Micron 114 (2018) 23–31

Contents lists available at ScienceDirect

Micron

journal homepage: www.elsevier.com/locate/micron

Exit-wave phase retrieval from a single high-resolution transmission electron microscopy image of a weak-phase object

F. Lin[a,b], X.B. Ren[b], W.P. Zhou[a], L.Y. Zhang[a], Y. Xiao[a], Q. Zhang[a], H.T. Xu[a], H. Li[a,*], C.H. Jin[b,*]

[a] College of Electronic Engineering, South China Agricultural University, Guangzhou, Guangdong, 510642, China
[b] State Key Laboratory of Silicon Materials, School of Materials Science and Engineering, Zhejiang University, Hangzhou, 310027, China

ARTICLE INFO

Keywords:
Phase retrieval
Exit-wave function
HRTEM
Weak-phase object

ABSTRACT

We propose a novel algorithm to numerically retrieve the phase of the exit-wave function from a high-resolution transmission electron microscopy (HRTEM) image of a weak-phase object material, e.g., graphene and hexagonal boron nitride monolayers. It theoretically only requires a single HRTEM image to retrieve the phase under the assumption of a weak-phase object. In addition, it can remove the effects of geometrical aberrations up to fifth order, and also improve the degraded information due to the finite temporal and spatial coherence. We further present its applications and successfully demonstrate the identification of the lattice atoms and line defects in single HRTEM image of graphene.

1. Introduction

The quantitative determination of the atomic structures of nanomaterials is essential to understanding complex microstructure-property correlations. To this end, high-resolution transmission electron microscopes (HRTEMs) have long been regarded as one of the most powerful tools for precisely characterizing microstructures at the atomic scale. After a high-energy electron beam passes through a specimen, the exit wave function at the exit surface carries the projected structural information of the specimen. This wave function is further transferred by the optical system and collected by imaging devices, such as charge-coupled devices (CCDs), in which only the wave amplitude is recorded and the phase information is lost (Spence, 2003). Also, modulation by the optical system is unavoidably included in the final images, i.e., the influence of the spherical aberration (Cs) and astigmatisms. This will reduce the attainable spatial resolution, and therefore, it is essential to eliminate these modulations to unambiguously resolve the lattice structure and precisely determine the atom position.

With the invention of hardware spherical aberration correctors (Uhlemann and Haider, 1998), the point resolution of transmission electron microscopes (TEMs) has been improved remarkably in the last two decades (down to approximately 0.5 angstroms) (Erni et al., 2009); this has opened up vast new research possibilities and capabilities. However, the residual geometric aberrations, e.g., two-fold and three-fold astigmatisms, and axial coma, become increasingly important in HRTEM image. For instance, to attain a resolution of 1.0 Å in an HRTEM at 200 kV, the upper limits of two-fold, three-fold, four-fold and five-fold astigmatisms are calculated to be no larger than 1.0 nm, 60 nm, 3.2 μm and 0.16 mm, respectively (Uhlemann and Haider, 1998); and even if the astigmatisms are smaller than these values, the shape of atomic columns may be distorted and it hampers the quantitative measurement of the centers of atomic columns (Lin et al., 2015).

The geometric aberrations can be corrected in the post-imaging correction by using numerical methods, and a two-step procedure is usually used (Jia et al., 2005; Thust et al., 1996): the exit-wave function is firstly reconstructed from a focal series of HRTEM images; then, the residual aberrations are corrected directly from the exit-wave function. In the first step, the complex exit-wave function is reconstructed by iterative algorithms (Coene et al., 1996; Op de Beeck et al., 1996; Coene et al., 1992; Allen et al., 2004; Hsieh et al., 2004), e.g., the maximum-likelihood method (Coene et al., 1996), the parabola method (Op de Beeck et al., 1996), and the steepest descent method (Coene et al., 1992). In these methods, the HRTEM image is simulated from the reconstructed exit-wave function using image simulation theory and subsequently compared with the experimental image to obtain the adaptation for further refining of the exit-wave function. Note that during this process, the image simulation theory contains the contribution of both the linear and non-linear interference components. In the reconstructed exit-wave function, the influence of Cs and focus can be removed, and then, the high-frequency information can be improved to quantitatively interpret the lattice structure. However, as serial images with a rather wide focal range are needed and the image

* Corresponding authors.
E-mail addresses: flin_rew@163.com (F. Lin), lee.361@163.com (H. Li), chhjin@zju.edu.cn (C.H. Jin).

https://doi.org/10.1016/j.micron.2018.07.001
Received 1 May 2018; Received in revised form 3 July 2018; Accepted 3 July 2018
Available online 08 July 2018
0968-4328/

alignment should be performed precisely for any two images recorded under similar focal conditions, the exit-wave reconstruction typically requires 20–30 images, with 6 images being the minimum (Lin et al., 2010). Capturing such a large number of images is time consuming and sometimes is a challenge for beam-sensitive materials and *in situ* dynamics. As such, it is highly beneficial to be able to quantitatively interpret the structure with as few images as possible, e.g. only one HRTEM image.

Currently, three methods are used for post-image processing for correcting lens aberration of one HRTEM image. (i) Based on the weak-phase object approximation (WPOA), a direct method can calculate the projected potential of the specimen from the HRTEM image by using the image deconvolution (Han et al., 1986; Tang et al., 2006); the disadvantage of this method results from the fact that the information near the frequencies of crossing-zero points cannot be retrieved since the contrast transfer function (CTF) may cross the zero several times due to fluctuations, and noise may be amplified during the deconvolution operation. (ii) Recently, a numerical method has been proposed to efficiently remove the contribution of the anti-symmetric geometric aberrations for a weakly scattering object (Lehtinen et al., 2015). By multiplying the Fourier transform of the HRTEM image by the part of the transfer function contributed by the anti-symmetric aberrations (the wave aberration function χ is represented as a series of aberration terms and those aberrations satisfy the relation of $\chi_{as}(\mathrm{u}) = -\chi_{as}(-\mathrm{u})$ are the anti-symmetric aberrations), this method rapidly removes these aberrations and facilitates analysis of atomic defects in graphene. However, there remain some limitations, e.g., the contribution from the symmetric geometric aberrations cannot be eliminated properly, and the attenuation of high-frequency information due to the finite spatial and temporal coherence of electron beams cannot be improved. (iii) Another non-iterative method has been proposed by Morgan et al. to recover the complex wave at the exit surface of the specimen from a single image (Morgan et al., 2011). This method mainly attempts to recover the image distortions caused by defocusing and should satisfy certain conditions to ensure the resolvability of a set of linear equations.

The post-imaging correction of geometric aberrations for only one single image is very necessary for beam-sensitive materials, and specially and potentially, the atomic structure in a dynamic process can be achieved on the phases retrieved from a sequential series of images since the retrieved phase has a better resolution. Here we propose a new phase retrieval method for a weak-phase object using only a single HRTEM image to improve the information limitation facing structure interception. In this algorithm, we retrieve the phase of the wave function at the exit surface of the specimen by completely neglecting the amplitude of the wave. All the aberrations and effects produced by the temporal and spatial coherence of the electron beams can be effectively eliminated so that the image resolution can be greatly improved. The main content will be organized as follows. We will first introduce the simulation theory of HRTEM images and the exit-wave phase retrieval algorithm. Simulated and experimental images of monolayer crystals of a hexagonal boron nitride (*h*-BN) and graphene are used to test the algorithm. Finally, we present further details of this algorithm.

2. Phase retrieval theory

2.1. HRTEM image simulation theory

The exit-wave function $\Phi(u)$ is directly linked to the projected potential of the atoms during the propagation of the incident beam within the specimen, which is then modulated by the electron optical system and interferes with itself before being recorded by a CCD camera. The rigorous treatment of image simulation theory should include the contribution of both linear and non-linear interferences to the HRTEM image formation, as well as various geometric aberrations and effects of temporal and spatial coherences. For the sake of fast computation, the image $I(u, \Delta f)$ is generally given by (Coene et al., 1996)

$$I(u, \Delta f) = FT\left\{\sum_{m=-M}^{M} f_\Delta(m)|FT^{-1}[\Phi(u)T_m(u, \Delta f)]|^2\right\} \tag{1}$$

where u is a two-dimensional vector in the frequency domain, Δf is the focus value, and T_m is the transfer function of the microscope. Here, the Fourier transform (FT) and inverse Fourier transform (FT^{-1}) are used to simplify the correlation calculation involved in the coherent imaging. To incorporate the effects of the focal spread due to the partial temporal coherence of the electron beams, the image is assumed to be formed by superimposing a series of independent images at different defocuses along the focal axis. Given that m is an integer ranging from *-M* to *M,* the coefficient $f_\Delta(m)$ is written as

$$f_\Delta(m) = \frac{\delta\varepsilon}{\sqrt{2\pi}\,\Delta}\exp\left[-\frac{(m\delta\varepsilon)^2}{2\Delta^2}\right] \tag{2}$$

representing the weight of each image. The selection of the M value influences the temporal coherence envelope (Coene et al., 1996), and generally, we use $M = 3$ in our calculation, which is sufficient and suitable for the maximum spatial resolution of 8 nm^{-1}. And Δ is the half width of the focal spread, and $\delta\varepsilon$ is the step size of deviation from the imaging defocus Δf. The effective foci of each image are given by $\Delta f + m\delta\varepsilon$ ($m \in [-M, M]$). As such, the term $T_m(u, \Delta f)$ in Eq. (1) can be written as

$$T_m(u, \Delta f) = \exp[-2\pi i\chi(u, \Delta f + m\delta\varepsilon, A_1, A_2, B_2, \ldots)] \times \exp\{-\left(\frac{\pi\alpha}{\lambda}\right)^2[\nabla\chi(u, \Delta f + m\delta\varepsilon, A_1, A_2, B_2, \ldots)]^2\} \tag{3}$$

And λ is the wavelength, and χ is the wave aberration function, which takes various types of aberration into account, e.g., defocus, Cs, two-fold astigmatism (A_1), three-fold astigmatism (A_2), and axial coma (B_2). The term $\exp\{-\left(\frac{\pi\alpha}{\lambda}\right)^2[\nabla\chi(u, \Delta f + m\delta\varepsilon, A_1, A_2, B_2, \ldots)]^2\}$ accounts for the effects of partial spatial coherence, $\nabla\chi$ is the gradient of the wave aberration function taken with respect to u, and α is the semi-convergence angle of the incident beam. It should be emphasized that we also consider high-order aberrations, e.g., four-fold astigmatism (A_3), star aberration (S_3), five-fold astigmatism (A_4), three-lobe aberration (D_4), axial coma (B_4), five-fold spherical aberration (C_5), and six-fold astigmatism (A_5). The wave aberration function χ can be further written as (Uhlemann and Haider, 1998).

$$\begin{aligned}\chi(u, \Delta f, A_1, A_2, B_2, \ldots) &= Re\Big\{\frac{1}{2}\Delta f\lambda uu^* + \frac{1}{2}A_1\lambda u^{*2} + \frac{1}{3}A_2\lambda^2u^{*3} + B_2\lambda^2u^2u^* \\ &+ \frac{1}{4}\mathrm{Cs}\lambda^3(uu^*)^2 + \frac{1}{4}A_3\lambda^3u^{*4} + S_3\lambda^3u^3u^* + \frac{1}{5}A_4\lambda^4u^{*5} \\ &+ D_4\lambda^4u^4u^* + B_4\lambda^4u^3u^{*2} + \frac{1}{6}\mathrm{C}_5\lambda^5(uu^*)^3 + \frac{1}{6}A_5\lambda^5u^{*6}\Big\} \\ &= \frac{1}{2}\Delta f\lambda uu^* + \frac{1}{4}\mathrm{Cs}\lambda^3(uu^*)^2 + \chi'(u, A_1, A_2, B_2, \ldots)\end{aligned} \tag{4}$$

Generally, Eq. (4) can be classified into two parts: the primary aberrations caused by Cs and defocus and the residual aberrations . This classification stems from the two-step procedure of exit-wave reconstruction introduced in the previous section. Note that the contribution of Cs is relatively small in a Cs-corrected TEM, as such the residual aberrations influencing the CTF of the electron optical system become substantially more important.

If Eq. (1) is expressed in the spatial domain by using the Fourier convolution theorem, it is:

$$\begin{aligned}I(x, \Delta f) &= \sum_{m=-M}^{M} f_\Delta(m)\left|\int\phi(x')t_m(x-x', \Delta f)dx'\right|^2 \\ &= \sum_{m=-M}^{M} f_\Delta(m)|\phi_m(x, \Delta f)|^2\end{aligned} \tag{5}$$

in which $\phi(x)$ and $t_m(x, \Delta f)$ in the spatial domain x are the inverse Fourier transform of $\Phi(u)$ and $T_m(u, \Delta f)$, respectively, and $\phi_m(x, \Delta f)$ is generated by $\phi(x) \otimes\ t_m(x, \Delta f)$ ($\otimes$ is the convolution operator). Hence,