Numerical correction of anti-symmetric aberrations in single HRTEM images of weakly scattering 2D-objects

Ossi Lehtinen a,*, Dorin Geiger a, Zhongbo Lee a, Michael Brian Whitwick b, Ming-Wei Chen b, Andras Kis b, Ute Kaiser a

a Central Facility for Electron Microscopy, Group of Electron Microscopy of Materials Science, University of Ulm, 89081 Ulm, Germany
b Electrical Engineering Institute, Ecole Polytechnique Federale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland

A R T I C L E   I N F O

Article history:
Received 14 July 2014
Received in revised form 5 September 2014
Accepted 26 September 2014
Available online 27 October 2014

Keywords:
HRTEM
Residual aberrations
Correction
Numerical image processing

A B S T R A C T

Here, we present a numerical post-processing method for removing the effect of anti-symmetric residual aberrations in high-resolution transmission electron microscopy (HRTEM) images of weakly scattering 2D-objects. The method is based on applying the same aberrations with the opposite phase to the Fourier transform of the recorded image intensity and subsequently inverting the Fourier transform. We present the theoretical justification of the method, and its verification based on simulated images in the case of low-order anti-symmetric aberrations. Ultimately the method is applied to experimental hardware aberration-corrected HRTEM images of single-layer graphene and MoSe2 resulting in images with strongly reduced residual low-order aberrations, and consequently improved interpretability. Alternatively, this method can be used to estimate by trial and error the residual anti-symmetric aberrations in HRTEM images of weakly scattering objects.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

Thanks to the practical realization of hardware aberration correction (AC), the point resolution of high-resolution transmission electron microscopy (HRTEM) is not limited by the positive spherical aberration (Cs) any longer [1–3], and materials can now be routinely imaged at atomic resolution with state-of-the-art instruments even at low acceleration voltages [4,5]. Once the strong Cs contribution is removed, other aberrations such as three-fold astigmatism (A2) and coma (B2) become measurable and can be corrected [3] down to a limit imposed by the measurement accuracy and the adjustment precision of the correcting elements. With advanced techniques, such as electron holography [6–9] and exit wave reconstruction from a focal series [10–14] direct information on the specimen, not obscured by the inevitable residual aberrations, can be acquired. However, these techniques require special arrangements/procedures during the acquisition of the images, and thus cannot be applied to a single HRTEM image post-situ.

The wave aberrations distort the information transferred from the specimen to the imaging device, such as a CCD camera, making interpretation of the acquired images difficult if not sometimes impossible [15]. In a corrected microscope this problem is to a large extent overcome, as is evident from the nearly aberration-free images of today. In the context of this paper, the word nearly should be emphasized, however. The aberration coefficients can be tuned down to zero only with some accuracy. Thus, in practice residual aberrations are always present even in corrected instruments [16,17]. Moreover, the instruments tend to drift away from the corrected state over time, and one can expect stronger residual aberrations as time passes after tuning the corrector [18]. Residual aberrations can be also of benefit, like in the case of Cs, as non-zero Cs is required for optimal phase contrast transfer [3].

On the other hand, residual aberrations such as A2 and B2, which can be present in the order of 100 nm after correction, can lead to undesired artefacts in the images, and the effect of these can be anything from a minor nuisance to complete misinterpretation of the image. Examples of such effects are presented in Fig. 1. In the first panel, an experimental aberration-corrected HRTEM (AC-HRTEM) image of a point defect in single-layer graphene can be seen. The presence of residual aberrations in the image is clear from the strongly non-hexagonal appearance of the graphene lattice. This effect can be attributed to the three-fold astigmatism A2, which has three axes of symmetry in the image plane (see Fig. 1c for a visualization); depending on the relative orientation of A2 to the graphene lattice, one of the two sublattices of graphene can produce a higher contrast than the other, resulting in the observed non-hexagonal pattern. The atomic structure of the defect can still be deduced from the experimental image, but the image is far from optimal.
In the second panel, an experimental AC-HRTEM image of the so-called mirror-twin boundaries in a single-layer MoSe$_2$ sample can be seen. In an aberration-free image the Se sites produce higher contrast as compared to the Mo sites, as there are always two Se ($Z=34$) atoms at the same location as opposed to single Mo ($Z=42$) atoms. However, in the recorded image the lattice has a completely different appearance at the different sides of the linear boundaries (compare, for example, the areas denoted by the two arrows). On one side the second sublattice has a much higher contrast, but on the other side the sublattices have almost equal contrast, resulting in the hexagonal appearance of the lattice. This effect can again be attributed to residual $A_2$, which is oriented matching the sublattices of the crystal with trigonal symmetry (see Fig. 1 panel c). On one side of the boundary the Se sites are further enhanced, increasing the contrast difference of the Mo and Se sites. However, when crossing the boundary, the Mo and Se sites are interchanged (see Fig. 1 panel d), and now the $A_2$ has the effect of reducing the contrast difference resulting in nearly equal contrast for the Mo and Se sites. Evaluation of the contrast could lead one to deduce that there are two different materials on the opposite sides of the boundary, which in fact is not the case. We would like to emphasize that these artefacts are not introduced by the microscopist, but on the other side the sublattices have almost equal contrast, resulting in the hexagonal appearance of the lattice. This effect can again be attributed to residual $A_2$, which is oriented matching the sublattices of the crystal with trigonal symmetry (see Fig. 1 panel c). On one side of the boundary the Se sites are further enhanced, increasing the contrast difference of the Mo and Se sites. However, when crossing the boundary, the Mo and Se sites are interchanged (see Fig. 1 panel d), and now the $A_2$ has the effect of reducing the contrast difference resulting in nearly equal contrast for the Mo and Se sites. Evaluation of the contrast could lead one to deduce that there are two different materials on the opposite sides of the boundary, which in fact is not the case. We would like to emphasize that these artefacts are not introduced by the microscopist. Under our experimental conditions, the $A_2$ coefficient is in the order of 100 nm, which is a typical state of correction, after tuning, in our microscope (FEI Titan 80–300, operated at 80 kV).

As mentioned above, numerous methods for acquiring truly aberration-free images in TEM have been proposed and implemented in practice. The common denominator of these methods is that they aim at recovering the object wave function, which then can be manipulated much more freely and interpreted more directly than a simple intensity image [8,9,12–14,19]. In electron holography [6,7], an interference pattern between the object wave and a reference wave is generated with the help of a bi-prism, and direct information on the object wave can be recovered numerically from this pattern. The $C_{1}$ corrector improves the phase-detection limit of electron holography significantly [8]. Another group of methods is in-line holography, i.e., exit wave reconstruction [10–14,20], where typically a series of images at equally spaced foci is acquired and subsequently the object wave is deduced iteratively. After the ‘raw’ electron wave is recovered, the influence of any residual aberrations can be corrected [19,8,9]. A common drawback of these methods, however, is that they all require special arrangements and/or procedures during acquisition of the images, and hence cannot be applied to a single HRTEM image during post-processing. Such a situation can be quite frustrating if the microscopist has managed to capture an elusive target during imaging, but the resulting image is of poor quality.

Here, we present a method for numerically removing the effects of a specific group of geometric aberrations (the anti-symmetric ones, as discussed later) after acquisition in single HRTEM images of weakly scattering objects. The requirement of the investigated object to be a weak scatterer is a strong one, and precludes a large number of typically studied materials from the applicability range of the method. However, one important and intensively studied group of materials, that is, the recently discovered class of 2D materials with graphene as the prime example, do approximately satisfy this condition at typically used acceleration voltages such as 80 kV [21], making the method potentially useful in a large number of studies.

In analogy to the phase reconstruction methods (focal series and off-axis holography) where residual aberrations are removed from the wave function after the reconstruction, we remove the contribution of the residual anti-symmetric aberrations using the fast Fourier transform (FFT), but instead of the full wave function, we work on a single image, which is possible in the case of a weakly scattering object, that is, when the linear imaging theory is valid [22]. Because the weakly scattering objects we investigate are 2D-objects like graphene, the tilt angles are usually small and the effect of object tilt will be neglected in the first approximation. We want to make clear that the effect of the symmetric aberrations, such as defocus or $C_{1}$, cannot be remedied by our method. This is clear already from the fact that the symmetric aberrations lead to loss of information at certain frequencies (the zero crossing points of the contrast transfer function), which of course cannot be recovered. We find our method to be particularly useful for removing the effects of residual aberrations in distorted AC-HRTEM images of weakly scattering 2D-materials, where information is available on how the image should look like, as demonstrated with the examples of graphene and MoSe$_2$.

2. Methods

First, we present the theoretical background and justification of our method. We follow closely the notation of Ref. [23].
Assuming coherent illumination and excluding the damping envelopes from the treatment, the wave function \( \psi'_o(x) \) at the image plane at a distance \( z \) from the object along the optical axis can be written as

\[
\psi'_o(x) = \psi_o(x) \otimes p_2(x) = (1 + \phi(x)) \otimes p_2(x),
\]

where \( \psi_o(x) \) is the wave at the exit surface, \( \phi(x) \) is the scattered wave, \( p_2(x) \) is the lens transfer function and \( \otimes \) denotes convolution. The image intensity \( I_z(x) \) is the square modulus of the wave at the detector plane, and its Fourier transform \( I_z(u) \) can be written as

\[
I_z(u) = \mathcal{F}[\psi'_o \cdot \psi_o^*] = \mathcal{F}[1 + \phi \otimes p_2 + \phi^* \otimes p_2^* + (\phi \otimes p_2)(\phi^* \otimes p_2^*)].
\]

In the case of a weakly scattering object (that is, when the weak phase object approximation is valid), an approximation can be made here [22]. That is, the last term on the second line of the equation can be deemed to be small, which is in effect a transition from the general non-linear imaging theory to the special case of linear imaging theory. Fortunately, 2D-materials such as graphene fulfill this condition at typical acceleration voltages such as 80 kV [21] and, as demonstrated by simulations and experiments below, the method can be applied on images of such materials. Eq. (2) now becomes

\[
I_z(u) = \delta(u) + \phi(u) \cdot e^{-i\phi(u)} + \phi^*(u) \cdot e^{i\phi(u)},
\]

where \( \phi(u) \) denotes the Fourier transform of \( \phi(x) \). The Fourier transform of \( p_2 \) is written in the explicit form \( e^{-i\phi(u)} \). The information about the microscope aberrations are encoded in the wave aberration function \( \chi(u) \).

Now, \( \chi(u) \) can be represented as a power series where each term represents a specific aberration. The terms of the series can be divided into symmetric \( \chi_s(u) = \chi(-u) \) and anti-symmetric \( \chi_a(u) = -\chi(-u) \) parts. Notably, the defocus and \( C_1 \) belong to the symmetric group and \( A_2 \) and \( B_2 \) to the anti-symmetric group. Thus, \( I_z(u) \) can be rewritten as

\[
I_z(u) = \delta(u) + \phi(u) \cdot e^{-i\phi(u)} e^{-i\chi_s(0)} + \phi^*(u) \cdot e^{i\phi(u)} e^{-i\chi_a(0)}.
\]

From this equation it is evident that the effect of all the anti-symmetric aberrations can be removed from the image by simply multiplying \( I_z(u) \) by \( e^{i\chi_a(0)} \) and calculating the inverse Fourier transform. That is, a corrected image \( I^c(x) \) can be generated by

\[
I^c(x) = |\mathcal{F}^{-1}[e^{i\chi_a(0)} \cdot \mathcal{F}[I_z(x)]]|.
\]

The problem of correcting the residual aberrations is two-fold, however. So far we have shown that, in principle, the anti-symmetric aberration contribution in a HRTEM image of a weakly scattering object can be freely manipulated in post-processing. However, in order to do actual correction, the aberration coefficients and orientations need to be determined. This is a shared challenge with all the methods aiming to correct aberrations in TEM, and of course one is limited by the accuracy at which the aberrations can be measured. Information on the residual aberrations is offered by the corrector based on a series of tilted-beam images [3,24]. Other methods for measuring the residual aberrations, e.g., from a focus/tilt series [25,26] or even a single HRTEM image [27] have been presented in the literature. The aberrations could, in principle, be removed with the corrector once determined, but, e.g., due to drift of the corrector state during searching of an area of interest and image acquisition [18] and finite precision of the corrector adjustment, typically residual aberrations in the final images are inevitable.

Post-acquisition correction by the hardware corrector is of course not possible, and one needs to find new strategies for fine-tuning the existing images if necessary. One alternative is to look at a feature of an image (e.g., the pristine graphene lattice) and find the coefficients which result in the expected appearance for an aberration-free image by trial and error. Great care should be taken with this approach, however, as the problem can be underdetermined. For example, in our example case of three-fold \( A_2 \) astigmatism and MoSe\(_2\) (see e.g. the illustration in Fig. 1 panel c) the same sublattice imbalance can be introduced by different combinations of amplitude and relative orientation of the \( A_2 \) and the MoSe\(_2\) lattice. A conservative approach in such a situation is to find the orientation where the smallest possible amplitude corrects the visible effect. If the actual orientation of the real residual aberration would be different from this, some residual aberration would remain (the new residual aberration is the vector sum of the real aberration and the correcting aberration), but the situation would still be improved and the aberration would not be ‘over corrected’. The situation is further complicated, if multiple dominant residual aberrations are present. However, often the third-order anti-symmetric aberrations, that is \( B_2 \) and \( A_2 \) are dominating (assuming that defocus and first-order astigmatism are properly tuned). In the special case of 2D-materials, which tend to have a three-fold symmetry, the three-fold symmetric \( A_2 \) has a distinctive effect on the appearance of the lattice, and thus can be distinguished from the effects of \( B_2 \). Thus, even if the aberration post-correction is not ideal, a significantly better result can be achieved in suitable situations, as shown below by experiments and simulations.

All the experimental images presented here were acquired using an FEI Titan 80–300 microscope equipped with an image-side hexapole aberration corrector [1]. The microscope was operated at 80 kV and a reduced extraction voltage of 2 kV, in order to reduce the energy spread of the electron beam. The \( C_4 \) was corrected to \( < 20 \) μm and the \( A_2 \) and \( B_2 \) were corrected to the order of 100 nm (unless stated otherwise).

The graphene samples were produced by the chemical vapor deposition method (acquired commercially from Graphenea S. A.) and transferred on Quantifoil TEM-grids by the procedure described in Ref. [28]. The defect in Fig. 1a is produced by carbon deposition, as will be described in detail in a separate paper [29]. The MoSe\(_2\) sample was produced by molecular beam epitaxy and transferred on to Pelco holey silicon nitride TEM grids (a detailed study on this sample will be presented also in a separate article [30]).

The image simulations were conducted using the QSTEM software package [31], using \( C_1 = 20 \) μm and a focus spread of 6 nm. The simulated structure consisted of 149 Mo and 288 Se atoms, and the simulation cell size was 6 nm × 6 nm with a sampling of 0.03 Å/pixel (the panels in Fig. 2 are cropped from the top right).

3. Results

To verify our approach in the case of the low-order \( A_2 \) and \( B_2 \), we first conducted image simulations, where one has precise control over the aberrations influencing the image. In Fig. 2a, a MoSe\(_2\) target with an embedded triangular mirror domain is simulated with \( A_2 = B_2 = 0 \). The domain boundaries are visible, and the lattice inside and outside looks identical except for the reversed order of the darker and brighter appearing sublattices. In order to emulate the effect observed in the experimental image of 1, a HRTEM image was simulated with \( A_2 = B_2 = 100 \) nm, and 30° as the azimuthal orientation of both of the aberrations (Fig. 2b). With these parameters the lattice appears completely different inside and outside the mirror domain. This image was then processed according to Eq. (5) using the known \( A_2 \) and \( B_2 \) coefficients and orientations, and the resulting image is shown in panel c of Fig. 2. The aberration-free appearance of the image is restored by the correction procedure, as predicted.

We also tested the procedure in more extreme conditions, that is, with \( A_2 = B_2 = 10 \) μm. This does not represent typical experimental conditions, but should be considered as an artificial test with very strong aberrations. The simulated image is shown in
Fig. 2d, where the whole image is scrambled by the aberrations and interpretation of the image is practically impossible. However, after applying Eq. (5) with the known parameters, the image is restored to an aberration-free state. Here, it should be pointed out that close inspection of the corrected image reveals some features which are absent in the reference image (panel a), e.g., in the vacuum area. This is due to some information being displaced outside the image frame by the (in this case) large point spread function/image delocalization induced by the large aberrations, which then is missing also in the corrected image. Generally speaking, one should avoid interpreting the corrected image within a margin at the image edges with a width determined by the point spread function/image delocalization.

In order to verify our method also in an experimental situation, we took the AC-HRTEM images of Fig. 1, and applied Eq. (5). In the case of graphene, we iteratively found the amplitude and orientation of \( A_2 = 150 \text{ nm/}17^\circ \) which resulted in the removal of visible asymmetry between the two sublattices of graphene. The corrected experimental image is presented in Fig. 3a, with the original image shown as inset. The graphene lattice has now the expected hexagonal appearance, and interpretation of the image is made more straight-forward. For the case of MoSe\(_2\), \( A_2 = 75 \text{ nm/}15^\circ \) was found to restore the expected appearance of the lattice on both sides of the mirror-twin boundaries in the experimental image.

An interesting observation in the case of MoSe\(_2\) is that the point defects are actually more visible in one of the domains in the original (non-corrected) image, and also the boundaries are easier to locate due to the different appearances of the lattice on the opposite sides of the boundaries. This bears resemblance to the well known effect of moving away from the ‘optimal’ Scherzer focus in order to make point defects more visible to the eye (see, e.g. Refs. [16,32]). No new information is added to the image by the anti-symmetric aberrations, but features are made more evident to the eye. A clear advantage of this approach is that the aberrations can subsequently be corrected with no information losses, as the anti-symmetric aberrations introduce only phase shifts and the phase contrast transfer function is not altered. If the presence of anti-symmetric residual aberrations is deemed to be beneficial in some situation, our method can just as well be used to increase the effect of these aberrations during post-processing.

Another experimental test was conducted by willfully introducing strong \( A_2 \) with the corrector of the microscope and acquiring

Fig. 3. Experimental verification of the correction method. (a) The AC-HRTEM image of Fig. 1a after application of the correction method with \( A_2 = 150 \text{ nm/}17^\circ \). The hexagonal appearance of the graphene lattice is restored and the structure of the defect is easier to analyze. (b) The AC-HRTEM image of Fig. 1b after the correction method was applied with \( A_2 = 75 \text{ nm/}15^\circ \). The MoSe\(_2\) lattice has identical appearance on both sides of the mirror-twin boundary as expected for an aberration-free image. The scale bars are 2 nm.

Fig. 2. Simulated verification of the correction procedure. (a) A simulated reference HRTEM image of MoSe\(_2\) flake with an embedded mirror domain with zero anti-symmetric aberrations. The same structure simulated with \( A_2 = B_2 = 100 \text{ nm/}30^\circ \) (b) and \( A_2 = B_2 = 10 \mu\text{m/}30^\circ \) (d). (c) and (e) The images of b and d after applying the correction, with the known parameters. The aberration-free appearance of the image was restored by the method in both cases. The scale bar is 1 nm.
images with such an ill-tuned system ($A_2 = 1.0 \pm 0.1 \, \mu m/48^\circ$, as measured by the corrector software). In Fig. 4a and b, two locations in the graphene sample are shown. In panel a it is clear that a tilt grain boundary runs through the center of the image, separating two areas with different lattice orientations. However, the structure of the boundary cannot be analyzed in the image, and the graphene lattice has a completely different appearance on the opposite sides of the boundary (due to different relative orientations of the lattices relative to the $A_2$ orientation). In panel b, a point defect in the graphene lattice can be detected, but again its structure remains hidden by $A_2$. The correction method was applied to both images using the $A_2$ value of 1.0701 $\mu m$ as measured by the corrector software. In both cases images free of $A_2$ are recovered and the imaged structure is easier to analyze. The scale bars are 2 nm.

4. Conclusion

To conclude, we have presented a method for correcting the anti-symmetric sub-group of aberrations in HRTEM images during numerical post-processing in the case of weakly scattering objects, and verified its applicability in the case of the often dominant low-order $A_2$ and $B_2$ aberrations through simulations and experiments with graphene and 2D MoSe$_2$. This procedure can be performed on a single conventional HRTEM image, that is, retrieval of the object wave is not necessary. The requirement of a weakly scattering object imposes a strong limitation on the range of materials where the method is applicable. However, the intensively studied class of 2D materials, with graphene as the prime example, do satisfy this condition at commonly used acceleration voltages such as 80 kV [21]. The contribution of the anti-symmetric aberrations is removed by applying the same aberrations with the opposite phase to the Fourier transform of the recorded intensity image and subsequently inverting the Fourier transform, that is, $I_0(x) = |F^{-1}(e^{i\chi(x)} \cdot F[I_0(x)])|$. We have presented the theoretical justification of the approach. The method was demonstrated on simulated and experimental HRTEM images suffering from residual anti-symmetric aberrations ($A_2$ and $B_2$). By applying the method, images with strongly reduced aberrations could be produced. In fact, the anti-symmetric aberration coefficients can be freely adjusted post-situ using the presented method.

A prerequisite for applying this method is prior knowledge of the aberration coefficients. In the presented examples, the values measured by the corrector software were used, or alternatively the coefficients were found by narrowing down the coefficients leading to the correct appearance of the graphene and MoSe$_2$ lattices through trial and error. The latter approach has an interesting implication: this method can actually be used to estimate the aberration coefficients when there is prior knowledge on how an aberration-free image should look like. Great care should be taken, however, as the problem can be underdetermined, and multiple aberration coefficient combinations can lead to the same final result.

The method allows some flexibility during the acquisition of images: as the anti-symmetric residual aberrations can be
corrected during post-processing, it is not imperative to have the corrector ‘perfectly’ tuned at all times. A reasonably good state of correction is important during acquisition, however, in order to enable the operator of the microscope to recognize the imaged features (compare, e.g., Fig. 4b and d). The loosened requirements for the tuning of the anti-symmetric aberration coefficients can accelerate the corrector tuning procedure. As the time available for the tuning is limited due to, e.g., gradual drift of the corrector state [18], an improved final state of correction can be achieved by the faster tuning procedure.

Acknowledgments

We are grateful to our group member, senior professor Harald H. Rose for invaluable advice during this work. We would also like to thank Hannu-Pekka Komsa, Arkady V. Krasheninnikov and Martin Link for helpful discussions, and Nilesh Vats for sample preparation. O.L., D.G., Z.L. and U.K. gratefully acknowledge the funding by the DFG (German Research Foundation) and the Ministry of Science, Research and the Arts (MKW) of Baden-Wuerttemberg in the framework of the SAIVE (Sub Angstrom Low-Voltage Electron Microscopy) project. O.L. acknowledges support from the Finnish Cultural Foundation. M.B.W., M.-W.C. and A.K. acknowledge funding from European Research Council (Grant no. 240076, FLATRONICS: Electronic devices based on nanolayers).

References