Direct structure inversion from exit waves. Part II: A practical example

A. Wang a,*, F.R. Chen b, S. Van Aert a, D. Van Dyck a

a Electron Microscopy for Materials Science (EMAT), University of Antwerp, Department of Physics, Groenenborgerlaan 171, 2020 Antwerp, Belgium
b National Tsing Hua University, Department of Engineering and System Science, No. 101, Sec. 2, Kuang-Fu Road, 300 Hsinchu, Taiwan

Abstract

This paper is the second part of a two-part paper on direct structure inversion from exit waves. In the first part, a method has been proposed to quantitatively determine structure parameters with atomic resolution such as atom column positions, surface profile and the number of atoms in the atom columns. In this part, the theory will be demonstrated by means of a Au[110] exit wave reconstructed from a set of focal-series images. The procedures to analyze the experimentally reconstructed exit wave in terms of quantitative structure information are described in detail.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

This paper is the second part of a two-part paper on direct structure inversion from exit waves. In Part I [1] the theoretical basis has been given and demonstrated by means of simulations of crystalline structures. It has been shown that structure information such as the surface profile and the number of atoms in the atom columns can be determined quantitatively with atomic resolution. In this part, the method will be applied to an experimental Au[110] exit wave reconstructed from a set of focal-series images.

In an experimentally reconstructed exit wave, information about structure parameters such as, positions of atom columns, number of atoms in an atom column and surface profile, is mixed in the amplitude and phase parts. The goal of this approach is to derive quantitative numbers for these parameters one by one. First, the atom column positions are estimated. Second, the distance from the exit position of each column to the experimentally reconstructed plane is determined, which reveals the surface profile. Next, the exit wave is propagated to the exact exit position for every column. Then, the pixel value at the atom column position of this corrected column wave is plotted in the Argand plot. From this plot, the number of atoms can be determined.

The paper is organized as follows. In Section 2, it is described how the Au[110] images have been acquired and how the exit wave has been reconstructed. Next, in Section 3, the theory is briefly recalled. Subsequently, in Section 4, the reconstructed exit wave is analyzed step by step using the proposed method. In Section 5, the results are described and discussed. Finally, in Section 6, conclusions are drawn.

2. Reconstructed Au[110] exit wave

The experimental high-resolution images were taken with the TEAM0.5 (Transmission Electron Aberration-corrected Microscope) electron microscope installed at Lawrence Berkeley National Laboratory in Berkeley, USA. This electron microscope is capable of producing images with a resolution of about 0.5 Å. It embodies technical advances including ultra-stable electronics, aberration correctors, a monochromator and a bright electron source [2]. Details of the microscope parameters of the TEAM0.5 can be found in [3]. A series of through-focus images was recorded from −9 Å to −169 Å with a focus step of 16 Å, with the microscope operated at 300 keV. Since the Au atoms at the edge were unstable and knocked-off during the recording, only 6 images, ranging from −12 Å to −97 Å, were used to reconstruct the exit wave. An assumption in the focal-series reconstruction method is that the images are taken from the same object at different focus values. This implies that the object should be stable during the recording process. However, even for heavy atoms such as Au, radiation induced diffusion [4,5] has been observed during the recording. This may lead to background noise in the reconstruction. For this reconstruction, an iterative procedure, programmed in MacTempas, has been used [6,7]. Fig. 1 shows the amplitude and phase of the reconstructed exit wave (REW). The experimental data enclosed in the white box have been used in this study. The pixel size is 0.147 Å.
3. Channelling theory with defocus effect

A general outline of the channelling theory, which is used to describe the electron–atom interactions and the effect of defocussing, has been given in Sections 2 and 5 of Part I of this two-part paper. Here, for convenience, only the necessary formulas will briefly be recalled.

Channelling theory [8,9] is often used to describe the dynamical electron scattering when the incident electron beam passes through atom columns, which are aligned parallel to the beam direction. The strong electrostatic potential of the atoms causes the atom column to act as a channel in which the electrons are trapped. Moreover, the electron wave oscillates periodically inside the channel. In a sense each atom acts as a small lens so that the electron wave is periodically focused and defocussed when propagating along the atom column. The period of oscillation is called the extinction distance [10] and is a function of the mass density of the column which is determined by the atom type and the interatomic distance along the column. When passed through an isolated atomic column, the exit wave can sufficiently accurately be expressed as follows [11]:

\[
\psi(r, z) = 1 + c_i \phi_i(r - \mathbf{b}) \left[ \exp \left(-in\frac{E_i}{E_0}kz\right) - 1 \right],
\]

(1)

with \( \phi_i \) the lowest bound state (s-state) of the two-dimensional projected electrostatic potential of the column. The entrance wave (plane wave) is set equal to 1, \( r \) is the two-dimensional vector in the plane perpendicular to the incident beam direction, \( \mathbf{b} \) is the atom column position, \( E_0 \) is the incident beam energy, \( k \) is the reciprocal wavelength and \( z \) is the thickness. Furthermore, \( c_i \) is the excitation coefficient and \( E_i \) is the eigenenergy of the s-state function \( \phi_i \). In most cases, the s-state function can be approximated as a Gaussian function, which is symmetric and peaked at the column positions,

\[
\phi_i(r) = \frac{1}{a\sqrt{2\pi}} \exp \left(-\frac{r^2}{4a^2}\right),
\]

(2)

where \( a \) is the column width and \( r = |r| \).

Exit wave reconstruction methods reconstruct the exit wave at a hypothetical plane close to the exit surface. In principle, however, one does not know the exact position of this plane. Moreover, as schematically shown in Fig. 8 of Part I, due to different vertical positions between neighboring columns, the offset between the plane of reconstruction and the actual exit face is different for every atom column. This offset is called the defocus value. The defocus \( \epsilon \) is the distance that the exit wave of the column has to propagate in order to reach the plane in which the electron wave is reconstructed. When taking this defocus effect correctly into account, Eq. (1) becomes

\[
\psi(r, z) = 1 + c_i \frac{2\sqrt{2\pi}a}{4\pi a^2 + i\epsilon k} \exp \left(-\frac{r^2}{4a^2 + \epsilon^2}\right) \left[ \exp \left(-in\frac{E_i}{E_0}kz\right) - 1 \right].
\]

(3)

As proposed in Section 3 of Part I, the Argand plot is the most convenient and sensitive way to visualize the exit wave. An Argand plot is a complex plane in which the real and imaginary parts of the complex pixel values are plotted in a \((x, y)\) coordinate system [12–14]. As shown in Fig. 2 of Part I, the complex pixel values at the column positions follow a circular locus where the angular increment is proportional to the mass of the atom column. For a known atom column type in a known crystal orientation and provided the resolution and sensitivity of the reconstruction method are sufficient, one can thus count the number of atoms directly from this plot. This is the so-called channelling map [10]. When the column wave is further defocussed, the point follows another circular curve which is called the defocus circle. This is shown in Fig. 9 of Part I. From this figure, it follows that the defocus circles of atom columns containing different number of atoms pass through a common point. By fitting the circular curve, the zero-defocus point \( \epsilon_0 \) can be obtained yielding the defocus value, or vertical position, of each atom column. However, in an experimentally REW, the exact intersection of the defocus circles is difficult to obtain. This is due, for example, to inelastic scattering effects which are not taken into account in the channelling model. However, since we are also able to model the distance between successive defocus points on the defocus circle in the neighborhood of the zero-defocus point, we can estimate this point by model based fitting. The distance between neighboring defocus points follows a cosine curve which is centered around the zero-defocus point which can then be determined from the estimated maximum position.

4. Data analysis

In this section, the procedure to analyze the REW will be described step by step. A flow-chart of the process is given in Fig. 2. Figs. 3–12 show the corresponding results for Sections 4.2–4.6.
4.1. Re-sampling of the reconstructed exit wave

Sampling of the wave directly affects the radius of the defocus circle. If the sampling is insufficient, the radius would be too small to visualize spiral-like sections on the defocus circle, which cannot be used in the analysis. These spiral-like curves are caused by overlapping of neighboring column waves and noise in the background [15]. This will be further explained in Section 4.4. For this reason, the experimentally REW is first resampled using a spline interpolation. A doubling of the sampling turned out to be sufficient.

4.2. Estimation of atom column positions

In order to measure the atom column positions quantitatively, statistical parameter estimation theory is used [16,17]. As such, the precision of the atom column positions can be orders of

---

Fig. 2. Flow chart procedure for analysis of the reconstructed exit wave.

Fig. 3. The crosses show the estimated atom column positions in the phase part of the area of interest.

Fig. 4. (a) Removal of central beam in Fourier space with a Gaussian filter; (b) Amplitude and (c) phase parts of the REW after removing the central beam.

Fig. 5. Defocus curves for four randomly chosen atom columns.

Fig. 6. Amplitude curves obtained from Fig. 5. The solid lines are the quadratic curves fitted to data points ±20 Å from the maximum points.
magnitude better than the information limit of the electron microscope provided the number of imaging particles (electron dose) is sufficient and the object is sufficiently stable [18]. To apply statistical parameter estimation theory, a proper parametric model is required. In this example, we use the channelling theory as a starting point to describe the interaction between the electron beam and the Au atoms. However, the width of the s-state of a Au atom column is smaller than the resolution of the electron microscope. Thus, the column wave is blurred by several factors such as defocus, thermal vibrations and residual aberrations. As a consequence, the column wave approaches a Gaussian function. The REW can then be modeled as a superposition of Gaussian peaks. Both the real and imaginary parts (or in other words, the amplitude and phase parts) of the REW are taken into account. The atom column coordinates and the width are fixed to be the same while the height of the atom column and the background are different for the real and imaginary parts. The expectation model $f_{kl}$ at pixel $(k, l)$ at the position $(x_k, y_l)$ is thus given by [19]

$$f_{kl}(\theta) = \left[ \sum_{n_c} \exp \left( -\frac{(x_k-x_{nc})^2 + (y_l-y_{nc})^2}{\rho_n^2} \right) \right] \left[ \sum_{n_m} \exp \left( -\frac{(x_k-x_{nm})^2 + (y_l-y_{nm})^2}{\rho_m^2} \right) \right].$$

(4)

where $n_c$ is the total number of atom columns and $\beta_{x_n}$ and $\beta_{y_m}$ are the $x$- and $y$-coordinates of the $n$th atom column, respectively. The parameters $x_{nc}$ and $y_{nm}$ are the constant backgrounds and $x_{nc}$ and $y_{nm}$ are the heights of the Gaussian peak for the real and imaginary parts, respectively. The width $\rho_n$ can be different for every atom column type due to a different amount of blurring. The unknown parameters $\theta = (\beta_{x1}, \ldots, \beta_{xN}, \beta_{y1}, \ldots, \beta_{yN})$ are estimated by minimizing the uniformly weighted least squares criterion

$$\tilde{\theta} = \arg\min_{\theta} \sum_{i=1}^{K} \sum_{k=1}^{L} \left[ (w_{kl}^R - f_{kl}^R(t))^2 + (w_{kl}^I - f_{kl}^I(t))^2 \right].$$

(5)

with $w_{kl}$ the restored complex value at the pixel $(k, l)$. The total number of pixels in $x$- and $y$-directions are given by $K$ and $L$, respectively. In [16,19], it is shown that the uniformly weighted least squares estimator is identical to the maximum likelihood estimator provided that the pixel values are independent and identically normally distributed about the expectations. The maximum likelihood estimator is known to have optimal statistical properties. The atom column positions are shown by crosses in Fig. 3. Note that the atom columns near the boundary of the region are not used for practical reasons which will be explained in the following steps.

4.3. High pass filter

As mentioned in Section 3 already, the common point of the defocus circles, corresponding to the entrance wave, may be difficult to define. There are many reasons for this. First, the focal-series reconstruction methods are only reliable for the high frequencies but not for low frequencies such as a background [20]. Second, the background also contains inelastically scattered electrons that are eliminated from the coherent image formation. For these reasons, the background is removed using a Fourier filter as shown in Fig. 4(a). A mask is applied to all the Fourier frequencies except the central beam. The edge of the mask is smoothed using a Gaussian function with a half width at half maximum equal to five pixels in order to avoid edge artifacts. The amplitude and phase parts of the REW after removing the central beam are shown in Fig. 4(b) and (c), respectively.

4.4. Propagation of reconstructed exit wave

Since one does not know beforehand the distance from the REW to the exact exit positions of the atom columns, the REW is propagated in a wide defocus range. As such, one can determine the exact vertical positions of the REW. The REW is computation- ally convoluted with the propagator for different defocus distances $z$. Artifacts near the boundary of the region are unavoidable because part of the wave is calculated beyond the boundary. Therefore, the atom columns near the boundary are not taken into account. In this example, the REW is defocused from $-100$ Å to $+201$ Å with a defocus step of $1$ Å. Defocus curves for four randomly chosen atom columns are shown in Fig. 5. This figure shows that the pixel at the column position follows a circular curve for every atom column. Moreover, for every defocus curve, there is a section where the curvature is large compared to other parts where the points follow spiral-like curves in addition. This is
due to the effect of neighboring columns. In [15], it is shown that for defocus values close to the zero-defocus, that is when $|e - e_0|$ is small, the s-state wave dominates over the defocus term. For defocus values further away from the zero-defocus, the atom column waves spread out far and depend on the configuration of the neighboring atom columns complicating the analysis. As a result, one needs to choose a suitable section on the defocus curve to obtain $e_0$. How this can be done will be explained in the following step.

### 4.5. Amplitude curve

In the last part of Section 5 of Part I, a method to obtain the zero-defocus point on the defocus curve without the use of a common point was derived. It turned out that, distances between neighboring points on a defocus curve follow a cosine. However, in practice, it turns out that these distance values are sensitive to noise. Intuitively, we see that the distance to the origin, corresponding to the amplitude of the wave, is maximal at the

![Fig. 9](image_url)
zero-defocus point. This suggests to plot the amplitude as a function of defocus. From Eq. (16) in Part I, the amplitude is equal to \( 2 \sqrt{1 + \chi^2} \) where \( \chi \) is proportional to \( \varepsilon \). As a function of defocus, the amplitude thus follows a polynomial function that peaks at \( \chi = 0 \), or \( \varepsilon = 0 \), which is the exact exit position. This will be called the amplitude curve. In conclusion, the zero-defocus point can be determined by finding the maximum of the amplitude curve. Fig. 6 shows the amplitude as a function of defocus for the same columns as in Fig. 5. It is shown that in the section where the curvature of the defocus curve is the largest, the amplitude curve gives a clear global maximum with a smooth curve close to that point. Other local maxima result from smaller spiral-like curves in the defocus curve. From the amplitude curve, a suitable section can easily be decided, which in this case is around \( \pm 20 \) Å from the global maximum. By fitting a 2nd degree polynomial (quadratic function) to this section, the maximum, corresponding to the zero-defocus point, can be determined with sub-angstrom precision. This is also shown in Fig. 6.

4.6. Defocus map and Argand plot

After the maxima are derived from the amplitude curves, the defocus value for every atom column can be calculated. The defocus value is the distance of propagation from the zero-defocus point to the experimental REW. Fig. 7 shows the defocus value for every atom column. Next, the REW is 'back-propagated' for every atom column using the obtained defocus value in order to retrieve the correct pixel value at the true exit position. The correct pixel values are then plotted in an Argand plot as shown in Fig. 8. These figures will be discussed in detail in the following section.
The defocus map, presented in Fig. 7, shows an increase of defocus values for atom columns further away from the edge. This indicates that the exit surface of the sample follows an upward inclined plane starting from the edge. After correction of the defocus, the spread of the pixel values in the Argand plot is reduced. This is shown in Fig. 8. Correcting the defocus will improve the accuracy of the number of atoms in the column and may avoid misinterpretation in the number of atoms in a column. The result indicates that the area of interest is rather uniformly thin. However, even after correction of the defocus, the clusters may avoid misinterpretation in the number of atoms in a column.

In order to study the exit face geometry (surface profile), we choose three small areas perpendicular to the edge to demonstrate single atom increment. These areas are shown in Fig. 9(a)–(c). The defocus values for the atom columns in the corresponding areas are shown in Fig. 9(d)–(f). Moreover, Fig. 9(g)–(i) show the phase values of the atom columns after correction of the defocus values. The phase values are nearly constant proving that the thickness of the area of interest is quite uniform. From Fig. 9(d)–(f), it can be concluded that the atom columns close to the edge give a sharp increase of the defocus values and after the 4th row the increase is fairly linear. The reason for this sharp increase in the first few rows is not clear at this moment. A possible reason may be that the edge columns are loosely bounded compared to other columns inside the lattice, so that the edge columns are not stable during the recording which may introduce errors in the exit wave reconstruction. The inner rows show a linear increase of defocus values corresponding to an angle of about 25°. Based on the results shown in Fig. 9, the authors propose a structure which is schematically shown in Fig. 11 when the structure is viewed from the side. This seems to indicate that it is a (111) slab viewed along the [110] direction. However, more experimental studies, for example, an extra analysis from another zone axis, should be made in order to confirm this proposed structure. Fig. 10(a)–(c) show the Argand plots for the three areas in Fig. 9(a)–(c), respectively. However, at this stage, the radius of the clusters cannot be further reduced.

Possible reasons are residual aberrations from the microscope, fluctuations resulting from data processing, the instability of the atoms during recording, etc. All these factors strongly influence the pixel value at the atom column position. Consequently, defocus curves and amplitude curves will be affected. Fig. 12 shows the Argand plot of the three areas together. In Fig. 12(b), a histogram of the data is inserted. This is done by first fitting a line to the data using linear regression and then projecting the data onto this line. From this figure, it is believed that there are three consecutive numbers of atoms in the columns. In addition, for this Au example, the thickness of the sample is expected to be less than one extinction distance (fewer than 10 atoms in a column), within which the absorption effect does not dominate over the s-state wave function.

In Section 4.5, the zero-defocus point was determined from the maximum of the amplitude curve. Moreover, the zero-defocus point could also be determined based on Eq. (18) of Part I

\[ \chi = \tan \frac{\theta}{\pi} \]  

where \( \chi = \frac{\Delta \ell}{\pi 4 \ell} \) and \( \theta \) is the angle between two neighboring defocus points drawn from the center of the defocus circle. This equation suggests that the phases of the defocus points follow an arc-tangent relationship with the defocus \( \varepsilon \). Hence, the zero-defocus point can be determined from the turning point of the arc-tangent curve. In this example, the values of the zero-defocus are the same using both procedure. However, in case of very light and thin atom columns, for example in graphene, the phases of the defocus points may still lay in the nearly linear region close to the turning point even when defocusing over a wide range. This makes the turning point difficult to define. Additionally, the amplitude curve will not have a clear global maximum. The reason for this is that the s-state wave of such atom columns cannot be approximated by a Gaussian function. Accordingly, the defocus curve may not follow a circle. For this type of atom columns, other methods will be proposed in a forthcoming paper [21].

The authors are aware of the possible presence of surface distortion which may either change the interatomic distance in the z-direction or positions in the x–y plane close to the surface. This may lead to a reduction of lattice symmetry as shown in [22] in the case of h-BN. It was shown that such distortions can cause significant phase shifts resulting in a misinterpretation of the defocus value. In order to rule out this possibility, we have done a simple multislice simulation [23] for two Au atoms in a column as illustrated in Fig. 13(a). The phase change due to surface distortion is given in Fig. 13(b). The interatomic distance between two Au atoms in the [110] column direction varies from 1 Å to 4.5 Å. The dotted curve shows the phase changes assuming only a change of interatomic distance \( d \) whereas the squared curve shows an additional 10pm lateral distortion in the x–y plane of the top atom. As expected, the phase value decreases with
increasing interatomic distance and also decreases with increasing lateral displacements. Moreover, Fig. 13(c) shows the Argand plots of the defocus waves with and without surface distortion. Indeed, surface distortion results in phase changes and may be confused with the defocus value since both are like blurring functions to the column wave. However, although distortion of Au(110) at the surface may be as deep as three layers \cite{24,25}, the expansion or contraction of the atom displacements compared to the bulk values is less than 20pm in the z-direction (indicated with vertical dashed lines in Fig. 13(b)) and less than 10pm in the x–y plane. Thus, the phase changes due to surface distortion are only marginal in this example.

One could also argue whether the profile at the entrance surface plays a role. However, the surface profile at the entrance face only gives an arbitrary phase shift for the incoming wave of the atom column leading to a relative rotation in the Argand plot. This has been tested by means of simulations but not presented here. The proposed method to obtain structure information uses the fact that the atom column wave at the exit position is peaked at the column position. The peakiness in itself is not altered by a constant phase shift. Only in case the exit waves of neighboring columns become important, the relative phase shift between columns, which is affected by the roughness of the entrance face, can play a role. Moreover, once the surface profile at the exit face and the number of atoms in the columns are obtained, one automatically obtains the surface profile at the entrance face.

6. Conclusions

In this second part of the two-part paper, an experimental exit wave was used to demonstrate the application of the theoretical method proposed in the first part. The REW was analyzed in several steps. First, the REW was interpolated in order to obtain a finer sampling and the atom column positions were obtained by applying statistical parameter estimation theory. Then, the central beam was removed with a high-pass filter applied in the Fourier space. Next, the REW was propagated over a large defocus range. The amplitude was plotted as a function of defocus. The defocus region around the maximal amplitude was then used for further analysis. By fitting a quadratic function to the amplitude curve or an arc-tangent function to the phase curve we could determine the zero-defocus point. The defocus values for every atom column could then be determined and a defocus map could be drawn. Later, for each atom column, the REW was back-propagated to the true exit position. Subsequently, the pixel value at the true exit position was plotted in an Argand plot. From these results, it could be concluded that the sample is rather thin and is inclined upwards with respect to the horizontal plane over about 25°. The clusters in the Argand plot are not clearly separated and the numbers of atoms in the columns cannot be well defined yet. This may be due to residual aberrations, local tilt and statistical noise which all affect the pixel value at the column position. Further refinement of the experiment and the analysis are needed to quantify these affects to the level of a single atom.

Acknowledgments

The authors would like to thank Dr. C. Kisielowski for providing the experimentally reconstructed exit wave. A. Wang, S. Van Aert and D. Van Dyck are grateful to the Fund for Scientific Research—Flanders (FWO) (project nos. G.0188.08, G.0064.10N and G.0393.11). A. Wang and F.R. Chen would like to thank to the National Science Council of Taiwan (project no. 96-2628-E-007-017-MY3).

References


