Measurement of charge density in nanoscale materials using off-axis electron holography

Zheng, Fengshan; Caron, Jan; Migunov, Vadim; Beleggia, Marco; Pozzi, Giulio; Dunin-Borkowski, Rafal E.

Published in: Journal of Electron Spectroscopy and Related Phenomena

Link to article, DOI: 10.1016/j.elspec.2019.07.002

Publication date: 2020

Document Version
Version created as part of publication process; publisher's layout; not normally made publicly available

Citation (APA):
Measurement of charge density in nanoscale materials using off-axis electron holography

Fengshan Zheng, Jan Caron, Vadim Migunov, Marco Beleggia, Giulio Pozzi, Rafal E. Dunin-Borkowski

1. Introduction

1.1. Charge density measurement

The development of a technique that can be used to measure charge density distributions in materials with high spatial resolution is important for understanding material properties such as conductivity, permittivity, ferroelectricity, piezoelectricity and spontaneous polarisation, as well as charge accumulation at interfaces in ferroelectric tunnel junctions [1] and p–n junctions [2], and charging and discharging processes in solid state battery devices [3].

Here, we illustrate recent progress in the development of an approach for local charge density measurement using off-axis electron holography through the study of an electrically biased needle-shaped specimen that was prepared for characterisation using atom probe tomography. Such a charge density measurement can be used to infer the spatial distribution of electric field around the specimen, which can be difficult to measure directly, in part because of its slow decay and its strong dependence on boundary conditions, including the shape and position of the counter-electrode. The electric field can then be used to determine the trajectories of ions that are emitted from the needle during atom probe tomography [4]. Such measurements can also be used to understand the relationship between the morphologies and electrical properties of field emitters.

1.2. Charge density measurement in the transmission electron microscope

The technique of off-axis electron holography involves the superposition of a highly coherent electron wave that has passed through an object of interest in a transmission electron microscope (TEM) with a reference electron wave using an electrostatic biprism, in order to form an interference pattern in the image plane, from which the phase of the object wave can be retrieved [5–7]. (A detailed explanation is provided in Section 2.1.) When examining a non-magnetic specimen, the phase is sensitive to the electrostatic potential projected in the incident electron beam direction. The technique has been used to study long-range electrostatic fields [8–12], such as those originating from triboelectric charges [13], trapped charges in specimens [14], charges at dislocations [15], electron-beam-induced charges in TEM specimens [16–19], p–n junctions [20,21], dynamic charging in Li ion battery materials [3], electrically biased tips [22–24], nanotips [25], and field emitters [26–30].

Related phase contrast techniques that are sensitive to electrostatic potential variations include in-line electron holography, as well as differential phase contrast (DPC) imaging [31] and ptychography in the scanning TEM (STEM). Iterative reconstruction algorithms have been developed for in-line holography [32] and comparisons between in-line and off-axis electron holography have been carried out [33,34]. DPC imaging is sensitive to the phase gradient rather than the phase shift,

ARTICLE INFO

Keywords: Charge density Off-axis electron holography Transmission electron microscopy Model-based iterative reconstruction Electric field Electrostatic potential

ABSTRACT

Three approaches for the measurement of charge density distributions in nanoscale materials from electron optical phase images recorded using off-axis electron holography are illustrated through the study of an electrically biased needle-shaped sample. We highlight the advantages of using a model-based iterative algorithm, which allows a priori information, such as the shape of the object and the influence of charges that are located outside the field of view, to be taken into account. The recovered charge density can be used to infer the electric field and electrostatic potential.
i.e., to the projected in-plane electric field rather than the projected potential [35]. Attempts have been made to use DPC imaging to measure polarisation fields [36], piezoelectric fields [37], built-in electric fields at p–n junctions [2], and atomic resolution signals [35,38,39]. The accuracy and precision of the technique are determined by the detector geometry and performance, experimental stability during scanning, calibration of the instrument, and careful interpretation [40].

The present paper is dedicated to the fundamental and practical aspects of quantitative charge density measurement at the nanoscale using off-axis electron holography. However, most of the conclusions are also relevant to experimental results obtained using other phase contrast techniques, including in-line electron holography, DPC imaging, and ptychography.

2. Basis of charge density measurement using off-axis electron holography

2.1. Theoretical considerations

Off-axis electron holography is a technique that allows retrieval of both the amplitude \( a \) and the phase \( \varphi \) of the wavefunction \( \psi = a \exp(i\varphi) \) that has passed through an electron-transparent specimen in the TEM. The experimental setup for the examination of an electrically-biased needle-shaped specimen (shown in red) is illustrated in Fig. 1. The specimen is illuminated by either a plane wave or a spherical wave that has a large radius of curvature \((I)\). Semi-transparent red surfaces show equipotential surfaces around the electrically-biased needle-shaped specimen. For clarity, the planar conducting counter-electrode that faces the tip (whose presence can be inferred by the flatness of the left equipotential surface) is not shown. An electron biprism \((EB)\) splits the electron wave into two parts: an object wave that passed through the specimen and a reference wave that passed through a region of vacuum outside it. Upon further propagation, the two parts of the electron wave overlap to form an interference pattern, or off-axis electron hologram, which encodes spatially-resolved phase and amplitude information \((Hol)\). The hologram is magnified by the projection lenses of the microscope (not shown) and recorded on a detector (typically a pixel array detector). It can then be processed digitally to retrieve real-space amplitude and phase information about the object.

In the absence of dynamical scattering and magnetic fields, the electron optical phase shift \( \varphi \) can be written in the form [5]

\[
\varphi(x, y) = C_E \int_{-\infty}^{\infty} V_I(x, y, z) dz. \tag{1}
\]

where \( s \) is the incident electron beam direction, \((x, y)\) are coordinates in the specimen plane, \( C_E \) is a constant that depends on the microscope accelerating voltage \((C_E = 6.53 \times 10^6 \text{ rad/(V m)} \text{ at } 300 \text{ kV})\) and \( V_I \) is the total electrostatic potential within and around the specimen, which includes contributions from the mean inner potential (MIP), fixed charges \((e.g., \text{ions})\), mobile charges \((e.g., \text{screening clouds})\), and polarisation charges.

If the electron biprism is oriented along the \( y \) axis, then the hologram intensity in an ideal imaging system can be expressed in the form \([8,9,41]\)

\[
I(x, y) = \left| \psi(x - \frac{d}{2}, y) \exp\left(\frac{imx}{s}\right) + \psi(x + \frac{d}{2}, y) \exp\left(-\frac{imx}{s}\right) \right|^2. \tag{2}
\]

where \( d \) is the interference distance (which depends on the biprism potential), and \( s \) is the spacing of the interference fringes in a vacuum reference hologram recorded with the specimen removed from the field of view \((\text{FOV})\). Eq. (2) describes two twin images of the object wavefunction.

If the object (specimen) is located at \( x + d/2 \), then the corresponding object wave \( \psi(x + \frac{d}{2}, y) \) is overlapped with the reference wave \( \psi(x - \frac{d}{2}, y) \). In order to retrieve the amplitude and phase of the object wave, the reference wave should ideally be equal to unity, or it should be known.

When long-range electrostatic fields originate from the specimen, as shown in Fig. 1, the reference wave may be perturbed. Analysis of the hologram then results in the reconstruction of a fictitious specimen, which can be described by the wavefunction \([8,9]\)

\[
\varphi(x, y) = \alpha(x, y) \exp\left[ -i \varphi(x + \frac{d}{2}, y) - i \varphi(x - \frac{d}{2}, y) \right], \tag{3}
\]

where \( \alpha(x, y) \) is the amplitude of the object wave, \( d \) is a two-dimensional “interference distance” vector that connects the two virtual sources created by the biprism, and \( \varphi(x + \frac{d}{2}, y) \) and \( \varphi(x - \frac{d}{2}, y) \) are the phases of the object and reference waves, respectively. The difference between these two phase distributions, rather than the true object phase, is then recovered. The influence of such a perturbed reference wave \((\text{PRW})\) on measurements of charge density and electric field is
The phase of the electron wave is typically reconstructed from an off-axis electron hologram by using a standard Fast Fourier Transform (FFT) based method [5]. First, the hologram is Fourier transformed. The Fourier transform contains two side-bands and a centre-band. The side-bands each contain complete information about the electron wave or its conjugate, while the centre-band is, to a first approximation, the Fourier transform of a conventional bright-field TEM image of the specimen. The electron wavefunction can be retrieved by selecting one of the side-bands using a digital (usually circular) mask, centring it and taking its inverse Fourier transform.

Once the object phase, which is proportional to the projected electrostatic potential, has been reconstructed, it can be analysed further to obtain the charge density distribution across the FOV. In classical electrodynamics, a potential obtained from the hologram has been reconstructed, it can be analysed further to take its inverse Fourier transform.

Furthermore, the Laplacian of the phase can be calculated directly from the reconstructed complex wavefunction using the expression [25]

$$\nabla^2 \psi = \frac{1}{\epsilon_0} \int_{V} \rho(x, y, z) dV,$$

where $$\epsilon_0$$ is the vacuum permittivity. According to Gauss' law, the flux of the electric field through a closed surface is proportional to the charge inside the volume of space bounded by that surface, according to the equation

$$\oint_{\partial V} \mathbf{E} \cdot d\mathbf{S} = \frac{1}{\epsilon_0} \iiint_{V} \rho(x, y, z) dV,$$

where $$\mathbf{E}$$ is the electric field and $$\partial V$$ is a surface that encloses volume $$V$$.

In the presence of a PRW, substitution of $$\psi(x, y)$$ in Eq. (3) by Eq. (1) results in an expression for the reconstructed phase $$\varphi_{\text{rec}}(x, y)$$ (i.e., the phase term of $$\psi$$ in Eq. (3)) of the form [22]

$$\varphi_{\text{rec}}(x, y) = C \int_{-\infty}^{\infty} \left[ V_0 \left( x + \frac{d}{2}, y, z \right) - V_0 \left( x - \frac{d}{2}, y, z \right) \right] dz.$$

By combining Eqs. (4)-(6), the relationship between the measured phase and the charge density distribution in the specimen can be expressed in the form [25,17]

$$\iint_{C} \nabla^2 \varphi(x, y)_{\text{rec}} dx dy = \frac{C}{\epsilon_0} \iint_{C} \rho_{\text{proj}}(x, y) dx dy = \frac{C}{\epsilon_0} Q_C,$$

where $$C$$ is the region of integration, $$\nabla^2$$ is a two-dimensional Laplacian operator, $$\rho_{\text{proj}}(x, y) = \int \rho(x - \vec{r}, y, z) dz$$ is the projected charge density, and $$Q_C$$ is the total charge present in region $$C$$. The volume over which Gauss’ law is evaluated, as discussed above, is an infinite cylinder (along the $$z$$ axis), of which $$C$$ is a cross-section.

Furthermore, the Laplacian of the phase can be calculated directly from the reconstructed complex wavefunction using the expression [25]

$$\nabla^2 \varphi = \text{Im} \left[ \frac{\nabla^2 \psi}{\psi} - \left( \frac{\nabla \psi}{\psi} \right)^2 \right] .$$

By making use of the divergence theorem, Eq. (7) can equivalently be written in the form [17]

$$Q_C = -\frac{\epsilon_0}{C} \int_{C} \nabla \cdot \nabla \varphi_{\text{rec}}(x(l), y(l)) \cdot \mathbf{n} (x(l), y(l)) dl,$$

where $$\nabla$$ is a two-dimensional gradient operator, $$\partial C$$ denotes an integration loop (coinciding with the boundary of the integration region $$C$$ in Eq. (7)), $$l$$ is a curvilinear coordinate along the contour and $$\mathbf{n}$$ is the outward normal to the contour.

2.2. Practical considerations

Parameters that can affect charge density measurements include the MIP contribution to the phase, the spatial resolution (i.e., the digital undersampling) of the recorded phase image, its signal-to-noise ratio (SNR), strong diffraction conditions (which can affect the measurement of the MIP contribution to the phase), electron-beam-induced specimen charging effects and the influence of sample imperfections (e.g., damage, contamination, and oxidation). Several of these considerations are now discussed.

2.2.1. Mean inner potential

The MIP of the specimen affects charge density measurements from electron optical phase images because it is associated with the presence of effective local dipole layers at the specimen surface [42]. Its influence is illustrated in Fig. 2. Fig. 2a shows part of a phase image of the end of an unbiased W needle recorded using off-axis electron holography. The needle is surrounded at its end by a layer of amorphous oxide and/or contamination. Fig. 2b shows the projected charge density distribution calculated directly from the Laplacian of the recorded phase image using Eq. (7). Evaluation of the Laplacian of the phase invariably results in a noisy image. In addition, Fig. 2b reveals that local variations in specimen thickness and MIP are visible in the form of.

Fig. 2. Apparent charge density distribution arising from the mean inner potential for an unbiased needle-shaped W specimen covered by an amorphous layer: (a) Part of an electron optical phase image recorded from the end of the needle using off-axis electron holography. (b) Charge density distribution calculated from the Laplacian of the phase, shown in units of e/px. An effective band of negative charge is situated on the vacuum side of the specimen edge, while an effective band of positive charge is situated on its inner side, thereby forming a dipole layer. Similar effective bands of charge are present at the interface between the needle and the surrounding oxide. (c) Cumulative charge profiles corresponding to integrals of the signal in the Laplacian of the phase across regions A (green) and B (red) marked in (b). The line profiles were in practice calculated from loop integrals (evaluated using Eq. (9)) applied to a median-filtered version of the phase. See text for details.
dipole layers, both at the specimen edge and at the boundary between the W core and the surrounding amorphous layer. Fig. 2c shows line profiles of the cumulative charge calculated using Eq. (9) by integrating the signal across Fig. 2b within the marked rectangles. The red curve from the amorphous layer alone in region B shows a peak and a dip at the specimen edge, whereas the green curve from region A shows additional similar features at the boundary between the amorphous layer and the W core. The MIP contribution to the phase therefore contributes additional effective negative or positive local charge wherever the sample thickness or MIP changes, as discussed elsewhere [43]. In each case, the total charge (illustrated here by the difference between the left and right sides of the line profiles in Fig. 2c) is zero. In the present example, the fact that the total charge is zero shows that there is no significant electron-beam-induced accumulation of charge in the W needle or the surrounding amorphous layer.

Several approaches can be used to remove the contribution to the measured charge density associated with the MIP. For the needle-shaped specimen examined here, we evaluated the difference between two aligned phase images that had been recorded with and without an electrical bias voltage applied between the needle and a counter-electrode, in order to measure the charge density distribution in the needle associated with the application of the electrical bias voltage alone. (See Section 3 below for details of the experimental setup). When using this approach, care is required to minimise any misalignment between the two phase images, which can result in artefacts in the phase difference image and the resulting charge density distribution, in particular at the specimen edge. Fig. 3 illustrates the influence of misalignment between two phase images on the measurement of the charge density distribution in the W needle examined above, in the form of cumulative charge profiles evaluated from phase difference images (Fig. 3a and b) that were calculated from deliberately misaligned phase images. The cumulative charge profiles shown in Fig. 3c were obtained in practice by applying loop integrals to median-filtered phase difference images that had been deliberately shifted with respect to each other by ±5 pixels, as shown in Fig. 3a and b. Each line profile shows four distinct peaks or dips at the specimen edge and at the interfaces between the W core and the surrounding amorphous layer. Just as for the influence of the MIP contribution to the phase, the total contribution to the cumulative charge across the FOV resulting from misalignment of the phase images is zero. However, local artefacts are present within the boundary of the specimen. Accurate alignment between two such phase images, often to sub-pixel precision, is therefore essential before evaluating the charge density distribution from their difference.

2.2.2. Spatial resolution

When using standard FFT-based reconstruction (see Section 2.1), the spatial resolution in the final phase image is determined by the size of the mask applied to the Fourier transform of the hologram, which can result in undersampling of the phase image, in addition to damping of high-frequency signal and noise. This effect is especially pronounced if coarser holographic interference fringes are used, resulting in a smaller separation between the side-band and centre-band, and therefore the need to use a smaller mask. For instance, the peak charge density at a p-n junction was damped when a small mask was chosen [21]. For this reason, the complementary techniques of in-line electron holography, DPC imaging and ptychography are sometimes better capable of retrieving high frequency phase information about a specimen [34,31].

2.2.3. Signal-to-noise ratio

Uncertainty in charge measurement is determined by factors that include noise in the original hologram, the sampling density of the phase image and the size of the integration region in the loop integral approach. If an experimental phase image (neglecting scattering absorption in the specimen) is considered to be a superposition of a noise-free ideal phase image and random normally-distributed noise with zero mean and standard deviation (SD) \( \delta \sigma \), then the Laplacian of the phase image can also be regarded as a noise-free charge distribution plus noise. Here, we show how the SD of the measured charge density distribution \( \delta p \) is related to the SD of the phase.

A discrete Laplacian is a one-step matrix algebra operator that maps each pixel in a phase image \( \varphi(i, j) \) onto the value \( \varphi(i+1, j) + \varphi(i-1, j) + \varphi(i, j+1) + \varphi(i, j-1) - 4 \varphi(i, j) \). If this operation is applied to a noisy phase image that has zero mean and SD \( \delta \varphi \), then the result is another noisy image, which is also normally distributed and has zero mean (because \( 1 + 1 + 1 + 1 = 0 \)), but which has a standard deviation that is \( \sqrt{20} \) times larger than the SD of the original image. This description is valid when each noise pixel is uncorrelated with its neighbours and when the SDs can be added in quadrature, such that \( 1^2 + 1^2 + 1^2 + 1^2 + 4^2 = 20 \). The relationship between the SDs of the phase and the charge density is given by the expression

\[
\delta \sigma = \sqrt{20} \frac{\delta \varphi}{C_e} = \sqrt{20} \frac{\delta p}{p^2}.
\]

Where \( p \) is the pixel size, and we define \( \delta q = \frac{\delta \varphi}{C_e} \) as the charge noise. For reference, \( \delta q = 118 \) mrad corresponds to \( \delta q = 1 \) e at 300 kV. Since experimental values of phase noise SD are typically well below 100 mrad, the achievement of single electron sensitivity in charge measurement appears to be relatively straightforward.

Eq. (10) is derived on the assumption of uncorrelated/white noise in the phase image. However, this situation does not strictly hold for Fourier-transform-based hologram reconstruction, as noise correlations are automatically introduced when a side-band is masked using an aperture. (A general description of the introduction of correlation in the reconstruction of holograms can be found elsewhere [44].) Such a mask may be “soft” (e.g., Gaussian, Hann, or Butterworth), or “hard” (e.g., top hat). Soft apertures are most commonly used because they are more efficient than hard apertures at suppressing phase noise without
introducing artefacts in the reconstructed phase. Fig. 4 shows a comparison between simulated white noise (Fig. 4a, top half), correlated noise resulting from the use of a Gaussian aperture (Fig. 4a, bottom half), and experimental noise (Fig. 4c) extracted from the vacuum region of a phase image. Correlations are visible in the granularity of the noise, which does not match the pixel size, as well as in the apparent width of the diagonal of the correlation matrix. White noise (Fig. 4b, top half) has a 1-pixel-wide diagonal, while correlated noise (Fig. 4b, bottom half) and Fig. 4d results in a thicker diagonal with a profile that decays in proportion to the smoothing parameter used to define the reconstruction mask.

Unfortunately, the calculation of a full correlation matrix would require several independent measurements, which are not available here. However, on the assumption of isotropic correlations, each column of a single image can be treated as one vacuum measurement, thereby reducing the problem from two dimensions to one dimension. The resulting (reduced) correlation matrix is shown in Fig. 4. Due to the reduced dimensionality, only coefficients $c_0, c_1$ and $c_2$ can be derived from this reduced matrix, where $c_0$ is the auto-correlation coefficient, $c_1$ is the correlation coefficient between adjacent pixels (e.g., $(i, j)$ and $(i+1, j)$) and $c_2$ is the correlation coefficient between non-adjacent pixels (e.g., $(i-1, j)$ and $(i+1, j)$). Based on the assumption of isotropic correlations, a Gaussian fit to these points can be extended to two dimensions, in turn allowing the determination of the $c_{ij}$ correlation coefficient between diagonally-neighbouring pixels (e.g., $(i, j)$ and $(i+1, j+1)$).

The noise correlations are not detrimental to the transfer of noise from the phase image to the charge distribution. On the contrary, since discrete differential operators involve taking differences between pixel values, covariances contribute to decreasing the transferred noise variance, as parts of the correlated noise cancel out. When a discrete Laplacian is applied to correlated noise, the noise transfer factor of 20 in Eq. (10) becomes $20 - 32c_1 + 8c_{11} + 4c_2$.

A representative value of phase noise in vacuum in a phase image reconstructed from a single off-axis electron hologram was found to be approximately 81 mrad (standard deviation). By using hologram series acquisition [45] and averaging 20 successive phase images, the phase noise was reduced to 17 mrad, in agreement with the expected $\sqrt{20}$ reduction for uncorrelated noise. The noise in each hologram is uncorrelated with others in the series and the averaging procedure does not introduce correlations. When a discrete Laplacian is applied to a selected region of vacuum in the phase image, for example to the region shown in Fig. 4c, the noise transfer factor is in general lower than $\sqrt{20}$, in agreement with the strong correlations that are visible in the experimental correlation matrix shown in Fig. 4d. The relevant correlation coefficients can be extracted from the experimental correlation matrix shown in Fig. 4d, resulting in values of $c_1 = 0.859, c_{11} = 0.746$ and $c_2 = 0.532$. ($c_{11}$ was estimated by applying second-order polynomial interpolation to $c_0 = 1, c_1$ and $c_2$). For these values, the noise transfer factor drops from $\sqrt{20}$ = 4.47 to $\sqrt{0.613}$ = 0.783. As a result, the SD of the charge density calculated using Eq. (10) drops from 0.64 e/p² to 0.11 e/p², in agreement with the experimental value of the charge density noise being 0.10 e/p².

Integration of the measured charge density distribution reduces noise, although it does not bring it back to the value that it had in the original phase image, both because the integration region is usually smaller than the FOV and as a result of noise correlations. (Even if the phase noise were uncorrelated, the charge density noise becomes correlated as a result of the use of the discrete Laplacian operator.) The right half of Fig. 4 illustrates how the SD of the measured charge $\delta Q$ is related to $\delta \sigma$ and $\delta \phi$. We consider a simple square 7 x 7 matrix with noisy pixels $\delta \phi$ representing the region of the phase image where we attempt charge measurement. The discrete Laplacian of this matrix, which is represented by a standard 3 x 3 kernel, is an edge-padded 5 x 5 matrix (the evaluation of boundary pixels is neglected for simplicity), in which each pixel value $L_{ij}$ is a linear combination of the original pixel values written above (e.g., $L_{44} = \delta \phi_{34} + \delta \phi_{35} + \delta \phi_{43} + \delta \phi_{45} - 4 \delta \phi_{44}$, etc.). Summing these 25 pixels gives another linear combination of pixel values. Isolating the coefficients of each pixel contributing to the sum and assembling them into a matrix yields what is shown in Fig. 4 at the end of the process diagram, which coincides with the discrete representation of the loop integral of the gradient (represented by the $(−1, 1)$ kernel) of the original image over the boundary of the chosen region. In addition to providing visual verification of the equivalence of the two methods for measuring $Q$, it implies that $\delta Q$ is also identical. Counting the number of pixels that contribute to the measurement of $Q$ and summing them in quadrature (for uncorrelated phase noise) provides the following relationship between the measurement uncertainty and the phase noise SD:

$$\delta Q = \sqrt{2n} \frac{\delta \phi}{C_{\phi}} = \sqrt{\frac{2n}{p}} \delta \phi,$$

(11)

where $n \gg 1$ is the number of pixels on one side of the square integration loop and $L$ is the total length of the loop. The noise transfer factor in the case of correlated phase noise becomes $\sqrt{8(1−c_1)n}$, where the only relevant correlations are those between adjacent pixels in the Laplacian. Since these correlations are also the strongest, $c_1$ can
be close to unity (it is 0.714 in the example considered above), supporting the performance of the charge measurement scheme with respect to noise, despite the use of discrete differential operators.

3. Reconstruction of charge density from electron optical phase images

Several approaches have been proposed for the reconstruction of charge density distributions from electron optical phase images and for removing the effect of the PRW, including the use of finite element simulations [30], a model-dependent approach [22,23,46] and a model-independent approach [17,25,43]. Here, we illustrate the application of a model-independent approach described by Eqs. (7) and (9) and an analytical model-dependent approach through the study of a needle-shaped LaB₆ specimen.

A LaB₆ specimen was prepared for use as a field emitter in a dual-beam focused ion beam (FIB) workstation. The specimen was milled to have an apex diameter of approximately 30 nm. An electrical bias voltage was applied to the field emitter in situ in the TEM using a specimen holder with a moveable tip that was also capable of scanning tunnelling microscopy (STM) [47], as shown in Fig. 5. A voltage could then be applied to the needle-shaped specimen in a geometry that is relevant for field emission and atom probe experiments. A voltage of between −200 and +200 V could be applied between the LaB₆ field emitter and a counter-electrode, which was positioned approximately 6 μm away from it. Off-axis electron holograms were recorded at 300 kV in an FEI Titan 60-300 TEM (Fig. 1) on a Gatan K2 camera using an exposure time of 8 s. The interference width was approximately 1.8 μm and the interference fringe spacing was approximately 2.7 nm, resulting in a nominal spatial resolution of approximately 8 nm.

The upper frame in Fig. 6a shows a phase image recorded from a LaB₆ field emitter without an electrical bias applied to it, corresponding only to the MIP contribution to the phase, i.e., to its thickness and shape. The lower right part of the specimen is covered by a region of amorphous contamination. The outline of the entire specimen, including the LaB₆ needle and the amorphous region, is marked using a dashed black line. Fig. 6b and c shows an equivalent phase image recorded in the presence of an applied voltage and the difference between the two phase images, respectively. In Fig. 6c, the mean inner potential contribution to the phase has been removed using the approach described in Section 2.2.1. Corresponding phase contour maps with a spacing of 2π radians are shown below each phase image. The phase contour map shown in Fig. 6a is not perfectly flat in the vacuum region around the specimen, suggesting that it has charged up slightly due to secondary electron emission in the TEM. This electron-beam-induced contribution to the charge density in the specimen is considered to be negligible here.

3.1. Model-independent reconstruction

Fig. 7a shows the charge distribution determined from the phase difference image shown in Fig. 6c using Eq. (7). The resulting charge distribution has poor signal-to-noise ratio, as expected from the use of the second-order derivative in the form of the Laplacian operator. Fig. 7b shows that, after applying a Gaussian filter to the charge distribution, negative charge can be seen to have accumulated at the edges and apex of the needle in the presence of the applied voltage of 50 V. Fig. 7c shows the cumulative charge profile in the specimen (blue) determined by integrating the charge density in Fig. 7a parallel to the emitter axis using the approach described in Section 2.1. The integration region is marked by a red dashed rectangle in Fig. 7a. The charge in the specimen is negative. The approximately constant slope of the cumulative charge profile suggests that the charge density in each slice of the three-dimensional volume of the specimen is the same, i.e., that it has an approximately constant linear charge density parallel to its axis.

3.2. Analytical model-dependent reconstruction

An analytical model-dependent approach for determining the charge density from an electron optical phase image relies on having access to a model that can be used to solve the Laplacian equation. The charged density in such a model can be adjusted until a best match is found between experimental and simulated images.

The value of $K$ in Eq. (12) was found from a best fit to the phase image. Fig. 8a shows a simulated contoured phase image corresponding to the best-fitting result, with $K = 35 e/\mu m^2$. The electric field strength $E$ induced by the counter-electrode and the base on which the field emitter sits can be calculated [23,46] and is approximately 0.4 MV/m, which is comparable to the electric field generated when 50 V is applied between two plates with a separation of 50 μm. Fig. 8b shows a stream plot of the electric field lines in the $z = 0$ plane for the best-fitting parameters. The maximum electric field strength at the apex $E_{\text{apex}}$ is approximately 3.6 GV/m. The ratio $E_{\text{apex}}/E$ corresponds to a field enhancement factor of 9000.

This field emitter was also transferred to a dedicated ultra-high-vacuum chamber to measure its field emission properties [48]. The electric field at the apex determined from a measured $I-V$ curve was 2.2 GV/m, which is slightly lower than that determined here. However, we did not detect a field emission current during our experiment, most likely due to the poorer vacuum level in the TEM column.

4. Numerical iterative model-based reconstruction

The accuracy of the analytical model-dependent approach described
Fig. 6. Upper row: Phase images recorded from a LaB$_6$ field emitter using off-axis electron holography, corresponding to: (a) the MIP contribution alone; (b) the MIP contribution and the effect of an electrical bias voltage of 50 V; (c) the difference between (b) and (c), i.e., the effect of the electrical bias voltage alone. The outline of the needle-shaped specimen is marked using a dashed black line. Lower row: Corresponding phase contour maps, displayed in the form of the cosine of the phase. The phase contour spacing is 2π radians.

Fig. 7. Application of the model-independent approach described in the text to measure the charge density distribution in a LaB$_6$ field emitter that has an electrical bias of 50 V applied to it in situ in the TEM. (a) Charge density distribution determined directly from the Laplacian of the phase; (b) Charge density distribution obtained by using a Gaussian filter with a standard deviation of 5 pixels (4.2 nm). (a) and (b) are both shown in units of e/pixel; (c) Cumulative charge profiles obtained using the model-independent approach (blue line) and model-based iterative reconstruction (red squares; see Section 4.4.6). The integration region is marked by a red dashed square in (a).

Fig. 8. Demonstration of the application of the analytical model-dependent approach described in the text to a LaB$_6$ field emitter that was electrically biased at 50 V. (a) Contoured phase image that provided a best fit in the vacuum region around the specimen to the experimental phase images shown in Fig. 6b and c. The phase contour spacing is 2π radians. (b) Streamlines showing a section through the electric field lines in the $z = 0$ plane calculated from the best-fitting parameters. The colour scale indicates the natural logarithm of the electric field strength. The shadow of the needle is indicated by the white region.
above relies on access to a suitable model for the charge density in the specimen, whereas the use of the Laplacian operator in the model-independent approach results in poor signal-to-noise ratio in the inferred charge density distribution. In order to tackle both of these limitations, a model-based iterative reconstruction (MBIR) algorithm was developed [49] to numerically retrieve the best-fitting projected or three-dimensional charge density distribution in a specimen from one or more electron optical phase (difference) images. This approach allows additional a priori information about the specimen to be included. The additional information can include a knowledge of the locations of untrustworthy parts of the image, such as regions that contain diffraction contrast, or a knowledge of the location and shape of the specimen, which can significantly reduce the number of unknowns when fitting the charge density distribution. This approach was first developed for the retrieval of magnetisation distributions from electron holographic phase images [49]. When applied to the reconstruction of three-dimensional magnetisation or charge density distributions, it avoids many of the artefacts that are present when using conventional backprojection-based tomographic reconstruction techniques.

4.1. Overview

In general, if a function \( F : \mathbb{R}^n \rightarrow \mathbb{R}^m \), which defines a forward model, maps a physical quantity onto a set of observable data, then the reconstruction of the quantity from the data is referred to as an inverse problem. In the present case, the dependence of one or more phase images \( \phi(x, y) \) on a charge density distribution \( Q(r) \) can be described by a forward model \( F(Q(r)) = \phi(x, y) \), in which the physical quantity \( Q \) is mapped onto a set of phase images \( \phi \). Such a forward model \( F \) is usually defined to operate on vectorised quantities according to the expression

\[
F(x) = y,
\]

where \( x \) represents the vectorised charge density distribution and \( y \) represents a concatenation of the pixels in the phase images.\(^1\) In the present case, the forward model is linear and can be expressed as a matrix-vector multiplication \( F(x) = \mathbf{Fx} \). The inverse problem involves retrieval of the charge distribution \( Q \) or \( x \) from the phase images \( \phi \) or \( y \). However, solving Eq. 13 for \( x \) is in general unfeasible due to the size of the matrix \( F \) and the fact that it is not of full-rank (indicating an “ill-posed” problem with no unique solution or no solution at all [50]).

Instead, a better-posed substitute problem can be defined in the form of a least-square minimisation of a cost function that guarantees the existence of a solution, in the form

\[
C(x) \equiv \| \mathbf{Fx} - y \|^2 + R_x,
\]

where the regularisation term \( R_x \) can be used to incorporate a priori knowledge about the sample, or other mathematical or physical constraints. In its simplest implementation, the regularisation term

\[
R_x = \lambda \| x \|^2
\]

is a simple Euclidean norm, or 0th order Tikhonov regularisation. The regularisation parameter \( \lambda \) then determines its weight in comparison to the first term of the cost function, which favours compliance with the measurements.

4.2. The forward model

The workflow of the reconstruction process is illustrated in Fig. 9. The forward model in this workflow, which is used to map a chosen charge density distribution onto a corresponding electrostatic phase image, serves as the basis for iterative reconstruction of the charge density. A simple and easy model that can be incorporated in the forward model is a dipole, which comprises a point charge and its image charge. The potential of a dipole, i.e., \( q(x_0, y_0, 0) \) and its image charge \( -q(x'_0, y'_0, 0) \), can be written in the form

\[
V_{\text{dipole}} = \frac{q}{4\pi\varepsilon_0} \left[ \frac{1}{\sqrt{(x-x_0)^2 + (y-y_0)^2}} + \frac{1}{\sqrt{(x-x_0)^2 + (y-y_0)^2 + z^2}} \right].
\]

Integration of the potential in the z direction from \( +\infty \) to \( -\infty \) then results in the expression

\[
\phi(x, y) = C_\phi \frac{q}{4\pi\varepsilon_0} \ln \left( \frac{(x-x_0)^2 + (y-y_0)^2}{(x-x_0)^2 + (y-y_0)^2 + z^2} \right).
\]

One should be aware that two singularities are present in the above equations. Here, we tackle this problem by treating a single voxel as a uniformly-charged sphere. (See Appendix.) Fig. 10 shows a simulated contour map and a corresponding phase image of a dipole, in which the charges are separated by a distance of 32 nm.

By making use of Eq. (17), the forward model can be divided into a projection of a charge density distribution in the electron beam direction and a subsequent phase mapping operation described by a convolution. In discretised form, the projection and convolution can be executed in two steps, which can be described in matrix form by the expression

\[
\mathbf{y} = \mathbf{Fx} = \mathbf{MPx},
\]

where the matrix \( \mathbf{F} \) is split into a projection matrix \( \mathbf{P} \) and a convolution matrix \( \mathbf{M} \), \( x \) is the charge state vector (i.e., \( q(x, y, z) \) in vectorised form) and \( y \) describes the calculated phase images in vectorised form. An efficient implementation of the projection step can be achieved by employing sparse matrix calculations, especially in the case of projections along \( z \) axis. In order to optimise the second step, the convolution kernels (see Eq. (17)) can be pre-calculated in real space and fast convolutions can be used in Fourier space [49,51].

4.3. Regularisator

Regularisation provides a way of making use of a priori information in the model-based inverse algorithm. The following regularisators were used here: the application of a mask to indicate the region that can contain charges (i.e., the location of the specimen); the application of a confidence mask to define trustworthy regions in the experimental phase; and the enforcement of physical or mathematical constraints by adding a Tikhonov regularisator [52] in the cost function.

The total electrostatic potential energy \( W \) of all of the charges can be written in the form [53]

\[
W = \frac{1}{8\pi\varepsilon_0} \sum_{i} \sum_{j} \frac{q_i q_j}{|\mathbf{r}_i - \mathbf{r}_j|},
\]

where \( q_i, q_j \) and \( \mathbf{r}_i, \mathbf{r}_j \) are the magnitudes and positions of the \( i, j \)th charges, respectively. The charges interact with each other through the Coulomb force, turning a linear term (Eq. (15)) into a non-linear one. For an ideal metal, in which charges are located only on the specimen surface, Eq. (19) can be reduced to the form

\[
w = \frac{s_a \sigma^2}{2},
\]

where \( w \) is the energy density and \( \sigma \) is the surface charge density. Minimisation of the total potential energy is a physical constraint that can be used to restrict the reconstruction of the charges. In order to enforce this constraint, we use Tikhonov regularisation of 0th order, which corresponds to the use of a scaled identity matrix in the regularisation term. The regularisation term is then exactly Eq. (15), i.e., a Euclidean norm. Although, in general, charges do not need to be

\[1\] Vectorised quantities of charge density \( x \) and phase \( y \) should not be confused with spatial coordinates \( (x, y) \).
located only on the specimen surface, 0th order Tikhonov regularisation is used here, as it aims to minimise the overall charge, which also has a physical meaning.

4.4. Reconstruction in projection

We first test charge density reconstruction in projection, as it is less complicated than three-dimensional (3D) reconstruction from a tilt series of phase images. It is also relevant for the study of two-dimensional materials. The parameters that can affect the fidelity of the reconstructed charge density distribution include:

- The mask that defines where charges can be located;
- Measurement noise in the phase image and the chosen regularisation strength;
- Other artefacts in the phase image;
- The presence of charges outside the FOV;
- The PRW.

In order to assess the influence of each of the above parameters, a simulated projected charge density distribution was generated from a uniform charge density distribution on a hollow sphere, as shown in Fig. 11. The inner and outer radius of the surface charge density distribution are ∼30 and 34 pixels, respectively, while the FOV is 128×128 pixels. The norm vector of the counter-electrode is (1, 1) mm, which is used in the forward model to define the positions of the image charges. The positions of both the charges and their images are symmetrical with respect to the plane of the counter-electrode. The charge distribution in the central slice of the original three-dimensional charge density distribution (at \( z = 0 \)) is shown in Fig. 11a, with a charge of \( 2 \times 10^{-3} \text{e/nm} \). The sampling density is 1 nm/pixel. The total charge in the shell is ∼100 electrons. Projection of the charge density distribution along the \( z \) axis was used to generate Fig. 11b. A line profile extracted across its centre (\( x = 0 \)), which contains features resulting from discretisation, is shown in Fig. 11c. Fig. 12 shows a corresponding calculated phase image.

4.4.1. Mask

A mask can be used to specify a priori information about the specimen geometry, i.e., the positions at which charges can be located. It can therefore lead to a significant reduction in the number of unknowns and to an improvement in the quality of the reconstruction. Although masking could be implemented as a term in the cost function, here we apply a mask by excluding these regions from the charge state vector \( x \), which corresponds to assuming a charge of zero in these regions. The algorithm then does not fit the regions outside the mask. Differences between input and reconstructed projected charge density distributions obtained from the phase image shown in Fig. 12, both with and without using a mask, are shown in Fig. 13. When a mask is not used, the error in the reconstructed charge is approximately 10%, while that in the phase is more than 150 μrad (not shown). There are also ripple-like artefacts in the reconstructed charge density. In contrast, when a mask (marked by the dashed circle in Fig. 12) is used, the error in the reconstructed charge falls to below 1%, while that in the phase falls to approximately 15 μrad (not shown). In addition, the ripple-like artefacts are absent.

4.4.2. Gaussian noise and regularisation strength

The noise in an experimental phase image can depend on the camera used and on the acquisition method (e.g., single vs multiple hologram acquisition). In the presence of noise, reconstruction without using a regularisator is found to result in a charge density distribution that can deviate greatly from the input. A 0th order Tikhonov regularisator was therefore used here. As discussed in Section 4.1, the regularisation strength, which is defined by the value of \( \lambda \), determines the ratio between the residual norm vector (the first term on the right of Eq. (14)) and the regularisation term. If \( \lambda \to 0 \), then the regularisation term vanishes and the cost function only relies on the residual norm vector, resulting in high frequency noise in the reconstructed charge density distribution. In contrast, if \( \lambda \to \infty \) then the cost function favours the regularisation term and the reconstructed result diverges from the experimental data. A good choice for the regularisation parameter corresponds to an optimal balance between compliance with the
measurements and enforcement of a priori constraints, and can be found from an L-curve plot [54], as illustrated in Fig. 14. Such a plot shows the normalised regularisation term $\|Fx - y\|_2^2$ as a function of the residual norm vector $\|Fx - y\|_2$, on a double logarithmic scale. (Details of the weighted matrix of $S^{-1}$ are given in Section 4.4.3). Fig. 14 was generated from the simulated phase image shown in Fig. 12, with a representative added Gaussian noise level of 0.05 rad. The almost-vertical part of the plot for smaller values of $\lambda$ corresponds to the reconstructed charge density being dominated by high frequency perturbations, whereas the formation of a quasi-plateau for increasing values of $\lambda$ corresponds to the high frequency perturbations being smoothed out. The optimal value of $\lambda$ is located at the corner of the plot, where the vertical line transitions into the plateau. This value represents a balance between compliance with the measurements and smoothness of the final charge density distribution. In the present example, the optimal value of $\lambda$ was determined to be approximately 5, resulting in the reconstructed charge density distribution shown in Fig. 15. At the edge of the mask, the reconstructed charge density deviates by 50% from the input distribution, while elsewhere the error is below 5%. The corresponding phase error is below 1%. Fig. 16 shows the influence on the reconstructed charge density distribution of using regularisation parameters of 0.5 and 50. For $\lambda = 50$, the charge distribution is too smooth, whereas for $\lambda = 0.5$ it is too noisy. It should be noted that the algorithm is designed to be insensitive to the presence of an arbitrary phase offset and an arbitrary phase ramp. Care in the interpretation of the result is therefore required if a real phase ramp may be present across the FOV.

4.4.3. Confidence mask

Experimental phase images can contain artefacts that originate from the specimen (e.g., unwanted effects of dynamical diffraction, contamination or electron-beam-induced charging), from the microscope (e.g., image distortions, or instabilities), from the detector (e.g., undersampling or dead pixels), or from image analysis (e.g., imperfect...
scaling or alignment of two phase images in magnification, position or angle before evaluating their difference). For example, it may not be possible to remove the MIP contribution to the phase everywhere across the field of view in the presence of changes in electron-beam-induced charging of the specimen. For all of these reasons, a confidence mask is used to define the degree to which the signal in each pixel in the phase image can be trusted. Regions that are trustworthy are assigned a value of unity, while other regions are assigned smaller values.

Here, we used a weighted norm for the cost calculation of the residuals, where the weighting matrix $S_1$ is diagonal, with each entry on the diagonal corresponding to a single pixel in each phase image. These values correspond to the entries of the confidence matrix. If the confidence value of a pixel is zero, then the corresponding residual does not contribute to the cost function. The weighted norm therefore takes the form

$$\frac{1}{2} \|F(x) - y\|_{S_1^{-1}}^2.$$ 

**Fig. 14.** L-Curve analysis of the reconstruction of the projected charge density distribution shown in Fig. 11. A good estimate for the optimal regularisation parameter $\lambda$ is 5. See text for details.

Density distribution is then determined only from the phase outside the charged sphere and can be seen to deviate significantly from the input charge density distribution. Interestingly, although the charge density distribution cannot be reconstructed reliably, the retrieved phase outside the charged sphere is consistent with the input phase, suggesting that the projected electric field can also be retrieved correctly outside the sphere. This is not surprising, since the algorithm always delivers a unique solution (in a mathematical sense) for a given type and strength of regularisation. However, without information about the phase in the interior of the object, reconstruction of the charge density inside it cannot be unique. For example, for a metallic ellipsoid the charge on the surface produces the same electric field distribution outside the object as a line of constant charge density located on its axis [22]. This statement is also consistent with the general property of a homogeneous Laplace equation that the values in the domain volume depend solely on the values or their derivatives on the domain boundary. The possibility of being able to reconstruct the potential and electric field outside a specimen without needing to make use of phase information inside it has significant implications for applications such as the characterisation of electrically biased needle-shaped specimens for atom probe tomography, for which the electric field outside the specimen rather than the charge density inside it may be the parameter of primary interest for providing experimental input for the simulation of ion trajectories. Nevertheless, whereas techniques based on differential phase contrast can be used to record the projected electric field directly, an argument in favour of reconstructing the charge density from an off-axis electron hologram before then using it to infer the projected electric field is that the charge density is localised within the specimen, rather than extending outside the FOV.

### 4.4.4. Charges outside the field of view

As a result of the limited lateral extent of a phase image, it is often not possible to include all of the specimen or all of the phase change associated with the charge density distribution inside the FOV. This problem is particularly apparent when examining electrically-biased needle-shaped specimens, such as atom probe tomography needles or field emitters. It is illustrated for a uniform shell-like charge density distribution in Fig. 18. The phase of the entire shell (not shown) is calculated using the charge density shown in Fig. 18a. However, only

**Fig. 13.** Difference between the reconstructed charge density distribution and the input charge density distribution shown in Fig. 11, first in the absence (top) and then in the presence (bottom) of a mask that defines the position of the specimen, shown in units of e/pixel. Corresponding line profiles across the centre ($x = 0$) are shown on the right. In each case, 1000 iterations of the reconstruction algorithm were used.
half of the calculated phase image was used for the reconstruction. The resulting reconstructed charge density distribution, which is shown in Fig. 18b, deviates significantly from the input charge density, as the algorithm assumes that charges are present only in the masked region, whereas charges outside the FOV also contribute to the phase. Since the forward model does not include any boundary conditions, with the exception of image charges, the presence of charges outside the FOV can be addressed by making use of additional buffer pixels, which are placed around the edge of the FOV. These buffer pixels can be used to introduce a distribution of additional charge density around the edge of the image, in order to take into account the influence of unknown charges outside the FOV. They are only used during the reconstruction and are discarded when displaying the final reconstructed charge density inside the part of the specimen that is within the FOV. In the present example, Fig. 18c shows that the reconstructed result is almost consistent with the original input charge density when 8 buffer pixels are used at the border of the image.

Fig. 15. Reconstruction starting from the phase image shown in Fig. 12 with Gaussian noise of 0.05 rad added, for a regularisation parameter λ of 5, showing the reconstructed projected charge density distribution and its deviation from the input charge density (upper row, shown in units of e/pixel), the reconstructed phase and its deviation from the input phase (middle row), and the charge profile across the centre of the image (x = 0) extracted from the reconstructed (red) and input (green) charge distributions and their difference (blue) (lower row).
4.4.5. Perturbed reference wave

As discussed in Section 2.1, the presence of a PRW can affect the interpretation of phase images if the vacuum reference wave that is used to form the off-axis electron hologram is perturbed significantly by the presence of long-range electrostatic (or magnetic) fields, which may originate from the object of interest itself. The influence of the PRW can be included in the reconstruction of the charge density by implementing Eqs. (1) and (6) in the forward model. A schematic illustration of the influence of the PRW on a recorded phase image is shown in Fig. 19 for a simple example of a single positive point charge within the FOV. In the absence of a PRW, i.e., without the tail of the electric field arising from the point charge affecting the vacuum reference wave, the measured phase distribution represents the projected potential of the positive charge faithfully and is symmetrical with respect to its position. However, the region from which the vacuum reference wave originates may be affected by the electric field of the point charge, which decays outside the FOV, as shown by the solid red line in Fig. 19. The measured phase is then the difference between the projected potential of the charge within the FOV and the projected potential in the vacuum reference wave region. The red dashed line in Fig. 19 represents the potential in the vacuum reference wave region, which has to be added to the potential within the FOV with a negative sign to take the PRW into account. The resulting phase image is asymmetrical with respect to the position of the charge. The red dashed line can be described as originating from a negative (virtual) point charge located on the other side of the biprism, as shown by a solid blue line in Fig. 19. The influence of the PRW can therefore be described by a region of virtual charges of opposite sign (mirror charges) that are located on the opposite side of the biprism at a distance that is equal to the interference distance. It can be treated in the same way as any other source of charge located outside the FOV, as described in Section 4.4.4, or alternatively by using a modified kernel that includes the PRW effect.

4.4.6. Reconstruction from an experimental phase image

A phase image recorded from a LaB₆ needle-shaped specimen that was electrically biased in situ in the TEM using an applied voltage of 50 V (Fig. 6c), from which the MIP contribution to the phase had been removed using the procedure describe above, was used for reconstruction of the charge density. The positions of image point charges in the forward model were calculated by assuming a distance between the needle and the contour-electrode of approximately 6 μm. The sampling density was 7 nm/pixel. A 4-pixel-wide buffer was also defined around the border of the image, in order to compensate for the presence of charges at unknown positions outside the FOV, as well as the presence of the PRW, as described in Sections 4.4.4 and 4.4.5. 0th order Tikhonov regularisation was used. An optimal value for the

Fig. 16. Comparison between reconstructed projected charge density distributions starting from the phase image shown in Fig. 12 for regularisation parameters λ of 50 and 0.5 (top row), shown in units of e/pixel. The lower panel shows the charge profiles across the centre for the input charge distribution (black) and the reconstructed charge distributions for three different regularisation parameters: 0.5 (green), 5 (red), and 50 (blue).
regularisation parameter $\lambda$ of 10 was determined from L-curve analysis, as shown in Fig. 20a. The reconstructed projected charge density distribution, which is shown in Fig. 20b, highlights the fact that charge accumulates around the outer edge of the electrically-biased needle-shaped specimen (including at the internal boundary between the needle and the amorphous contamination at its lower right side), but most strongly at its apex. The horizontal band of charge visible at the lower boundary of the image results from compensation for the effect of the charge outside the FOV, in particular in the wider base of the specimen, as described in the context of boundary pixels above. If the reconstructed charge is integrated along the axis of the specimen, then the resulting cumulative charge profile is found to increase linearly along the needle and to agree quantitatively with the result obtained using model-independent analysis above, as shown in Fig. 7. This agreement provides confidence in the result obtained using the MBIR algorithm. Although the reconstructed phase deviates slightly from the original experimental image (see Fig. 20c and d), this discrepancy is at a level of below 1% and may result from a slight error in the selection of the mask, or from the finite sampling of the phase at the narrow apex of the needle. If the reconstructed charge density distribution obtained using the MBIR algorithm (Fig. 20) is compared with that obtained from the Laplacian of the phase (Fig. 7), it is clear that the noise in the fitted charge distribution is greatly reduced, as a result of the use of a priori knowledge (in particular, the mask and the regularisation parameter) when performing the reconstruction. Nevertheless, it should be noted that the result depends strongly on the values of several input parameters, which should be chosen carefully when applying the MBIR approach.

4.5. Reconstruction of charge density in three dimensions

Three-dimensional charge density distributions can in principle be...
reconstructed using each of the approaches described above by applying standard backprojection-based tomographic reconstruction algorithms to projected charge density distributions measured as a function of specimen tilt angle. However, the MBIR approach offers significant advantages, as it allows the three-dimensional charge density to be reconstructed directly, while making use of all of the additional constraints and a priori information that were described above. It should be noted that a reconstructed three-dimensional charge density distribution will not only be affected by all of the parameters and sources of error that were discussed above, but also by errors in the alignment of images in the tilt series, incorrect tilt axis determination and incomplete datasets.

Here, we illustrate the influence of the choice of a three-dimensional mask on the reconstruction of charge density for a uniform charge density distribution on the surface of a sphere, as shown in Fig. 11. A corresponding phase image is shown in Fig. 12. As the charge distribution is symmetrical, phase images recorded in any direction are identical. Gaussian noise of magnitude 0.05 rad was added to each image in the tilt series. The tilt angle range was chosen to be ±50° about one axis, as this range can be achieved experimentally. The angular sampling was chosen to be 10°, resulting in an input dataset to the MBIR algorithm of 11 phase images in total. Three different three-dimensional masks were used: the shell, the outer surface of the sphere and the full three-dimensional volume. 0th order Tikhonov regularisation was used, as described in Section 4.3. The regularisation parameter $\lambda$ was set to be 100 for all three cases.

Figs. 21 and 22 show two-dimensional slices and line profiles through the reconstructed three-dimensional charge density. As expected, the use of a shell mask, which defines the true positions of the charges, delivers the best results. If only the outer surface of the sphere is chosen as a mask, then the algorithm retrieves the key feature of the charge distribution (the homogeneous surface charge) correctly. However, the reconstructed charge is smoothed slightly into the volume of the sphere and exhibits small oscillations next to the shell region. If the full three-dimensional volume is used, then the basic features of the charge density are reproduced (see Fig. 22), but additional spreading of the charge and high frequency artefacts are present across the FOV. Although further tests are required to optimise the use of the MBIR approach for three-dimensional charge density reconstruction, the results presented here are highly encouraging.

5. Summary and conclusions

Three different approaches have been described for the measurement of charge density distributions in nanoscale materials from electron optical phase images recorded using off-axis electron holography:
(i) an analytical model-dependent approach (Section 3.2), in which a mathematical model is used to describe the charge density and phase shift; (ii) a model-independent approach (Section 3.1), which is based on the application of a Laplacian operator to a recorded phase image; and (iii) a model-based iterative reconstruction approach (Section 4), in which the charge density in a forward model used to simulate phase images is varied until a best match to the experimental measurements is obtained. The analytical model-dependent approach (Section 3.2) relies on access to an analytical solution for the charge density and phase distribution for the experimental specimen geometry and requires the presence of a perturbed reference wave to be included in the model. The model-independent approach (Section 3.1) is the most direct and unbiased approach for retrieving the projected charge density distribution from a phase image and is insensitive to the presence of a perturbed reference wave and charges outside the field of view. However, the measured charge density can be noisy (since the approach relies on the evaluation of derivatives). In the MBIR approach (Section 4), the forward model approximates each charged voxel as a homogeneously charged sphere and a mirror charge. It can incorporate a priori knowledge through the use of masks, regularisation parameters and other physical constraints, resulting in lower noise but requiring care in the selection of parameters to avoid the introduction of artefacts. A further advantage is that boundary pixel regions can be used to take account of the presence of charges outside the field of view and the perturbed reference wave (Section 4.4.4). Artefacts can be tackled by assigning zero confidence to regions of phase images that contain untrustworthy information (Section 4.4.3). It is important to note that different charge distributions inside an object can result in the same electrostatic potential and phase distribution outside it.

The three approaches have been tested on an experimental phase image of an electrically biased needle-shaped LaB₆ specimen and have been shown to provide consistent results for the charge density. The phase shift of a line charge is used as a simple model in the analytical model-dependent approach. Projected charge density distributions retrieved using the model-independent approach (Fig. 7a and b) and the MBIR approach (Fig. 4.4.6) show that most of the charge is located close to the surface of the needle, with charge accumulation at its apex. The result obtained using the MBIR approach has much less noise than that obtained using the model-independent approach. Three-dimensional charge density distributions can in principle be reconstructed using each approach, either by applying a standard backprojection-based tomographic reconstruction algorithm to projected charge density data or using an iterative reconstruction approach. A further advantage is that boundary pixel regions can be used to take account of the presence of charges outside the field of view and the perturbed reference wave (Section 4.4.4). Artefacts can be tackled by assigning zero confidence to regions of phase images that contain untrustworthy information (Section 4.4.3). It is important to note that different charge distributions inside an object can result in the same electrostatic potential and phase distribution outside it.
density distributions measured as a function of specimen tilt angle or directly using the MBIR algorithm. We have used simulated phase images to show the advantages of using the MBIR approach, in which a priori information can be specified about the boundary of the object when reconstructing three-dimensional charge density distributions (Section 4.5, Figs. 21 and 22).

Acknowledgments

We are grateful to Prof. Michael Farle and AG Farle at the University of Duisburg-Essen for technical help and to Gopal Singh of the Max Planck Institute for the Structure and Dynamics of Matter in Hamburg, Germany and Dr. Urs Ramsperger of ETH Zurich, Switzerland for the provision of specimens and for valuable discussions. We acknowledge Werner Pieper and Rolf Speen for technical assistance and Maximilian Kruth for focused ion beam preparation of the LaB₆ specimen. This project was carried out within the framework of a scientific service agreement between the Ernst Ruska-Centre for Microscopy and Spectroscopy with Electrons in Forschungszentrum Jülich and Gatan Inc. The authors acknowledge the European Union for funding through the Marie Curie Initial Training Network SIMDALEE2 (Marie Curie Initial Training Network (ITN) Grant No. 606998 under FP7-PEOPLE-2013-ITN). V.M. thanks the Deutsche Forschungsgemeinschaft for funding within the framework of the SFB 917 project “Nanoswitches”. R.D.-B. thanks the Deutsche Forschungsgemeinschaft for a Deutsch-Israelische Projektkooperation (DIP) Grant and the European Union’s Horizon 2020 Research and Innovation Programme Q-SORT (Grant No. 766970 under H2020-FETOPEN-2016-2017). This project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 823717-ESTEEM3.

Appendix

The phase shift of a uniformly charged sphere of radius \( R \) and charge \( q \) can be written in the form

\[
\varphi(x, y) = 2C_0 \frac{q}{4\pi \varepsilon_0} \begin{cases} 
\ln \frac{d_1}{d_i} & \text{if } d_i, d_2 \geq R \\
\ln \frac{d_1}{d_i} + \frac{z_1}{R} + \frac{z_1^3}{3R^3} & \text{if } d_i < R, \ d_2 \geq R \\
\ln \frac{d_2}{d_i} - \frac{z_2}{R} - \frac{z_2^3}{3R^3} & \text{if } d_2 < R, \ d_i \geq R
\end{cases}
\]

where \( d_1 = \sqrt{(x-x_0)^2 + (y-y_0)^2} \) and \( d_2 = \sqrt{(x-x_0)^2 + (y-y_0)^2} \) are the projected distances to the charge and its image charge, respectively, \( z_1 = \sqrt{R^2 - d_1^2} \) and \( z_2 = \sqrt{R^2 - d_2^2} \) are the heights at which an electron enters the sphere and its image, respectively, and \((x_0, y_0)\) and \((x'_0, y'_0)\) are the coordinates of the centres of the sphere and its image, respectively.

References


