Electronic diffraction tomography by Green’s functions and singular values decompositions

A. Mayer*
Laboratoire de Physique du Solide, Facultés Universitaires Notre-Dame de la Paix, Rue de Bruxelles 61, B-5000 Namur, Belgium
(Received 5 July 2000; revised manuscript received 27 September 2000; published 2 January 2001)

An inverse scattering technique is developed to enable a three-dimensional sample reconstruction from the diffraction figures obtained for different sample orientations by electronic projection microscopy, thus performing a diffraction tomography. In its Green’s-functions formulation, this technique takes account of all orders of diffraction by performing an iterative reconstruction of the wave function on the observation screen and in the sample. In a final step, these quantities enable a reconstruction of the potential-energy distribution, which is assumed real valued. The method relies on the use of singular values decomposition techniques, thus providing the best least-squares solutions and enabling a reduction of noise. The technique is applied to the analysis of a three-dimensional nanometric sample that is observed in Fresnel conditions with an electron energy of 40 eV. The algorithm turns out to provide results with a mean relative error around 3% and to be stable against random noise.

DOI: 10.1103/PhysRevB.63.035408

I. INTRODUCTION

Electron projection microscopes1–3 make use of the quasiradial far propagation of electrons projected out of small tips to obtain highly magnified projections of small samples. The tip and sample dimensions are of the order of a few nanometers (the tip having a single-atom termination to ensure the pointlike character of the emission area). The bias responsible for the field-emission process is typically between a few tens volts and a few hundreds volts.

Below a critical tip-sample distance, the incoming wave is essentially spherical on the scale of the sample, giving rise to Fresnel diffraction images, still highly correlated with direct-space representations of the object. The best resolution obtained so far by this direct approach is around 0.5 nm. An inverse scattering treatment should improve this resolution, by using the information present in all diffraction fringes and by considering simultaneously the projections obtained for different positions of the source.

Theoretical support enabling the understanding of the image formation (i.e., direct scattering simulations) was given within the Fresnel-Kirchhoff flat-object formalism.4 Consideration of three-dimensional aspects is possible in a Green’s-functions formulation.5–8 The main disadvantage of this approach is that it requires excessive storage space when the sample is large. Our approach to the problem was based on both the transfer-matrix and Green’s-functions formalisms.9–14 This approach enables direct scattering simulations15–17 with reduced storage requirements.

Inverse scattering has been developed among others18–20 for x rays,21,22 ultrasonic,23,24 optical,25 and electron transmission26–28 techniques. The problem of inverse scattering in a projection configuration was already considered by Bleloch29 and Spence.30 Their method relies on the Fresnel-Kirchhoff formalism, so only a two-dimensional reconstruction of the sample (i.e., a projection) is achieved, without consideration of multiple scattering effects. In a previous paper,31 we introduced a technique based on the Green’s-functions formalism that is suited for a three-dimensional reconstruction of the sample with consideration of all multiple scattering effects.

This paper is an extension of that technique to take into account the diffraction figures obtained for different tip positions or sample orientations, thus performing a diffraction tomography. We take the point of view that it is the sample that is rotated while the source and screen are fixed. The idea is to relate the scattered intensities (obtained for the various sample orientations) to the sample potential energy in the Green’s-functions formalism and to solve the corresponding linear system of equations by a singular values decomposition32 (thus obtaining the best, least-squares solution). The screen and sample wave functions are reconstructed iteratively to match the given intensities. After completion of this step, a reconstruction of the sample potential-energy distribution is achieved. All information obtained for the different sample orientations is always considered simultaneously, in all steps of the algorithm. This paper starts with the description of this technique. In Sec. III, it is applied to the reconstruction of a three-dimensional nanometric sample, from clean Fresnel diffraction figures and from noisy data. The algorithm turns out to provide results with a mean relative error around 3% and to be stable against random noise.

II. DIFFRACTION TOMOGRAPHY BY GREEN’S FUNCTIONS AND SINGULAR VALUES DECOMPOSITIONS

The Green’s-functions description of the scattering process with the associated reconstruction technique were given with details in Ref. 31. Only the main lines of the reconstruction technique, with the adaptations required for the consideration of nview sample orientations, are given in this section.

A. Notations and fundamental equation

Let us refer to \( \Psi_{\text{scrt}},s \) and \( \Psi_{\text{scrt}},s \) as two vectors containing the values of the total and incident wave function on the \( n_{\text{scrt}} \) points of a distant screen, for a sample orientation with label \( s \). The two vectors \( \Psi_{\text{obj}},s \) and \( \Psi_{\text{obj}},s \) contain the corresponding values on the \( n_{\text{obj}} \) discretization cells of the object. The values of the incident wave function on the screen...
are not dependent on the sample orientation. However, for the homogeneity of the notation, we will keep the label \( s \) in \( \Psi_{\text{scr}}^0 [s] \).

Let \( V \) be a \( n_{\text{obj}} \times n_{\text{obj}} \) diagonal matrix containing the unknown values of the potential energy in the \( n_{\text{obj}} \) discretization cells of the sample, each one being associated with a volume \( \Delta V_i \). The values of the Green’s functions \( G \) are contained in the \( n_{\text{obj}} \times n_{\text{obj}} \) matrix \( G_{\text{obj, obj}} \) whose elements are the quantities \( G(r_i, r_j, E) \Delta V_i \) (the \( r_i \) and \( r_j \) referring both to points in the object) and in the \( n_{\text{obj}} \times n_{\text{scr}} \) matrix \( G_{\text{scr, obj}} \) whose elements are the quantities \( G(r_i, r_{i[s]}, E) \Delta V_i \) (the \( r_i \) referring to points on the screen and the \( r_{i[s]} \) to points in the object for the orientation \( s \), these points being attached to the sample).

Let us finally define \( \Psi_{\text{scr}}^{k-1} = \sum_{k=0}^{k} \Psi_{\text{obj}}^{k} \) and \( \Psi_{\text{obj}}^{k} = \sum_{k=0}^{k} \Psi_{\text{obj}}^{k} \), where \( \Psi_{\text{obj}}^{k} \) is the \( k \)th Born expansion term of the wave function in the object (in its orientation \( s \)), i.e.,

\[
\Psi_{\text{obj}}^{k} = \sum_{k=0}^{\infty} \Psi_{\text{obj}}^{k} = G_{\text{obj, obj}} V^{k} \Psi_{\text{obj}}^{0} \tag{1}
\]

and \( \Psi_{\text{scr}}^{k} \) the \( k \)th Born expansion term of the wave function on the screen, i.e.,

\[
\Psi_{\text{scr}}^{k} = G_{\text{scr, obj}} V^{k} \Psi_{\text{scr}, \text{obj}}^{0} \tag{2}
\]

With these notations, the Green’s functions description of the scattering process leads then to the following fundamental relation:

\[
\Psi_{\text{scr}}^{k} = \Psi_{\text{scr}}^{k-\infty} = \Psi_{\text{scr}}^{k-0} G_{\text{scr, obj}} \Psi_{\text{obj}}^{k-\infty} \tag{3}
\]

B. Iterative reconstruction of the scattered wave function

In the first step of the iteration (\( k = 0 \)), we have the equation \( \Psi_{\text{scr}} = \Psi_{\text{scr}}^{0} + G_{\text{scr, obj}} \Psi_{\text{obj}}^{0} \), where \( \Psi_{\text{obj}}^{0} \) is approximated by its major contribution, i.e., \( \Psi_{\text{obj}}^{0} \approx \Psi_{\text{obj}}^{0} \), which is associated with \( V = 0 \).

In this first and subsequent steps of the algorithm, we have an approximation of \( \Psi_{\text{scr}} \), \( \Psi_{\text{obj}} \), and \( \Psi_{\text{obj}} \) that is associated with an approximation of \( \Psi_{\text{scr}} \) by \( \Psi_{\text{scr}}^{k+1} = G_{\text{scr, obj}} \Psi_{\text{obj}}^{k+1} \).

The intensities on the \( n_{\text{scr}} \) points of the screen (for each sample orientation \( s \)) are given by

\[
\begin{align*}
|\Psi_{\text{scr}}^{k}(r_i)|^2 &= \left| \Psi_{\text{scr}}^{k+1}(r_i) \right|^2 + \text{Re} \left\{ \Psi_{\text{obj}}^{k+1}(r_{i[s]}) \right\} E \sum_{l=1}^{n_{\text{obj}}} 2 \text{Re} \left\{ \Psi_{\text{obj}}^{k+1}(r_{i[s]}) \right\} \Psi_{\text{scr}}^{k+1}(r_{i[s]}) V(r_{i[s]}),
\end{align*}
\]

provided the potential energy is real valued.

This system of \( n_{\text{view}} \times n_{\text{scr}} \) equations is linear in \( V(r_i) \) (whose values do not change with the sample orientation since the \( r_{i[s]} \) are attached to the sample) and can be solved by a singular values decomposition (see Appendix A of Ref. 31), thus providing the \( V(r_i) \) that give the best agreement (in the least-squares sense) with the known intensity values.

The potential-energy matrix \( V \) is then used to provide the estimation of \( \Psi_{\text{scr}}^{k+1} \left[ s \right] \), \( \Psi_{\text{obj}}^{k+1} \), \( \Psi_{\text{obj}}^{k+1} \), and \( \Psi_{\text{obj}}^{k+1} \) for the next iteration and for the final determination of the potential energy by:

\[
\begin{align*}
\Psi_{\text{scr}}^{k+1} &= \Psi_{\text{scr}}^{k+1} + G_{\text{scr, obj}} \Psi_{\text{obj}}^{k+1}, \tag{5}
\Psi_{\text{obj}}^{k+1} &= \Psi_{\text{obj}}^{k+1} + G_{\text{obj, obj}} \Psi_{\text{obj}}^{k+1}, \tag{6}
\Psi_{\text{obj}}^{k+1} &= \Psi_{\text{obj}}^{k+1} + G_{\text{obj, obj}} \Psi_{\text{obj}}^{k+1}, \tag{7}
\end{align*}
\]

C. Final determination of the potential energy

Once the wave function on the screen and in the sample have been reconstructed, the potential energy is determined by solving for \( V(r_i) \) the system of equations:

\[
|\Psi_{\text{scr}}^{k}(r_i)|^2 = |\Psi_{\text{scr}}^{k}(r_i) + \Psi_{\text{obj}}^{k+1}(r_{i[s]})|^2 = |\Psi_{\text{scr}}^{k}(r_i)|^2 + |\Psi_{\text{scr}}^{k+1}(r_i)|^2 + \sum_{l=1}^{n_{\text{obj}}} 2 \text{Re} \left\{ \Psi_{\text{scr}}^{k+1}(r_i) \right\} \Psi_{\text{obj}}^{k+1}(r_{i[s]}) \}
\]

where the quantities \( \Psi_{\text{scr}}^{k+1}(r_i) \), \( \Psi_{\text{obj}}^{k+1}(r_{i[s]}) \), and \( \Psi_{\text{obj}}^{k+1}(r_{i[s]}) \) result from the preceding section, with \( k \) corresponding to the last iteration.

To improve the efficiency of the algorithm, the long-range intensities should be renormalized to the same asymptotic \( l/r \) factor (so this geometrical factor is removed from the information the singular values decompositions have to deal with) and only the singular values that are higher than 1/20 of their maximum should be considered (or even a more restricted set of singular values in the case of strong noise) to prevent high-frequency oscillations to appear in the reconstruction. Finally, the algorithm can be run iteratively by evaluating the starting quantity \( \Psi_{\text{obj}}^{0} \)

\[
\Psi_{\text{obj}}^{0} = \sum_{k=0}^{\infty} \left( G_{\text{obj, obj}} \Psi_{\text{obj}}^{0} \right) \Psi_{\text{obj}}^{0} \]
perpendicular to the screen. The wavelength $\lambda$ is 0.196 nm in the vacuum and 0.185 nm in the sample, so some of its parts have dimensions $\approx 0.1$ nm lower than the electron wavelength and lower than the resolution limit $\Delta d = 1/2 \sqrt{\lambda d}$.

The diffraction figure is measured on a flat 10 cm distant screen (on a regular cartesian grid) but the intensity values are renormalized to those that would have been measured on a sphere with a 10 cm radius.

**B. Reconstruction from a clean diffraction figure**

The first reconstruction is achieved by considering a source-sample distance of 3 nm, thus giving rise to Fresnel diffraction. The screen is 30 cm wide and has a 35 $\times$ 35 resolution. Five sample orientations are considered, i.e., rotations of $-72, -36, 0, 36,$ and $72^\circ$ around the $y$ axis. The reconstruction is achieved in a 0.9 $\times$ 0.9 $\times$ 0.5 nm$^3$ sample support and the potential-energy values are restricted to the $[-5.0]$ eV range. This is justified when the physical properties of the sample (i.e., its internal potential energy) are known. The result of only two iterations of the algorithm given in the previous section is presented in Fig. 2. The three parts of this figure are $xz$ sections of the reconstruction corresponding, respectively, to $y = -0.2, -0.1,$ and 0 nm. The different parts of the sample are perfectly recovered.

The mean relative error is evaluated by $E = (1/n_{obj}) \sum_{i=1}^{n_{obj}} |V_{i,i}^{\text{true}} - V_{i,i}^{\text{recons}}|/5 \text{ eV}$, where $n_{obj}$ is the number of discretization cells in and around the sample and where the $V_{i,i}^{\text{true}}$ and $V_{i,i}^{\text{recons}}$ refer to the corresponding true and reconstructed potential-energy values.

The mean relative error of the reconstruction presented in Fig. 2 is 2.7%. A 3.9% relative error is achieved by considering only three sample orientations (again with two iterations of the algorithm). The maximal deviation from the exact result is then increased from 22 to 38% of the 5 eV reference value. It turns out that the quality of the reconstruction increases with the number of sample orientations. The range of the sample rotations should be as large as possible, although results obtained by a rotation range of 180° are not significantly better than those corresponding to a 90° range.

A relative error around $10^{-2}$ is achieved when the true wave function (on the screen and in the sample) is used in Sec. II C, in the case of a single sample orientation and no singular values restriction. Considering five sample orientations results in a $10^{-3}$ relative error! The technique of Sec. II E of Ref. 31 only requires the knowledge of the wave function on the screen. Considering the exact values and a single sample orientation provides the same $10^{-3}$ relative error. The technique of the present paper gives, however, better results when processing with diffraction intensities only, since all singular values decompositions can be performed while those required in Sec. II E of Ref. 31 fail with three-dimensional samples, thus limiting the application of the algorithm to a single iteration.

The singular values decomposition required in Sec. II B is essential for the reconstruction. This decomposition fails when the dimensions of the screen are not large enough. In the case of Fresnel diffraction, the screen has to be at least 2.5 times larger than the geometrical projection of the sample support from the point source. The simulations presented so far considered a screen around 10 times as large as this geometrical projection. In the case of Fraunhofer diffraction, the screen should be large enough to include the first-order fringes.

**C. Reconstruction from a diffraction figure with random noise**

To test the stability of the algorithm, a random noise was added to the diffraction intensities. This random noise consists of a relative and absolute part. The relative part of the
noise is characterized by its amplitude in each screen pixel being proportional to the corresponding true intensity value. The absolute part of the noise is characterized by its amplitude (everywhere) being proportional to the true maximal intensity encountered on the screen.

In this last simulation, a combination of 5% relative and 3% absolute noises was considered, these values being typical of a ‘‘correct’’ experiment. In general, the effects of noise can be reduced by keeping only the singular values $\sigma_i$ whose ratio to their maximal value is higher than the noise level. However, the noise level being in this case lower or equal to the 1/20 reference value, no improvement results from a further restriction of the singular values. The reconstruction was done by running three iterations of the algorithm. The result (presented as three $xz$ sections corresponding to $y = -0.2, -0.1,$ and 0 nm) is illustrated in Fig. 3. The mean relative error on the sample reconstruction is 6.1%. This is reasonable compared with the 2.7% value obtained with clean diffraction figures. The noise is, however, responsible for the maximal deviation on the whole reconstruction to be increased from 22 to 47 % of the 5 eV reference value.

The stability of our algorithm was tested with more details in Ref. 31. Although the technique of this paper is somewhat different, the conclusions are the same. Our algorithm is stable against random noise, whose effects are controlled mainly by an appropriate selection of the singular values. Enforcing limits in the potential energy (when knowledge of the physical properties of the sample allows so) and increasing the resolution and dimensions of the screen also lead to improvements in the results.

IV. CONCLUSION

In this paper, the inverse scattering technique presented in our previous paper 31 was developed to achieve an electronic wave tomography. In its formulation, this technique takes into account all orders of diffraction by performing an iterative reconstruction of the wave function. In the case of thin samples, where the scattering process is essentially reduced to its first-order contribution, the results are not significantly improved by running more than a few iterations of the algorithm.

In Fresnel conditions, the algorithm is very stable against random noise, whose effects are reduced by neglecting the contributions associated with too small singular values. To help the algorithm separating useful information from noise, it is also useful to increase the resolution and size of the screen (so more diffraction fringes are taken into account) and to increase the electron energy (thus reducing both the wavelength and resolution limit and also reinforcing the predominance of first-order diffraction).

The limits of this approach come essentially from the space needed to store and treat the two arrays $G_{\text{obj, obj}}$ and $G_{\text{scr, obj}}$ whose size is, respectively, proportional to $n_{\text{obj}}^2$ and $n_{\text{view}} \cdot n_{\text{scr}}$, $n_{\text{obj}}$, $n_{\text{view}}$ being the number of views, $n_{\text{obj}}$ is the number of discretization cells in the sample, and $n_{\text{scr}}$ is the number of pixels on the screen. For the same amount of storage space, it turns out that increasing the number of views is more efficient than increasing the resolution of the screen. Taking simultaneously into account projections obtained for different energy values $E$ is possible, by letting $E$ appear as a new parameter beside the sample orientation $s$. However, since there would be as many $G_{\text{obj, obj}}$ arrays as energy values $E$, it is probably more efficient to increase the number of views. Extending significantly the range of the technique will require further developments or the use of out-of-core techniques.

An exact knowledge of the incident wave is assumed. The amplitude and the angular dependence of the incident wave can be determined by measuring the electron beam obtained without any sample. The source-sample distance $d$ can then be determined approximately from the changes induced in the projection by a lateral source displacement. In a final step, this distance $d$ can be refined by selecting the value providing a reconstruction with a minimum of oscillations. In its present form, this technique is limited to real-valued potential energies (see Secs. II B and II C) and will therefore apply essentially to small transparent samples (for which the most interpretation is needed).

ACKNOWLEDGMENTS

A.M. was supported by the Belgian National Fund for Scientific Research (FNRS). The author acknowledges the national program on the Interuniversity Research Project (PAI) and the use of the Namur Scientific Computing Facility, a common project between the FNRS, IBM-Belgium and the FUNDP.
5 E.N. Economou, Green's Functions in Quantum Physics (Springer-Verlag, Berlin, 1983).