Inverse electronic scattering by Green's functions and singular values decomposition

A. Mayer* and J.-P. Vigneron
Laboratoire de Physique du Solide, Facultés Universitaires Notre-Dame de la Paix, Rue de Bruxelles 61, B-5000 Namur, Belgium
(Received 22 May 2000)

An inverse scattering technique is developed to enable a sample reconstruction from the diffraction figures obtained by electronic projection microscopy. 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. This scattered wave function is then backpropagated to the sample to determine 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 two-dimensional nanometric sample that is observed in Fresnel conditions with an electronic energy of 25 eV. The algorithm turns out to provide results with a mean relative error of the order of 5% and to be very stable against random noise.

I. INTRODUCTION

Electron projection microscopes[1–3] 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 hundred 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 or different energy values.

Theoretical support enabling the understanding of the image formation (i.e., direct-scattering simulations) was first 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 becomes large. Our approach of the problem was based on both the transfer-matrix and Green's functions formalisms.[9–14] This approach enables direct-scattering simulations[15–17] with reduced storage requirements. However, due to the fact that the potential energy remains explicitly present in the Green's functions formulation, this latter will be used in this inverse scattering approach.

Inverse scattering[18–28] has been developed among others for x rays, ultrasonic, and electron transmission techniques. Diffraction tomography refers to the treatment of a set of diffraction figures obtained for different positions of the source. Among the works that inspired this paper, we refer to Maleki and Devaney[18,19] and Wedberg and Stamnes[20] who use an iterative scheme to reconstruct the unknown phase of the scattered (acoustic/optical) field. The backpropagation of the scattered field from the observation plane to the sample is achieved by the (hybrid) filtered backpropagation algorithm[19] and the sample is assumed to be a weak scatterer. In the formulation of x-ray inverse scattering by Tegze and Faigel[21,22] (where weak scattering is also assumed), the scattered intensity is related linearly to the atomic scattering factors (more precisely to the real part of a combination of these factors). In their Fourier analysis of electron transmission micrographs, Klug *et al.*[23–25] make interpolations in the reciprocal space (by generalized inversions[29]) to complete the data obtained from a limited set of projections.

The idea of this paper is to relate the scattered intensity to the sample potential energy in the Green's functions formalism and to solve the corresponding linear system of equations by a singular values decomposition[30] (providing the best least-squares solution). The scattered wave function is reconstructed iteratively to match the given intensity. This paper starts with the derivation of this technique. In Sec. III, it is applied to the reconstruction of a nanometric sample, from a clean diffraction figure first and then from a figure with random noise. Only Fresnel diffraction is considered and a comparison between our results and those obtained by the techniques of Bleloch *et al.*[31] and Spence *et al.*[32] is achieved. In the conditions of this paper, it turns out that a sample reconstruction from a Fresnel figure is more stable against noise, whose effects can be reduced by an appropriate singular values selection and by increasing the electron energy or the resolution and size of the diffraction figure.

II. INVERSE SCATTERING BY GREEN'S FUNCTIONS AND SINGULAR VALUES DECOMPOSITION

A. Direct scattering by Green's functions

Let us consider the stationary electron Schrödinger equation:

$$\left[-\frac{\hbar^2}{2m}\nabla^2+V(\mathbf{r})\right]\Psi(\mathbf{r})=E\Psi(\mathbf{r}). \qquad (1)$$

If $\Psi_0(\mathbf{r})$ is the solution for $V=0$ agreeing with the scattering boundary conditions (i.e., vanishing value and deriva-

0163-1829/2000/62(23)/16023(8)/$15.00 ©2000 The American Physical Society

tive ad infinitum), the solution of the full perturbed equation is given by the Lippmann-Schwinger equation:

$$\Psi(\mathbf{r})=\Psi_0(\mathbf{r})+\int\int\int dv' G(\mathbf{r},\mathbf{r}',E)V(\mathbf{r}')\Psi(\mathbf{r}'). \quad (2)$$

The kernel in this integral equation is the vacuum Green's function G which reads:

$$G(\mathbf{r},\mathbf{r}',E)=-\frac{m}{2\pi\hbar^2}\frac{1}{|\mathbf{r}-\mathbf{r}'|}e^{i\sqrt{(2mE/\hbar^2)}|\mathbf{r}-\mathbf{r}'|}. \quad (3)$$

The integration is naturally reduced to regions of space where $V(\mathbf{r}')\neq 0$.

B. Matricial formulation of the scattering process

The Lippmann-Schwinger equation (2) is a Fredholm equation of the second kind[33,34] and the Fredholm method can be applied to obtain a numerical solution. The first step is to perform a discretization of the regions where $V(\mathbf{r}')\neq 0$. Each discretization cell is assumed to identify a volume Δv_i. Labeling them by $i=1,\ldots,n_{\mathrm{obj}}$, the discrete Lippmann-Schwinger equation is

$$\Psi(\mathbf{r})=\Psi_0(\mathbf{r})+\sum_{i=1}^{n_{\mathrm{obj}}}\Delta v_i G(\mathbf{r},\mathbf{r}_i,E)V(\mathbf{r}_i)\Psi(\mathbf{r}_i). \quad (4)$$

The usual technique used to deal with the direct-scattering problem is to solve the linear system of n_{obj} equations obtained by writing Eq. (4) for $\mathbf{r}=\mathbf{r}_i$, $i=1,\ldots,n_{\mathrm{obj}}$. The resulting set of values $\{\Psi(\mathbf{r}_i)\}_{i=1,n_{\mathrm{obj}}}$ can then be used to compute $\Psi(\mathbf{r})$ anywhere. The elements $G(\mathbf{r}_i,\mathbf{r}_i,E)$ are calculated by integration over a spherical cell with volume Δv_i in order to avoid the divergence appearing when $|\mathbf{r}-\mathbf{r}'|$ cancels.

In our inverse scattering problem, the unknown quantities are the potential-energy values $V(\mathbf{r}_i)$, which are to be determined from the amplitude of the wave function $\Psi(\mathbf{r}_j), j=1,\ldots,n_{\mathrm{scr}}$ that is measured on n_{scr} points of a distant screen. We will assume the potential energy to be real valued (i.e., neglect energy transfers between the imaging electron and the sample) and the unperturbed solution $\Psi_0(\mathbf{r})$ to be known.

Let us refer to $\mathbf{\Psi}_{\mathrm{scr}}$ and $\mathbf{\Psi}^0_{\mathrm{scr}}$ as two vectors containing the values of $\Psi(\mathbf{r}_j)$ and $\Psi_0(\mathbf{r}_j)$ on the n_{scr} points of a distant screen. The two vectors $\mathbf{\Psi}_{\mathrm{obj}}$ and $\mathbf{\Psi}^0_{\mathrm{obj}}$ contain the corresponding values on the n_{obj} discretization cells of the object. Let $\mathbf{V}$ be a $n_{\mathrm{obj}}\times n_{\mathrm{obj}}$ diagonal matrix containing the unknown values of the potential energy in these n_{obj} diagonalization cells. If the $n_{\mathrm{scr}}\times n_{\mathrm{obj}}$ matrix $\mathbf{G}_{\mathrm{scr,obj}}$ contains the quantities $G(\mathbf{r}_j,\mathbf{r}_i,E)\Delta v_i$ (the $\mathbf{r}_j$ referring to points on the screen and the $\mathbf{r}_i$ to points in the object) and if the $n_{\mathrm{obj}}\times n_{\mathrm{obj}}$ matrix $\mathbf{G}_{\mathrm{obj,obj}}$ contains the quantities $G(\mathbf{r}_j,\mathbf{r}_i,E)\Delta v_i$ (the $\mathbf{r}_j$ and $\mathbf{r}_i$ referring both to points in the object), Eq. (4) leads to the system

$$\mathbf{\Psi}_{\mathrm{obj}}=\mathbf{\Psi}^0_{\mathrm{obj}}+\mathbf{G}_{\mathrm{obj,\,obj}}V\mathbf{\Psi}_{\mathrm{obj}}, \quad (5)$$

$$\mathbf{\Psi}_{\mathrm{scr}}=\mathbf{\Psi}^0_{\mathrm{scr}}+\mathbf{G}_{\mathrm{scr,\,obj}}V\mathbf{\Psi}_{\mathrm{obj}}, \quad (6)$$

where these two matricial equations correspond to the two steps of the resolution, i.e., determination of the wave-function values $\Psi(\mathbf{r}_i)$ in the object and far propagation to the screen.

C. Formulation as a Born series

The matricial equation (5) enabling the computation of the wave function in the sample can be written in the form

$$\mathbf{\Psi}_{\mathrm{obj}}=[\mathbf{I}-\mathbf{G}_{\mathrm{obj,\,obj}}\mathbf{V}]^{-1}\mathbf{\Psi}^0_{\mathrm{obj}} \quad (7)$$

$$=\mathbf{\Psi}^0_{\mathrm{obj}}+[\mathbf{G}_{\mathrm{obj,\,obj}}\mathbf{V}]\mathbf{\Psi}^0_{\mathrm{obj}}+[\mathbf{G}_{\mathrm{obj,\,obj}}\mathbf{V}]^2\mathbf{\Psi}^0_{\mathrm{obj}}+\cdots, \quad (8)$$

provided this last series is convergent (i.e., $\mathbf{G}_{\mathrm{obj,\,obj}}\mathbf{V}$ is a contraction).

The terms in this last expression refer to the Born expansion of the wave function. Let us refer by $\mathbf{\Psi}^k_{\mathrm{obj}}$ to the kth Born expansion term $(k=0,\ldots,\infty)$ of the wave function in the object, i.e.,

$$\mathbf{\Psi}^k_{\mathrm{obj}}=[\mathbf{G}_{\mathrm{obj,\,obj}}\mathbf{V}]^k\mathbf{\Psi}^0_{\mathrm{obj}}, \quad (9)$$

with the convention that $[\mathbf{G}_{\mathrm{obj,\,obj}}\mathbf{V}]^0=\mathbf{I}$, the identity matrix. According to Eq. (6), one can write the kth Born expansion term of the wave function on the screen as

$$\mathbf{\Psi}^k_{\mathrm{scr}}\overset{k=0}{=}\mathbf{\Psi}^0_{\mathrm{scr}} \quad (10)$$

$$\overset{k>0}{=}\mathbf{G}_{\mathrm{scr,\,obj}}\mathbf{V}\mathbf{\Psi}^{k-1}_{\mathrm{obj}}=\mathbf{G}_{\mathrm{scr,\,obj}}\mathbf{V}[\mathbf{G}_{\mathrm{obj,\,obj}}\mathbf{V}]^{k-1}\mathbf{\Psi}^0_{\mathrm{obj}}. \quad (11)$$

Defining $\mathbf{\Psi}^{i-j}_{\mathrm{scr}}=\Sigma^j_{k=i}\mathbf{\Psi}^k_{\mathrm{scr}}$ and $\mathbf{\Psi}^{i-j}_{\mathrm{obj}}=\Sigma^j_{k=i}\mathbf{\Psi}^k_{\mathrm{obj}}$, one finds the relation

$$\mathbf{\Psi}_{\mathrm{scr}}=\mathbf{\Psi}^{0-\infty}_{\mathrm{scr}}=\mathbf{\Psi}^{0-k}_{\mathrm{scr}}+\mathbf{G}_{\mathrm{scr,\,obj}}\mathbf{V}\mathbf{\Psi}^{k-\infty}_{\mathrm{obj}}. \quad (12)$$

This relation will enable a term by term construction of the wave function on screen from the screen intensity distribution.

D. Iterative reconstruction of the scattered-wave function

In the first step of the iteration $(k=0)$, we have the equation $\mathbf{\Psi}_{\mathrm{scr}}=\mathbf{\Psi}^0_{\mathrm{scr}}+\mathbf{G}_{\mathrm{scr,\,obj}}\mathbf{V}\mathbf{\Psi}^{0-\infty}_{\mathrm{obj}}$, where $\mathbf{\Psi}^{0-\infty}_{\mathrm{obj}}$ is approximated by its major contribution, i.e., $\mathbf{\Psi}^{0-\infty}_{\mathrm{obj}}\simeq\mathbf{\Psi}^0_{\mathrm{obj}}$, which is associated with $\mathbf{V}=\mathbf{0}$.

In this first and subsequent steps of the algorithm, we have an approximation of $\mathbf{\Psi}^k_{\mathrm{obj}}$ and $\mathbf{\Psi}^{k-\infty}_{\mathrm{obj}}$ that is associated with an approximation of $\mathbf{\Psi}^{(k+1)-\infty}_{\mathrm{scr}}$ by $\mathbf{\Psi}^{(k+1)-\infty}_{\mathrm{scr}}=\mathbf{G}_{\mathrm{scr,\,obj}}\mathbf{V}\mathbf{\Psi}^{k-\infty}_{\mathrm{obj}}$.

The intensities on the n_{scr} points of the screen are given by

$$|\mathbf{\Psi}_{\mathrm{scr}}(\mathbf{r}_j)|^2=|\mathbf{\Psi}^{0-k}_{\mathrm{scr}}(\mathbf{r}_j)+\mathbf{\Psi}^{(k+1)-\infty}_{\mathrm{scr}}(\mathbf{r}_j)|^2$$

$$=|\mathbf{\Psi}^{0-k}_{\mathrm{scr}}(\mathbf{r}_j)|^2+|\mathbf{\Psi}^{(k+1)-\infty}_{\mathrm{scr}}(\mathbf{r}_j)|^2$$

$$+2\ \mathrm{Re}[\mathbf{\Psi}^{0-k}_{\mathrm{scr}}(\mathbf{r}_j)^*\mathbf{\Psi}^{(k+1)-\infty}_{\mathrm{scr}}(\mathbf{r}_j)]$$

$$=|\mathbf{\Psi}^{0-k}_{\mathrm{scr}}(\mathbf{r}_j)|^2+|\mathbf{\Psi}^{(k+1)-\infty}_{\mathrm{scr}}(\mathbf{r}_j)|^2$$

TABLE I. Scheme of the reconstruction algorithm.

Initialize:		$\mathbf{V}=0$
Loop n from 0 to N:	Initialize:	
		$\mathbf{\Psi}_{\rm obj}^{0-\infty}=\begin{cases}\mathbf{\Psi}_{\rm obj}^{0} & \text{if } n=0\\ \Sigma_{l=0}^{\infty}[\mathbf{G}_{\rm obj,\,obj}\mathbf{V}]^{l}\mathbf{\Psi}_{\rm obj}^{0} & \text{if } n>0\end{cases}$
	Loop k from 0 to n:	
		$\mathbf{\Psi}_{\rm scr}^{(k+1)-\infty}=\mathbf{G}_{\rm scr,\,obj}\mathbf{V}\mathbf{\Psi}_{\rm obj}^{k-\infty}$
		solve $\lvert\mathbf{\Psi}_{\rm scr}(\mathbf{r}_j)\rvert^2=\lvert\mathbf{\Psi}_{\rm scr}^{0-k}(\mathbf{r}_j)+\mathbf{\Psi}_{\rm scr}^{(k+1)-\infty}(\mathbf{r}_j)\rvert^2$ to determine V
		$\mathbf{\Psi}_{\rm scr}^{0-(k+1)}=\mathbf{\Psi}_{\rm scr}^{0-k}+G_{\rm scr,\,obj}V\mathbf{\Psi}_{\rm obj}^{k}$
		$\mathbf{\Psi}_{\rm obj}^{k+1}=[\mathbf{G}_{\rm obj,obj}\mathbf{V}]\mathbf{\Psi}_{\rm obj}^{k}$
		$\mathbf{\Psi}_{\rm obj}^{(k+1)-\infty}=[\mathbf{G}_{\rm obj,obj}\mathbf{V}]\mathbf{\Psi}_{\rm obj}^{k-\infty}$
	Finalize:	
		solve $\mathbf{\Psi}_{\rm scr}^{0-(n+1)}-\mathbf{\Psi}_{\rm scr}^{0}=\mathbf{G}_{\rm scr,\,obj}[\mathbf{V}\mathbf{\Psi}_{\rm obj}]$ to determine $[\mathbf{V}\mathbf{\Psi}_{\rm obj}]$
		$V(\mathbf{r}_j)=\dfrac{[\mathbf{V}\mathbf{\Psi}_{\rm obj}]_j}{[\mathbf{\Psi}_{\rm obj}^{0}+\mathbf{G}_{\rm obj,obj}\mathbf{V}\mathbf{\Psi}_{\rm obj}]_j}$

$$+\sum_{l=1}^{n_{\rm obj}}\{2\ \mathrm{Re}[\mathbf{\Psi}_{\rm scr}^{0-k}(\mathbf{r}_j)^{*}G(\mathbf{r}_j,\mathbf{r}_l,E)\times\Delta v_l\mathbf{\Psi}_{\rm obj}^{k-\infty}(\mathbf{r}_l)]\}V(\mathbf{r}_l). \tag{13}$$

This equation is linear in $V(\mathbf{r}_l)$ and can be written in the form $\mathbf{Ax}=\mathbf{b}$, where $\mathbf{x}$ is a vector containing the $n_{\rm obj}$ values of $V(\mathbf{r}_l)$, $\mathbf{b}$ a vector containing the $n_{\rm scr}$ values of $\lvert\mathbf{\Psi}_{\rm scr}(\mathbf{r}_j)\rvert^2-\lvert\mathbf{\Psi}_{\rm scr}^{0-k}(\mathbf{r}_j)\rvert^2-\lvert\mathbf{\Psi}_{\rm scr}^{(k+1)-\infty}(\mathbf{r}_j)\rvert^2$ and $\mathbf{A}$ a $n_{\rm scr}\times n_{\rm obj}$ matrix containing the coefficients $2\ \mathrm{Re}[\mathbf{\Psi}_{\rm scr}^{0-k}(\mathbf{r}_j)^{*}G(\mathbf{r}_j,\mathbf{r}_l,E)\Delta v_l\mathbf{\Psi}_{\rm obj}^{k-\infty}(\mathbf{r}_l)]$. This equation can be solved by a singular values decomposition (see, the Appendix), thus providing the $V(\mathbf{r}_l)$ that give the best agreement (in the least-squares sense) with the known values of $\lvert\mathbf{\Psi}_{\rm scr}(\mathbf{r}_j)\rvert^2$.

The potential-energy matrix $\mathbf{V}$ is then used to update the estimation of the wave function on the screen by

$$\mathbf{\Psi}_{\rm scr}^{0-(k+1)}=\mathbf{\Psi}_{\rm scr}^{0-k}+\mathbf{G}_{\rm scr,\,obj}\mathbf{V}\mathbf{\Psi}_{\rm obj}^{k}. \tag{14}$$

It is also used to provide the estimation of $\mathbf{\Psi}_{\rm obj}^{k+1}$ and $\mathbf{\Psi}_{\rm obj}^{(k+1)-\infty}$ for the next iteration by

$$\mathbf{\Psi}_{\rm obj}^{k+1}=[\mathbf{G}_{\rm obj,obj}\mathbf{V}]\mathbf{\Psi}_{\rm obj}^{k}, \tag{15}$$

$$\mathbf{\Psi}_{\rm obj}^{(k+1)-\infty}=[\mathbf{G}_{\rm obj,obj}\mathbf{V}]\mathbf{\Psi}_{\rm obj}^{k-\infty}. \tag{16}$$

E. Determination of the potential energy from the scattered-wave function

Once the scattered-wave function $\mathbf{\Psi}_{\rm scr}$ has been reconstructed, the equation

$$\mathbf{\Psi}_{\rm scr}-\mathbf{\Psi}_{\rm scr}^{0}=\mathbf{G}_{\rm scr,\,obj}[\mathbf{V}\mathbf{\Psi}_{\rm obj}]. \tag{17}$$

can be solved by computing the singular values decomposition of the matrix $\mathbf{G}_{\rm scr,\,obj}$, thus providing the estimation of $[\mathbf{V}\mathbf{\Psi}_{\rm obj}]$ giving the best agreement (in the least-squares sense) with the long-range (scattered) wave function values $\mathbf{\Psi}_{\rm scr}-\mathbf{\Psi}_{\rm scr}^{0}$.

The potential energy is then found by

$$V(\mathbf{r}_j)=\frac{[\mathbf{V}\mathbf{\Psi}_{\rm obj}]_j}{[\mathbf{\Psi}_{\rm obj}^{0}+\mathbf{G}_{\rm obj,obj}\mathbf{V}\mathbf{\Psi}_{\rm obj}]_j}. \tag{18}$$

F. Comments on the algorithm efficiency

Although an estimation of V is obtained during the reconstruction of the long-range wave function, it is of poor quality compared to the final result obtained by Eq. (18). The estimation of V resulting from Eq. (13) is only an intermediate result enabling the iterative reconstruction of the wave function on the screen. Its accuracy is limited by the starting approximation on $\mathbf{\Psi}_{\rm obj}^{0-\infty}$.

When repeating the whole reconstruction process, a better accuracy can be achieved by evaluating this starting quantity by $\mathbf{\Psi}_{\rm obj}^{0-\infty}=\Sigma_{l=0}^{\infty}[\mathbf{G}_{\rm obj,\,obj}\mathbf{V}]^{l}\mathbf{\Psi}_{\rm obj}^{0}$, where $\mathbf{V}$ results from Eq. (18) in the final step of the preceding reconstruction. The algorithm should run from $k=0$ to $k=0$ the first time, then from $k=0$ to $k=1$ the second time and so on (see Table I for a scheme of the reconstruction algorithm). It is found that running the algorithm a single time already provides a good result, whose accuracy is comparable with that obtained by performing more iterations. This is related to the high-order terms of the Born series having a negligible contribution to the diffraction figure, in the situations considered in the present paper.

The long-range intensities should be renormalized to the values that would have been observed on a spherical screen. In this way, the $1/r$ factor appearing in the asymptotic behavior of the wave function takes the same value on the whole screen so this geometrical factor is removed from the information the singular values decomposition has to deal with.

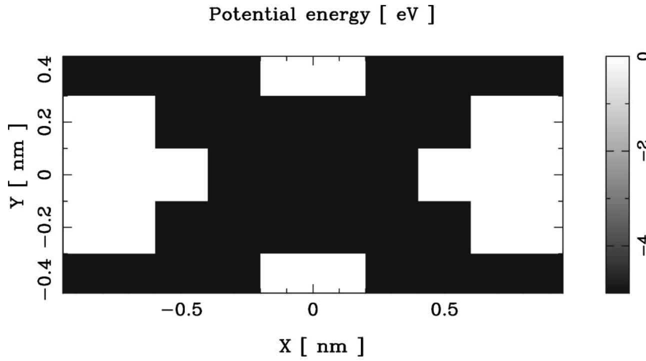

FIG. 1. Potential-energy distribution (in eV) in the central xy plane of the sample.

To prevent high-frequency oscillations to appear in the reconstruction, it was found that only the singular values that are higher than 1/19 of their maximum should be considered (or even a more restricted set of singular values in the case of strong noise). The only situations where the reconstruction completely fail occur when the singular values decompositions required for the reconstruction of the long-range wave function cannot be achieved, usually due to the dimensions of the screen being too small.

III. APPLICATION TO THE OBSERVATION OF A HOMOGENEOUS OBJECT

A. Description of the situation

In this paper, we will consider the reconstruction of a nanometric sample, whose shape is given in Fig. 1. It is an X-like object characterized by a 5 eV potential well (this is the typical value for the electron affinity of a carbon material). There is no surrounding electric field so the potential energy goes directly from -5 eV in the sample to 0 eV outside. The width of the object is 0.1 nm.

The imaging electrons have an energy of 25 eV and are described by a spherical wave whose point source is on the z axis. This axis is perpendicular to the sample and has its origin in the object center. The wavelength λ is 0.248 nm in the vacuum and 0.226 nm in the sample, so some of its parts have dimensions that are lower than the electron wavelength and lower than the resolution limit $\Delta d = 1/2\sqrt{\lambda d}$ for source-sample distances d higher than 0.65 nm.

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 corresponding (renormalized) intensity distribution is illustrated in the left part of Fig. 2. The screen is 30 cm wide and has a 35×35 resolution. The reconstruction resulting from five iterations of the algorithm given in the previous section is presented in the right part of this figure. A horizontal section of this reconstruction is given in Fig. 3, where it is compared with the true potential-energy distribution and with the result obtained with 40 eV energy. The reconstruction turns out to be improved by this rise of energy.

The mean relative error is evaluated by

$$\bar{E} = \frac{1}{n_{\text{obj}}} \sum_{i=1}^{n_{\text{obj}}} |\mathbf{V}_{i,i}^{\text{true}} - \mathbf{V}_{i,i}^{\text{recons}}|/(5\ \ \text{eV}),$$

where n_{obj} is the number of discretization cells in and around the sample and where the $\mathbf{V}_{i,i}^{\text{true}}$ and $\mathbf{V}_{i,i}^{\text{recons}}$ refer to the corresponding true and reconstructed potential-energy values.

When the sample is reconstructed in a $2\times1\times1$ nm^3 support (with the correct z position), the mean relative error is 6.8%. This mean relative error is reduced to 3.2% when the electron energy is increased from 25 to 40 eV and to 5.7% when the potential energy is not allowed to take values lower than -5 eV and higher than 0 eV. Increasing the electron energy thus results in a better sample reconstruction, which is expected by the wavelength and resolution limit being reduced. The next simulations will assume a 25 eV energy. Enforcing the potential energy to take values restricted to the $[-5,0]$ eV range is justified when the physical properties of the sample (i.e., its internal potential energy) are known. Increasing the resolution of the diffraction figure does not improve the reconstruction, unlike what is observed with noisy

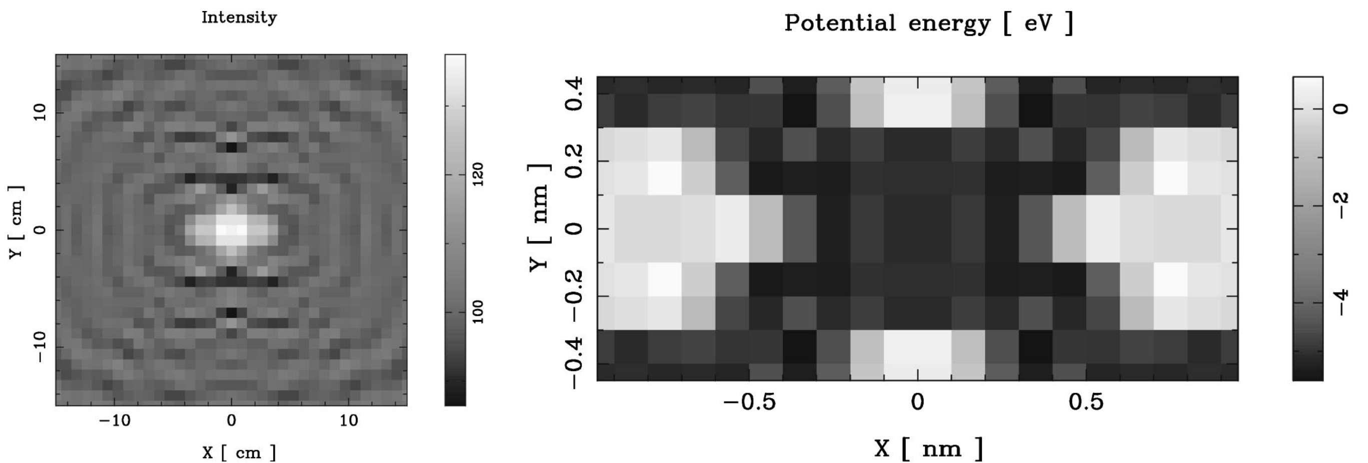

FIG. 2. Left: renormalized Fresnel diffraction figure corresponding to a 3-nm-source-sample distance and a 25-eV electron energy. The screen is 10 cm away from the sample; right: reconstruction of the potential energy in the central xy plane of the sample.

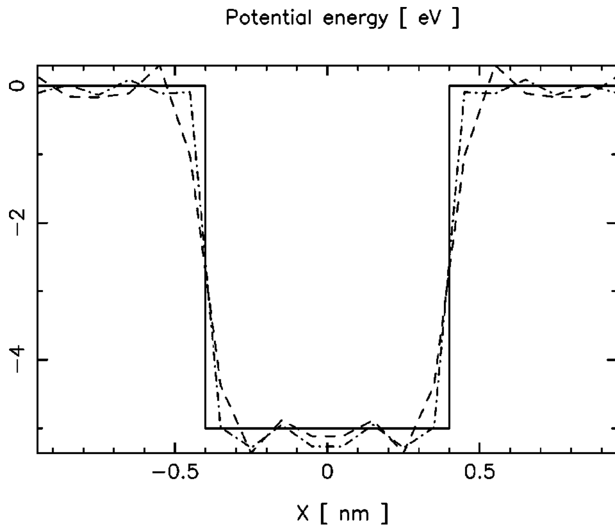

FIG. 3. Potential energy on the central x axis of the sample: exact result (full), reconstruction at an energy of 25 eV (dashed), reconstruction at an energy of 40 eV (dot dashed).

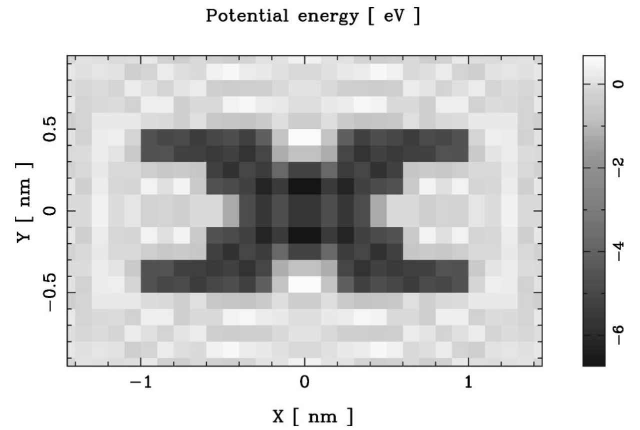

FIG. 4. Reconstruction of the potential energy in the central xy plane of the sample in a $3\times2\times1$ nm^3 extended support. The screen is 30 cm wide and the electron energy is 25 eV.

data. Finally, making a reconstruction from the *exact* wave function on the screen (i.e., applying only the technique of Sec. II E with the true $\mathbf{\Psi}_{\rm scr}$) provides a result undistinguishable from Fig. 1, with a mean relative error of 10^{-9}!

When the dimensions of the sample support is extended from $2\times1\times1$ nm^3 to $3\times2\times1$ nm^3, the mean relative error takes the value of 5.6% when there is no restriction on the potential-energy values and 3.1% otherwise. A single iteration of the algorithm was performed, because the singular values decomposition required in Sec. II E could not be achieved (the solution proposed in the Appendix was applied, but running the algorithm several times does not improve the results when this situation occurs). The reconstruction is presented in Fig. 4.

For the reconstruction to be successful, the singular values decomposition required in Sec. II D must be achieved (the success of the decomposition required in Sec. II E is less important). This decomposition fails when the dimensions of the screen are not large enough. In the case of Fresnel diffraction, the screen has to be around 2.5 times larger than the geometrical projection of the sample *support* from the point source. The simulations presented so far considered a screen, respectively, 4.5 and 3 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. In the first case, a *relative* noise was considered, i.e., a noise whose amplitude in each screen pixel is proportional to the corresponding true intensity value. In the second case, an *absolute* noise was taken into account, i.e., a noise whose amplitude (everywhere) is proportional to the true maximal intensity encountered on the screen. Finally, a *mixed* noise containing both relative and absolute noises was considered.

In the first simulation of this series, the extreme situation of a 30% relative noise was considered. The resolution of the screen was increased to 105×105 pixels. The mean relative error on the sample reconstruction is 22% without potential-energy restrictions and 14% otherwise. However, by keeping only the singular values σ_i whose ratio to their maximal value is higher than the noise level (i.e., $\sigma_i/\sigma_{\rm max}>0.3$), the mean relative error on the reconstruction is reduced to 9.6% with the potential-energy restriction (and 15% without). The diffraction figure and the corresponding reconstruction are presented in Fig. 5.

In the second simulation, a 7% absolute noise was considered. All conditions are the same as in the previous simulation. The mean relative error on the sample reconstruction is 29% without potential-energy restrictions and 11% otherwise. Again, by keeping only the singular values σ_i whose ratio to their maximal value is higher than the noise level (i.e., $\sigma_i/\sigma_{\rm max}>0.07$), the mean relative error on the reconstruction is reduced to 9.2% with the potential-energy restriction (and 23% without). The diffraction figure and the corresponding reconstruction are presented in Fig. 6.

Finally, a combination of 5% relative and 3% absolute noises was considered, these values being more typical of a ''correct'' experiment. The observation conditions are unchanged. The mean relative error on the sample reconstruction is 14% without potential-energy restrictions and 7.2% otherwise. By keeping these potential-energy restrictions and increasing the electron energy from 25 to 40 eV, a 4.2% mean relative error is achieved. The noise level being lower than the 1/19 reference value, no improvement results from a further singular values selection. The diffraction figure and the corresponding reconstruction are presented in Fig. 7.

These three simulations thus prove the stability of our algorithm against random noise. The quality of the reconstruction can be substantially improved first by using exclusively the singular values whose ratio to their maximal value is higher than the noise level and also by restricting the potential energy when knowledge of the physical properties of the sample allows so. It also turns out that increasing the resolution and dimensions of the screen leads to better results. Our algorithm is however less efficient in Fraunhofer conditions, where random noise levels of a few percent can spoil the reconstruction.

D. Comparison with two other reconstruction techniques

We compared these results with those obtained by two other reconstruction techniques. The first is due to Bleloch

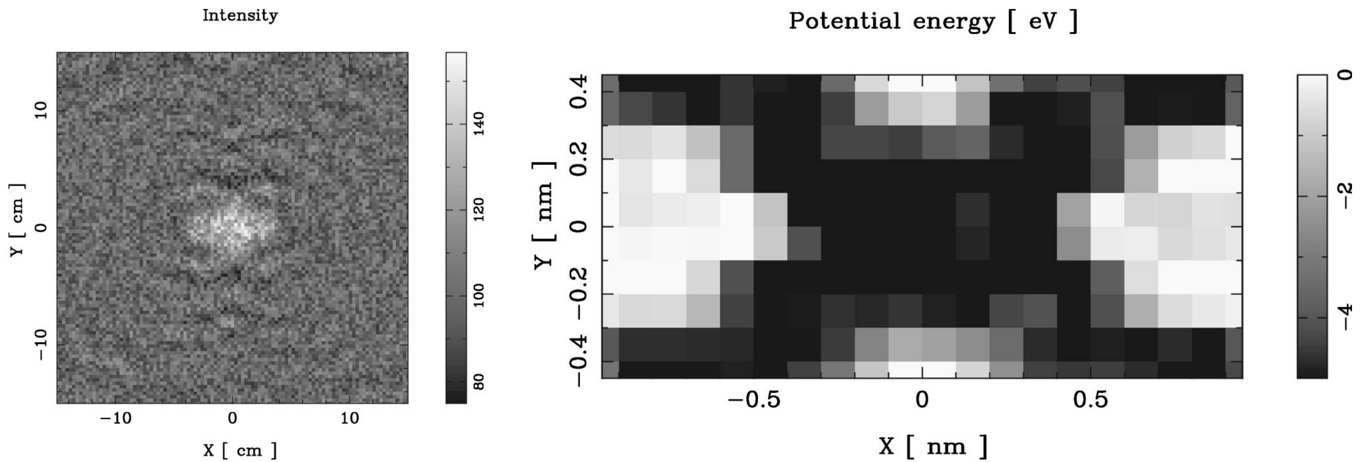

FIG. 5. Left: renormalized Fresnel diffraction figure corresponding to a 3-nm-source-sample distance and a 25-eV electron energy. The screen is 10 cm away from the sample and contains a 30% relative noise; right: reconstruction of the potential energy in the central xy plane of the sample (with restriction to the $[-5,0]$ eV range and consideration of $\sigma_i > 0.3\sigma_{\max}$).

et al.[31] and consists of an iterative back-and-forth propagation of the wave function between the observation screen and the sample support. The wave function is enforced on the screen to have the amplitude corresponding to the measured intensity distribution, while on the sample support it is enforced to have the phase of the incident wave. The second reconstruction technique is due to Spence *et al.*[32] and consists of a backpropagation of the intensity distribution from the screen to the sample. The reconstruction is not iterative and is affected by the twin image and second-order effects. The underlying formalism of these two techniques is essentially the Fresnel-Kirchhoff theory.[4]

In both cases, the description of the sample is reduced to a two-dimensional mask, whereas the technique of this paper is suited for a three-dimensional reconstruction. Due to the underlying Fresnel-Kirchhoff theory and due to the phase of the incident wave being not affected by the sample (in Bleloch's method) and second-order effects having to be negligible (in Spence's method), these two techniques are essentially first-order treatments while our method deals with all orders of diffraction. They provide the shape of a two-dimensional projection of the sample, by the amplitude (of arbitrary magnitude) of the wave function in the sample and not by a reconstructed potential-energy distribution.

For these reasons, these two techniques are better suited for the reconstruction of (large) opaque samples or holes in opaque supports, where the corresponding assumptions are justified. They are not suited for the reconstruction of the transparent sample considered in this paper, where the wave function has comparable amplitude inside and outside the sample. We only got these techniques to work by considering the diffraction figure corresponding to the situation where the sample is a hole (with the same dimensions) in an opaque support. The best reconstruction was achieved with Bleloch's method: with a 35×35 resolution on the 15-cm-wide screen, a mean relative error around 15% was obtained (instead of our 6.8% value). In this situation, the number of iterations (around 20) required to reach convergence was larger than with our method.

Let us say finally that due to the use of discrete Fourier transforms in these two techniques, there are constraining relations between the resolution in the sample and the reso-

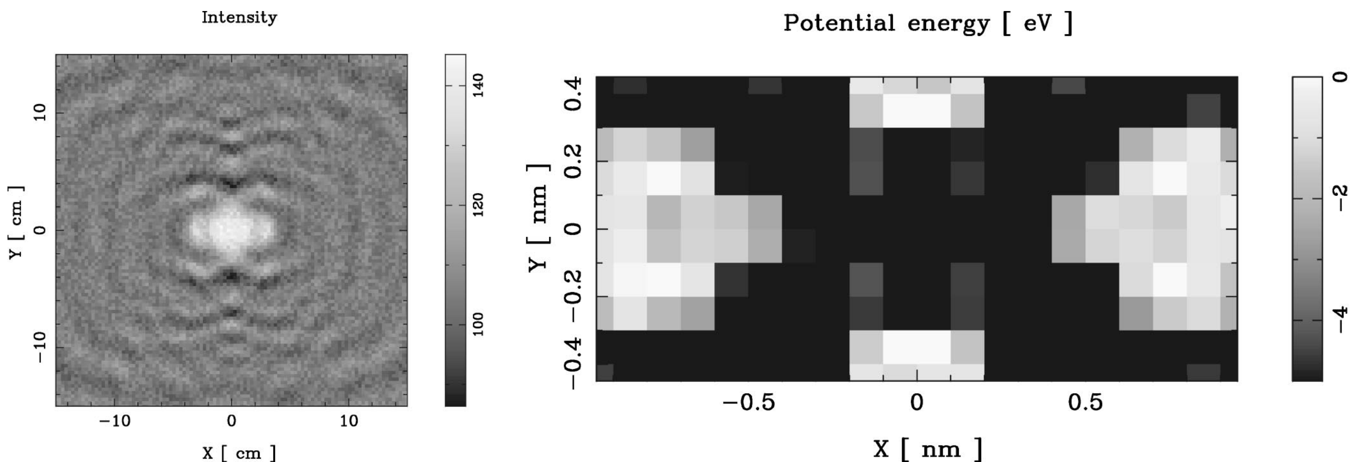

FIG. 6. Left: renormalized Fresnel diffraction figure corresponding to a 3-nm-source-sample distance and a 25-eV electron energy. The screen is 10 cm away from the sample and contains 7% absolute noise; right: reconstruction of the potential energy in the central xy plane of the sample (with restriction to the $[-5,0]$ eV range and consideration of $\sigma_i > 0.07\sigma_{\max}$).

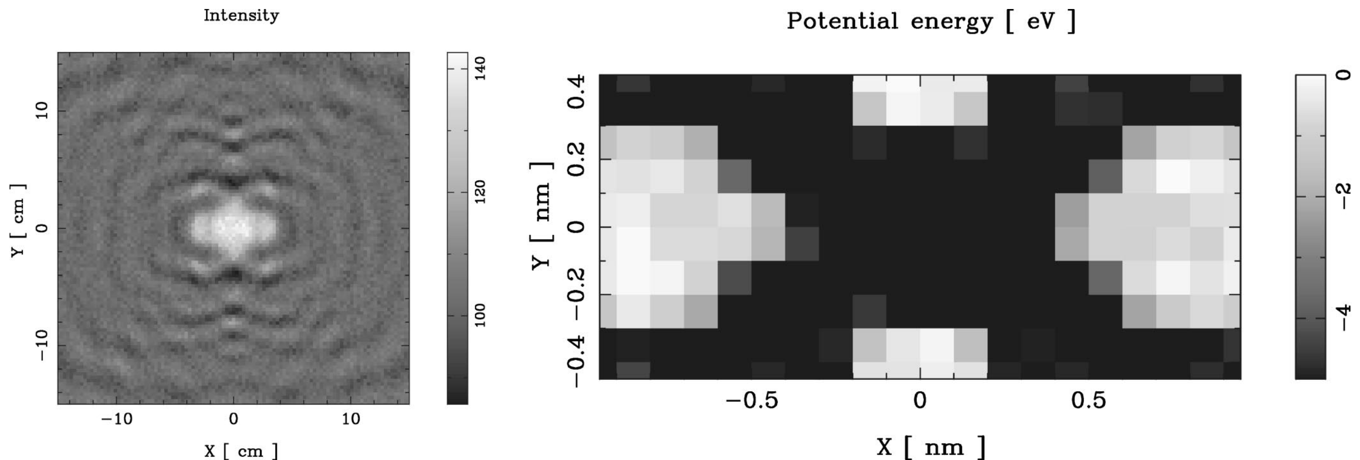

FIG. 7. Left: renormalized Fresnel diffraction figure corresponding to a 3-nm-source-sample distance and a 25-eV electron energy. The screen is 10 cm away from the sample and contains 5% relative and 3% absolute noises; right: reconstruction of the potential energy in the central xy plane of the sample (with restriction to the $[-5,0]$ eV range).

lution on the screen, while these parameters are independent in our method. The space requirements of the Bleloch and Spence's techniques are however far less important.

IV. CONCLUSION

In this paper, an inverse scattering technique was presented. In its formulation, this technique takes account of all orders of diffraction by performing an iterative reconstruction of the wave function on the observation screen. For the present application, it turns out that the results are not significantly improved by running more than a few iterations of the algorithm. This is related to the sample being so thin (0.1 nm thickness) that the scattering process is essentially reduced to its first-order contribution, the high-order contributions being smaller in magnitude than the noise or numerical accuracy.

The algorithm turns out to give better results in the case of Fresnel diffraction. This could be related on one side to the inhomogeneity of the incident wave in the sample (which plays an important role as a reference for the reconstruction) and on the other side to the diffraction figure being more contrasted than in Fraunhofer diffraction.

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. This idea is similar to applying a low-pass filter to the diffraction figure. This filtering of information is however achieved directly by the technique and does not require any preliminary treatment. To help the algorithm separating useful information from noise, it is useful to increase the resolution and size of the screen (so more diffraction fringes are taken into account). The results are also improved by increasing 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 invert/decompose the two matrices $\mathbf{G}_{\mathrm{obj,obj}}$ and $\mathbf{G}_{\mathrm{scr,obj}}$ whose size is, respectively, proportional to n_{obj}^2 and $n_{\mathrm{scr}}n_{\mathrm{obj}}$, n_{obj} and n_{scr} being, respectively, the number of discretization cells in the sample and the number of pixels on the screen. Extending significantly the range of the technique will require further developments or the use of out-of-core techniques.

Due to the phase-recovering part of the algorithm (see Sec. II D), this technique is limited to real valued potential energies. Absorption processes are therefore not considered and in its present form this technique will apply essentially to small transparent samples (for which the most interpretation is needed).

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 the present paper the sample is essentially two dimensional (no variation of its structure along z). In its formulation, our technique is not limited to this peculiar case and, due to the inhomogeneity of the incident wave, should already provide good results for true three-dimensional samples. However, it should be useful to consider simultaneously the diffraction figures obtained for different source positions (thus performing a diffraction tomography). The present formalism is already suited to such a description since nothing prevents the $\mathbf{\Psi}$ vectors to contain values associated with different source positions. It could also be useful to take simultaneously into account projections obtained for different energy values. Finally enforcing the sample to have a minimal curvature could still improve the reconstruction.

ACKNOWLEDGMENTS

A.M. was supported by the Belgian National Fund for Scientific Research (FNRS). The authors acknowledge 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. The authors thank A. Istace and G. Gilquin for useful discussions.

APPENDIX: SOLUTION OF AX = B BY SINGULAR VALUES DECOMPOSITION

Computing the singular values decomposition (SVD) of a $m\times n$ matrix $\mathbf{A}$ consists of expressing it as

$$\mathbf{A}=\mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\dagger}, \tag{A1}$$

where $\mathbf{U}$ and $\mathbf{V}$ are, respectively, two $m\times n$ and $n\times n$ matrices with orthonormalized columns. The $n\times n$ diagonal matrix $\mathbf{\Sigma}$ contains the singular values σ_i.

Given an m-vector $\mathbf{b}$, the SVD technique can be used to find a solution $\mathbf{x}$, of the system $\mathbf{Ax}=\mathbf{b}$. It is an n-vector given by

$$\mathbf{x}=\mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^{\dagger}\mathbf{b}. \tag{A2}$$

The elements of the diagonal matrix $\mathbf{\Sigma}^{-1}$ are given by

$$\mathbf{\Sigma}^{-1}_{\mathrm{i,i}}=\begin{cases}\sigma_{\mathrm{i}}^{-1} & \text{if } \sigma_i\neq 0\\ 0 & \text{if } \sigma_i=0.\end{cases} \tag{A3}$$

This solution $\mathbf{x}$ has the property to minimize (in the least-square sense) the norm of the ''error'' $|\mathbf{Ax}-\mathbf{b}|$.

In some cases (in this paper when the screen is not large enough), the matrix $\mathbf{\Sigma}$ can only be reduced to a bidiagonal (instead of a diagonal) form. The matrix $\mathbf{\Sigma}^{-1}$ is then the usual inverse of $\mathbf{\Sigma}$. These situations do not spoil the reconstruction if they occur in the part of the algorithm described in Sec. II E.

*Author to whom correspondence should be addressed. Electronic address: alexandre.mayer@fundp.ac.be

[1] H.-W. Fink, W. Stocker, and H. Schmid, Phys. Rev. Lett. **65**, 1204 (1990).
[2] V.T. Binh, V. Semet, and N. Garcia, Ultramicroscopy **58**, 307 (1995).
[3] V.T. Binh and V. Semet, Ultramicroscopy **73**, 107 (1998).
[4] V.T. Binh, V. Semet, N. Garcia, and L. Bitar, in *Optics at the Nanometer Scale*, edited by M. Nieto-Vesperinas and N. Garcia (ESA, Paris, 1996), pp. 277–296.
[5] E.N. Economou, *Green's Functions in Quantum Physics* (Springer-Verlag, Berlin, 1983).
[6] H.J. Kreuzer, K. Nakamura, A. Wierzbicki, H.-W. Fink, and H. Schmid, Ultramicroscopy **45**, 381 (1992).
[7] A. Mayer, A. Castiaux, and J.-P. Vigneron, Comput. Phys. Commun. **109**, 81 (1998).
[8] C. Adessi and M. Devel, Phys. Rev. A **60**, 2194 (1999).
[9] A. Mayer and J.-P. Vigneron, Phys. Rev. B **56**, 12 599 (1997).
[10] A. Mayer and J.-P. Vigneron, J. Phys.: Condens. Matter **10**, 869 (1998).
[11] A. Mayer and J.-P. Vigneron, Phys. Rev. B **60**, 2875 (1999).
[12] A. Mayer and J.-P. Vigneron, Phys. Rev. E **59**, 4659 (1999).
[13] A. Mayer and J.-P. Vigneron, Phys. Rev. E **60**, 7533 (1999).
[14] A. Mayer and J.-P. Vigneron, Phys. Rev. E **61**, 5953 (2000).
[15] A. Mayer and J.-P. Vigneron, J. Vac. Sci. Technol. B **17**, 506 (1999).
[16] A. Mayer and J.-P. Vigneron, Ultramicroscopy **79**, 35 (1999).
[17] A. Mayer, P. Senet, and J-P. Vigneron, J. Phys.: Condens. Matter **11**, 8617 (1999).
[18] M.H. Maleki and A.J. Devaney, J. Opt. Soc. Am. A **10**, 1086 (1993).
[19] A.J. Devaney, Ultrason. Imaging **4**, 336 (1982).
[20] T.C. Wedberg and J.J. Stamnes, Pure Appl. Opt. **4**, 39 (1995).
[21] M. Tegze, G. Faigel, S. Marchesini, M. Belakhovsky, and A.I. Chumakov, Phys. Rev. Lett. **82**, 4847 (1999).
[22] G. Faigel and M. Tegze, Rep. Prog. Phys. **62**, 355 (1999).
[23] R.A. Crowther, L.A. Amos, J.T. Finch, D.J. De Rosier, and A. Klug, Nature (London) **226**, 421 (1970).
[24] R.A. Crowther, D.J. De Rosier, and A. Klug, Proc. R. Soc. London, Ser. A **337**, 319 (1970).
[25] A. Klug and R.A. Crowther, Nature (London) **238**, 435 (1972).
[26] D.G. Fischer, Pure Appl. Opt. **7**, 1151 (1998).
[27] B. Chen, J.J. Stamnes, and K. Stamnes, Pure Appl. Opt. **7**, 1161 (1998).
[28] B. Chen, J.J. Stamnes, A.J. Devaney, H.M. Pedersen, and K. Stamnes, Pure Appl. Opt. **7**, 1181 (1998).
[29] A. BenIsrael and T. N. E. Greville, *Generalized Inverses. Theory and Applications* (Krieger, New York, 1980).
[30] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran*, 2nd ed. (Cambridge University Press, Cambridge, 1986), pp. 51–63.
[31] A.L. Bleloch, A. Howie, and E.M. James, Appl. Surf. Sci. **111**, 180 (1997).
[32] J.C.H. Spence, X. Zhang, and W. Qian, *Electron Holography*, edited by A. Tonomura, L.F. Allard, G. Pozzi, D.C. Joy, and Y. A. Ono (Elsevier Science B.V., Amsterdam, 1995), pp. 267–276.
[33] L. Rodberg and R. Thaler, *Introduction to the Quantum Theory of Scattering* (Academic, New York, 1967), Chap. 5.
[34] G.B. Arfken and H.J. Weber, *Mathematical Methods for Physicists*, 4th ed. (Academic, New York, 1995), Chap. 16.