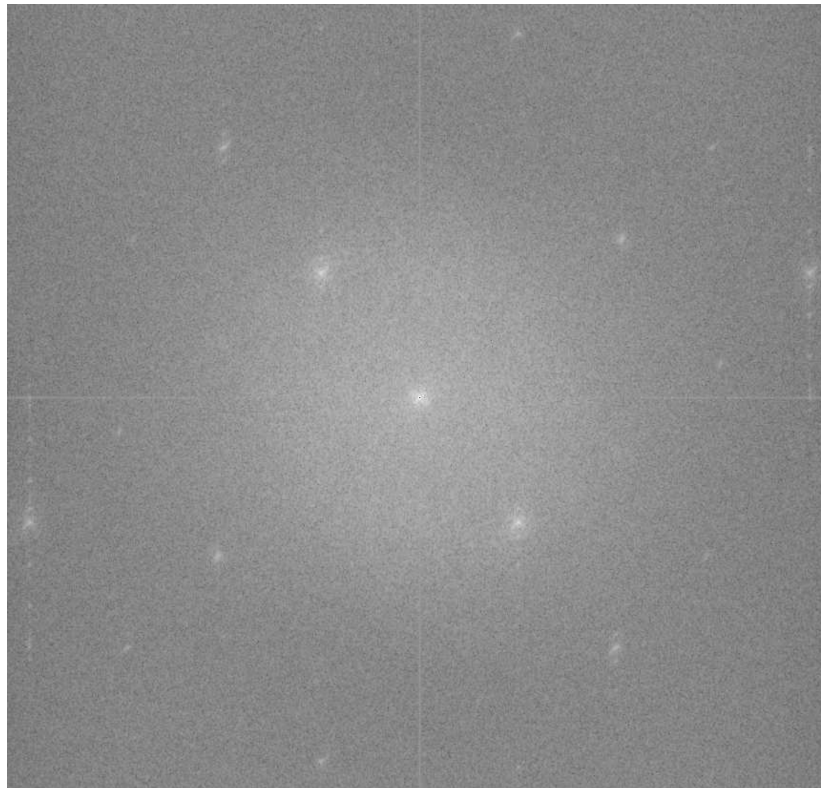


Figure 29: power spectrum

The dark edge at the lower left is the sapphire. As sapphire is hard compared

to the layers that have been deposited, this edge marks the first GaN/sapphire interface. The layers have been removed during sample preparation in this area. If we extrapolate this line into the image we can measure the distance of the marked regions from the interface. The micrograph has been digitized with a resolution of $2400dpi$ and the magnification is about $240k$ times. From these parameters a scale factor of $0.441\mathring{A}/pixel$ is obtained. From that we can determine the distances of the regions from the interface and the material:

A	B	C
$61nm$	$12nm$	$75nm$
AlGaN	GaN	AlGaN

Image 29 shows the PS of the image discussed previously. We can see the spot corresponding to the fringes along the diagonal and the first harmonic. We will not use this harmonic here because the SNR is about $27dB$ while the SNR for the primary spot is $38dB$.

The spacing measurements in area A and C have been performed with an aperture radius of 20, while a radius of 15 was used in area B. Wiener filtering has been applied to all measurements.

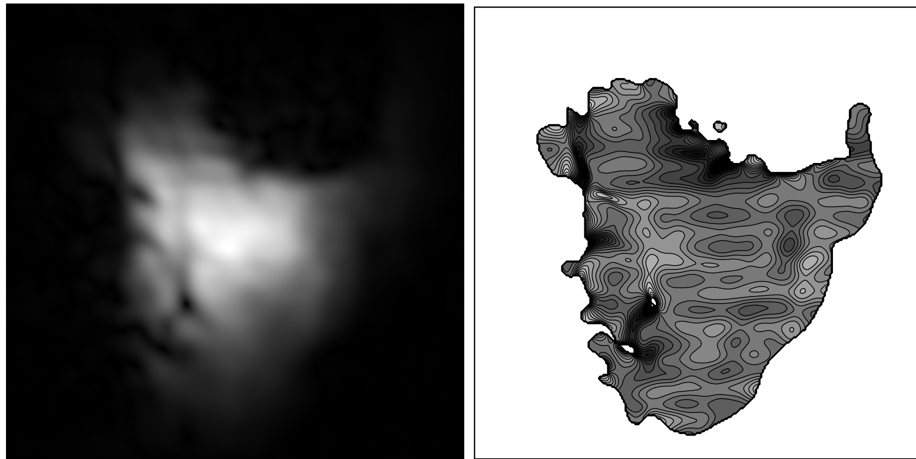


Figure 30: spacing measurements: area A in figure 28

Figure 30, 31 and 32 show the dark field images of the various regions and the results of the spacing measurements. We are not able to see big changes in spacings as in the previous examples. The following range for the spacings can be determined in the three regions:

region	A	B	C
spacing [pixel]	6.04 - 6.13	6.07 - 6.19	6.07 - 6.16

In chapter 3.3 we examined the effect of noise on spacing measurements and

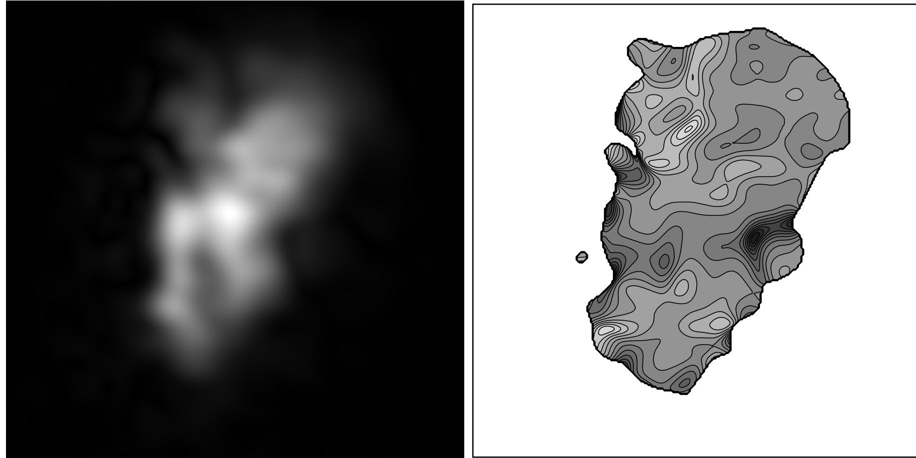


Figure 31: spacing measurements: area B in figure 28

found a gaussian distribution of the measurements. A statistical analysis can be performed in order to give another expression for the spacings in the three different regions. First the average value for each column of data points is calculated, see figure 33. Then the distribution of these mean values is determined using the frequency count feature. The maximum of the obtained distribution is taken to be the average spacing in the corresponding region. The following average spacings can be determined in the different regions:

region	A	B	C
spacing [pixel]	$6.05 \pm 1\%$	$6.14 \pm 1\%$	$6.13 \pm 1\%$

The spacings in all three areas are in the same range. This might be because diffusion occured during sample preparation. Images taken later revealed a super lattice structure, which is another sign of difusion. As the sapphire is much harder than the deposited layers it is more resistant to ion milling than the layers. Therefore thin regions containing the layer structure are hard to find and the thick sapphire absorbs beam intensity which leads to thermal expansion and drift.

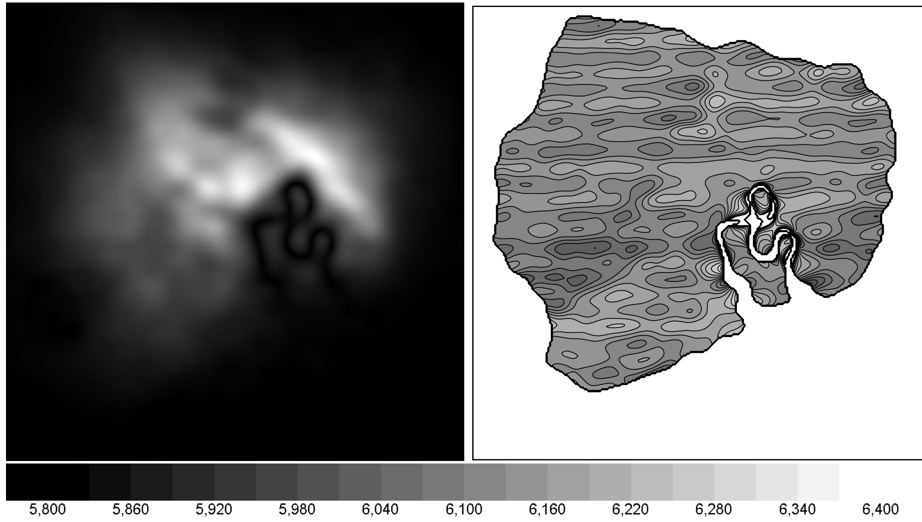


Figure 32: spacing measurements: area C in figure 28

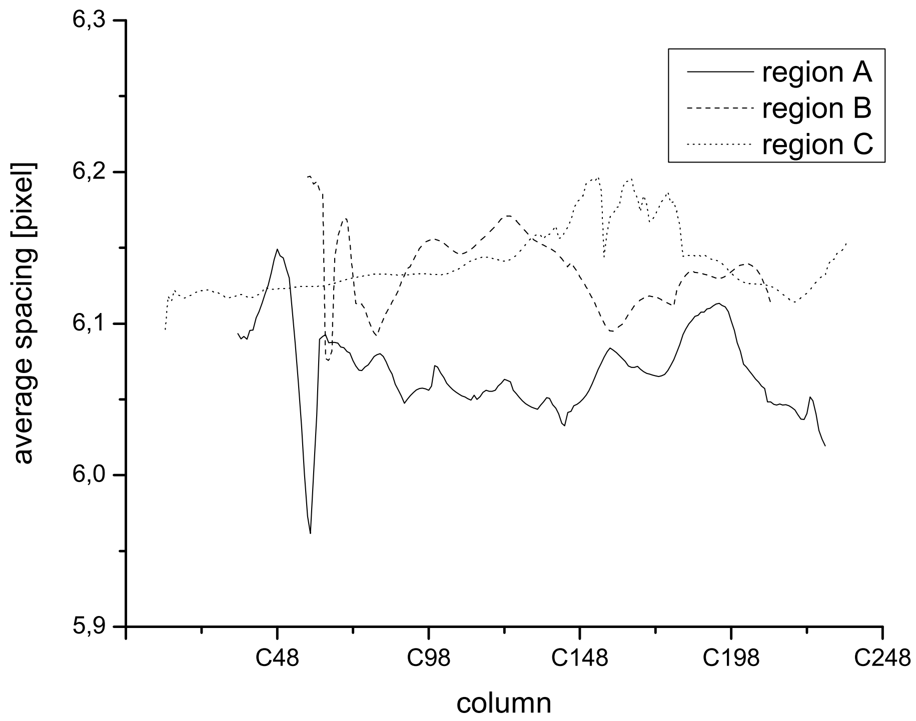


Figure 33: mean spacings in each column

5 Summary, conclusions and outlook

I was able to work out the mathematical foundation of the DF method which allows the phase gradient of the darkfield image to be interpreted as a position dependent wave vector. This interpretation is just an approximation which is valid if the change in the wave vector is small. If the wave vector changes by 5% within 6 pixels the intrinsic error is about 2%. In this case the error introduced by noise is smaller and can be neglected.

If smaller changes are examined the effect of noise becomes more important. I was able to show that the error due to noise increases linearly with the radius of the aperture and the $\hat{N}/\hat{S}$ ratio. The relative error can be determined from simulations using the wide step geometry in combination with added noise.

A new alternate approach to reduce the effect of noise has been developed: The multi spot method uses multiple, in the simple case three, periodicities to determine the spacings more accurately without loosing spacial resolution. This method is very useful if the periodicity that is to be examined is very noisy.

The DF method was applied to simulated images in order to examine the effect of noise and to get a better understanding of the systematic error. When the DF method was applied to experimental images I was able to determine spacings and to give a range for the error. In the case of the three buried quantum dots all three dots and the unstrained area between them could be identified. In the case of the free standing quatum dot the area of constant spacing and the increase towards the top of the dot could be observed. Finally Dan Leopold's sample has been examined. No change in spacings in three different regions could be determined.

I want to give a list of the key features of the DF method that make it complementary to other methods described before:

- The DF method is easy to use as the only parameters the user has to adjust are the center and the radius of the aperture. The input can be read from a larger image.
- The user has the option to add or remove noise, either for simulation purposes or to improve the measurement.
- The multi spot method uses multiple periodicities in different directions to determine the spacings. It adds two more parameters for each spot but it allows to reduce the effect of noise in a very elegant way.
- The method is fast compared to others as it requires only the calculation of two FFTs and some linear algebra. Therefore the method can be used to take a first look at spacings before more advanced methods are applied.

Because of the features listed above the method is suited for implementation

in a TEM equipped with a digital camera. This would allow the user to identify the regions he is interested in more easily. The imaging parameters like defocus could be tuned online to obtain the best results. Using this online image acquisition, the effect of noise could be reduced drastically.

On the other hand ways to reduce the intrinsic error need to be explored. One way to do that in the case of the wide step geometry would be to rotate the micrograph so that the examined spacing runs parallel to one edge of the image. In this case the wave vector would change in one direction only. Then it should be possible to determine its exact value by combining nearby points.

6 Acknowledgment

First of all I want to thank my parents for the continued support throughout my studies, the motivation they gave me and the freedom to follow my mind.

I want to thank Dr. P. Fraundorf for providing this interesting topic, the numerous enlightening dicussions and the support.

I want to thank these people to appreciate their support:

Ginfeng Wang for the introduction at the beginning and for preparing Dan Leopold's sample together with Sam and for taking the images.

Dr. F. Phillipp and Dr. N.Y. Jin-Phillipp, Max-Planck-Institut für Metallforschung at Stuttgart for providing the HRTEM images of the free standing and the buried quantum dots (figure 21 and 25). These examples added a lot to this thesis.

Thanks to the Landesstiftung Badden-Württemberg for the financial support from 08/05 through 02/06.

Finally I want to thank everybody from the International Office at St.Louis and Stuttgart for making the exchange between UMSL and the University of Stuttgart possible.

A Implementation of the DF method

The DF method has been implemented in C++ using the GNU toolchain under Suse Linux 10. It uses the FFTW library (www.fftw.org) to efficiently compute the FFT. The LibTIFF library (www.libtiff.org) is used to read and write TIFF images. The code is portable, it can be compiled in the Cygwin shell under Windows, or under MacOS, if the libraries mentioned before are installed.
The program which performs the DF method is called *DDF* and has the following command line:

DDF Filename N xDat yDat xSoi ySoi AppRad [keywords]

The input denoted by *filename* must be a 16bit TIFF image. It can be bigger than NxN. The keywords are optional while all other parameters are required.

Filename	filename of the input image
N	size of input, should be a power of 2
xDat, yDat	position in the input image where the inputshould be read
xSoi, ySoi	center of the aperture in the power spectrum
AppRad	radius of the aperture

The following keywords can be used in combination.
Grid outputs spacing measurements on a rectangular grid.
GetSpace x y outputs spacing measurements at a certain position.
NoPix does not create any pictures. This saves time when large images are processed.
PSonly aborts the method after the power spectrum has been calculated and saved.
PSav calculates the azimuthal average of the PS.
PSavAp calculates the azimuthal average of the PS inside the aperture.
RemNoise applies a Wiener filter with noise level N.
AddNoise adds noise with a level of N.
Global creates a DF pattern from small regions in Fourier space.

For example

DDF test.tif 1024 123 294 600 200 15 GetSpace 300 600

loads a 1024x1024 image from the file *test.tif*. The input file is bigger than 1024x1024 so the point 123/294 is chosen as the upper left corner of the input. The aperture is placed at 600/200 with a radius of 15. The spacing is measured at 300/600. The spacing and some other information like the maximum inside the aperture is written to the screen. The input image, the PS and the DF image is saved to the disk.

List of Figures

1	single steps of the DF method	10
2	Bravais lattice and the examined spacings	16
3	local spacings for the sharp step geometry	17
4	local spacings for the wide step geometry	18
5	PS of the step geometry, right log scale, left lin scale	20
6	the various 1D window functions. (N=1024)	22
7	different window functions for the input, no window on aperture. (N=1024)	23
8	different windows applied to the aperture, Hamming window for the input. (N=1024)	23
9	(a) 5% sharp ($w = 3$) step (b) 1% wide ($w = 100$) step	25
10	width: 40 (a) 1% step (b) 5% step	26
11	width: 1 (a) 1% step (b) 5% step	27
12	experimental PS	31
13	theoretical PS	31
14	azimuthally averaged PS of images showing a hole in the specimen	32
15	(a) negative 5001, fringes (b) negative 5029, fringes and amorphous material	33
16	(a) statistical distribution of measurements (b) linear increase of error with N/S	38
17	linear increase of error with aperture, $N/S \approx 0.006$ SNR=44bB	39
18	(a) azimuthally averaged PS and noise model (b) extracted signal (c) filter factor $\Phi(\nu)$ (d) filtered signal	41
19	DF images for different settings and noise levels	45
20	PS and azimuthal average of PS	46
21	HRTEM image of 3 buried InP quantum dots	47
22	spacing measurements in GaInP/InP heterostructure	48
23	multi spot measurements: top left 1 spot, right 3 spots; bottom left 5 spots, right 7 spots	50
24	spacing measurements in GaInP/InP heterostructure	51
25	HRTEM image of a free standing quantum dot	52
26	left: 10 spot, right: 10, 01 and 11 spot	53
27	overlay of the input and the spacing map	54
28	HRTEM image of a heterostructure	55
29	power spectrum	56
30	spacing measurements: area A in figure 28	57
31	spacing measurements: area B in figure 28	58
32	spacing measurements: area C in figure 28	59
33	mean spacings in each column	59

References

[1] Originpro, www.originlab.com.

[2] N.W. Ashcroft and N.D. Mermin. *Solid State Physics*. Saunders College Publishing, 1976.

[3] P.E. Batson, N. Dellby, and 0.L. Krivanek. Sub angstrom resolution using aberration corrected optics. *Nature*, 418:270–273, 2002.

[4] R. Bierwolf, M. Hohenstein, F. Phillipp, et al. Direct measurement of local lattice distortions in strained layer structures by HRTEM. *Ultramicroscopy*, 49:273–285, 1993.

[5] Bradley G. Boone. *Signal Processing Using Optics*. Oxford, 1998.

[6] W.J. de Ruijter, R. Sharma, et al. Measurements of lattice-fringe vectors from digital hrem images. *Ultramicroscopy*, 57:409–422, 1995.

[7] K. Du, Y. Rau, N.Y. Jin-Phillipp, and F. Phillipp. Lattice distortion analysis directly from High Resolution TEM images - the ladia program package. *J. Mater. Sci. Technol.*, 18:135–138, 2002.

[8] P. Fraundorf. Digital darkfield decomposition. *arXiv:cond-math/0403017*, 2004.

[9] Brent Fultz and James Howe. *Transmission Electron Microscopy and Diffractometry of Materials*. Springer, 2001.

[10] Martin J. Hÿtch, Jean-Luc Putaux, and Jean-Michel Pénisson. Measurement of the displacement field of dislocations to 0.03å by electron microscopy. *Nature*, 423:270–273, 2004.

[11] M.J. Hÿtch and T. Plamann. Imaging conditions for reliable measurement of displacement and strain in high-resolution electron microscopy. *Ultramicroscopy*, 87:199–212, 2001.

[12] M.J. Hÿtch, E. Snoeck, and R. Kilaas. Quantitative measurement of displacement and strain fields from HRTEM micrographs. *Ultramicroscopy*, 74:131–146, 1998.

[13] N.Y. Jin-Phillipp and F. Phillipp. Defect formation in self-assembling quantum dots of InGaAs on GaAs. *Journal of Microscopy*, 194:161–170, 1998.

[14] N.Y. Jin-Phillipp and F. Phillipp. Strain distribution in self-assembld InP/GaInP quantum dots. *J. Appl Physics*, 88:710–715, 2000.

[15] Craig L. Johnson, Martin J. Hÿtch, and Peter R. Buseck. Displacement and strain fields around a [100] dislocation in olivine measured to sub angstrom accuracy. *American Mineralogist*, 89:1374–1379, 2004.

[16] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery. *Numerical Recipes in C - The art of scientific computing*. Cambridge University Press, 1997.

[17] W.O. Saxon. Images of distorted crystals - recovering the displacement fields. *Synoptics Newsletter*, 1990.

[18] E. Snoek. Quantitative analysis of strain field in thin films from HRTEM micrographs. *Thin Solid Films*, 319:157–162, 1998.