Nonsquare transfer-matrix technique applied to the simulation of electronic diffraction by a three-dimensional circular aperture

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 11 March 1999; revised manuscript received 21 December 1999)

The transfer-matrix methodology is frequently used to deal with elastic scattering problems that require a solution of the Schrödinger or homogeneous Maxwell equations in the continuous part of their spectra. Until now, this technique was limited to representations associated with square transfer matrices. This paper extends the transfer-matrix methodology to enable consideration of general representations associated with nonsquare matrices. The theory is illustrated by the diffraction of a field-emitted electronic beam by a three-dimensional circular aperture. The application focuses on the dependence of the long-range angular spread on the aperture radius, by highlighting the effects of the field-emission tip shape and dimensions.

PACS number(s): 02.70.−c, 03.65.Fd, 61.14.Dc, 79.70+q

I. INTRODUCTION

Linear systems of differential equations are frequently encountered in theoretical physics. Such equations appear, indeed, when dealing with the Schrödinger equation in quantum mechanics or with the Maxwell equations in electromagnetism. A useful property that appears in these situations is the additivity of solutions. When an analytic solution is not obtainable, several numerical techniques exist to deal with these equations in the energy or frequency continuum.

The transfer-matrix methodology [1–7] is one of these techniques. To apply this methodology, the physical system considered should be located between two separate boundaries. Given a set of basis states used for the wave function expansion, the transfer matrices contain, for each state incident on one boundary of the system, the amplitudes of the corresponding transmitted and reflected states.

The method depends essentially on the additivity property of solutions and proceeds in two steps. In the first step, intermediate solutions associated with particular boundary conditions are constructed. These solutions are combined in the subsequent step to derive those corresponding to the incidence of a single basis state. Since a matrix inversion is involved in this second step, the number of basis states had until now to be the same at the two boundaries of the system so that the matrix to invert is square.

Considering different numbers of basis states at the two boundaries of the system can be useful when a large number of those states is required at one boundary only or when propagating all states through the system would make the matrix to invert singular. The extension of the transfer-matrix methodology presented in this paper makes it possible to consider different numbers of basis states at each boundary, by providing a technique to invert nonsquare matrices and enforcing a reflection of the states that cannot be propagated through the system.

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

This paper first presents in Sec. II the transfer-matrix technique with the extensions required to proceed with nonsquare matrices. For reasons of clarity, the theory is developed with the particular objective of computing the scattering through a narrow aperture but provides results that apply to general situations. In Sec. III, the technique is implemented to simulate electronic field emission from a conical/elliptical emitter and the subsequent beam diffraction by a three-dimensional circular aperture. The simulations focus on the dependence of the long-range angular spread on the aperture radius. The variations due to the shape and dimensions of the emitter are studied. The main features of the results turn out to be explained by Fraunhofer and Fresnel diffraction. The occurrence of Fresnel diffraction is strongly dependent on the spherical shape of the electronic beam when incident on the aperture. A high degree of sphericity requires both a reduced field-emission area and radial surrounding electric fields. In the conditions of this paper, the position of the virtual projection point turns out to be strongly dependent on this last factor. The stability of the technique is demonstrated by comparing results obtained with different representations of the wave function.

II. THE GENERALIZED TRANSFER-MATRIX METHOD

A. Introduction

Let us consider scattering in a physical system made of three adjacent regions and let us assume the intermediate region to be the only diffusive part. The scanning tunneling microscope [8] and the Fresnel projection microscope [9] provide examples of such situations. Let us refer to the intermediate region as ''region II'' and the two other regions as ''region I'' and ''region III.'' Let z be a coordinate axis oriented from region I to region III, so that region II corresponds to the interval $0 \leqslant z \leqslant D$.

At this point, we should make the choice of simple basis states to represent the wave function in regions I and III. Let us write these states $\Psi_j^{\mathrm{I},\pm}$ in region I and $\Psi_j^{\mathrm{III},\pm}$ in region III. The sign $\pm$ stands for the direction of propagation relative to the z axis.

To motivate the generalization of the transfer-matrix methodology to nonsquare matrices, region III is assumed to

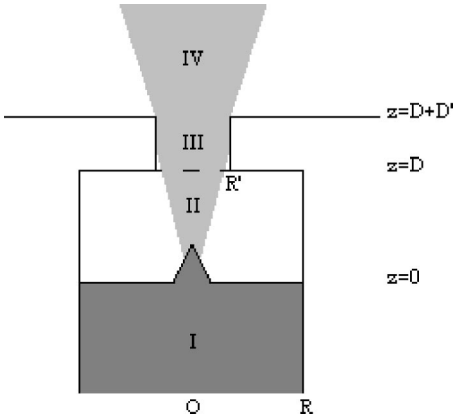

FIG. 1. Situation considered. A metal (region I) is followed by region II (which contains a field-emission tip), a circular aperture (region III), and an open space (region IV). The number of basis states needed for the wave function expansion is lower in region III than in region I and II.

be confined (compared to regions I and II) so that the number of basis states required for the wave function expansion in region III is lower than in the two other regions. This situation is illustrated in Fig. 1, where a metal (region I) supports a field-emission tip (in region II). This latter region is followed by a circular aperture (region III) and an open space (region IV). The quantization of the basis states in regions I and III is enforced by assuming the wave function to be confined in a cylinder with radius R in regions I and II and $R'<R$ in region III.

Our objective is to compute the coefficients in the wave function expansion of the reflected and transmitted states $\Psi_i^{\mathrm{I},-}$ and $\Psi_i^{\mathrm{III},+}$ corresponding to a single incident state $\Psi_j^{\mathrm{I},+}$ in region I. These coefficients will be stored in the so-called transfer matrices $\mathbf{t}_{\mathrm{I,III}}^{-+}$ and $\mathbf{t}_{\mathrm{I,III}}^{++}$. In the same way, the transfer matrices $\mathbf{t}_{\mathrm{I,III}}^{+-}$ and $\mathbf{t}_{\mathrm{I,III}}^{--}$ will contain the coefficients of the reflected and transmitted states $\Psi_i^{\mathrm{III},+}$ and $\Psi_i^{\mathrm{I},-}$ corresponding to a single incident state $\Psi_j^{\mathrm{III},-}$ in region III.

The computation will proceed in three stages. In the first stage, propagation through region II will be described by four square transfer matrices $\mathbf{t}_{\mathrm{I},D}^{\pm\pm}$, where $z=D$ is the boundary between regions II and III. A set of intermediate basis states $\Psi_i^{D,\pm}$ is required for the wave function expansion at the boundary $z=D$. Their number has to be identical with that relevant to region I. In a second stage, the connection between these states and the states $\Psi_i^{\mathrm{III},\pm}$ in region III will be described by four rectangular transfer matrices $\mathbf{t}_{\mathrm{D,III}}^{\pm\pm}$. Finally, the four transfer matrices $\mathbf{t}_{\mathrm{I,III}}^{\pm\pm}$ of interest are obtained by an appropriate combination of the eight matrices $\mathbf{t}_{\mathrm{I},D}^{\pm\pm}$ and $\mathbf{t}_{D,\mathrm{III}}^{\pm\pm}$.

B. Propagation through the intermediate region II by square transfer matrices

Scattering in the intermediate region II is described by four transfer matrices that contain the coefficients of the reflected and transmitted parts of the following solutions:

$$\Psi_j^{+\overset{z\leqslant 0}{=}}\Psi_j^{\mathrm{I},+}+\sum_i (t_{\mathrm{I},D}^{-+})_{i,j}\Psi_i^{\mathrm{I},-\overset{z=D}{=}}\sum_i (t_{\mathrm{I},D}^{++})_{i,j}\Psi_i^{D,+},$$

$$\Psi_j^{-\overset{z\leqslant 0}{=}}\sum_i (t_{\mathrm{I},D}^{--})_{i,j}\Psi_i^{\mathrm{I},-\overset{z=D}{=}}\Psi_j^{D,-}+\sum_i (t_{\mathrm{I},D}^{+-})_{i,j}\Psi_i^{D,+}, \tag{1}$$

where the solutions in the first set are those associated with a single incident state $\Psi_j^{\mathrm{I},+}$ in region I, while the solutions in the second set correspond to a single incident state $\Psi_j^{D,-}$ at $z=D$.

In order to derive these solutions, one starts by constructing two other sets of solutions

$$\overline{\Psi}_j^{+\overset{z\leqslant 0}{=}}\sum_i A_{i,j}^{+}\Psi_i^{\mathrm{I},+}+\sum_i B_{i,j}^{+}\Psi_i^{\mathrm{I},-\overset{z=D}{=}}\Psi_j^{D,+}, \tag{3}$$

$$\overline{\Psi}_j^{-\overset{z\leqslant 0}{=}}\Psi_j^{\mathrm{I},-\overset{z=D}{=}}\sum_i A_{i,j}^{-}\Psi_i^{D,-}+\sum_i B_{i,j}^{-}\Psi_i^{D,+} \tag{4}$$

that correspond to a single transmitted state $\Psi_j^{D,+}$ at $z=D$ or $\Psi_j^{\mathrm{I},-}$ in region I. These solutions are obtained by considering each transmitted state individually and propagating it (by using the relevant propagation equation) backward to the other boundary, where the corresponding solution is written as a combination of incident and reflected states.

Since the relevant propagation equation is linear, these solutions can be combined in order to derive the first two sets of solutions corresponding to a single incident state. The transfer matrices turn out to be related to the matrices $\mathbf{A}^{\pm}$ and $\mathbf{B}^{\pm}$ by

$$\mathbf{t}_{\mathrm{I},D}^{++}=(\mathbf{A}^{+})^{-1}, \tag{5}$$

$$\mathbf{t}_{\mathrm{I},D}^{-+}=(\mathbf{B}^{+})(\mathbf{A}^{+})^{-1}, \tag{6}$$

$$\mathbf{t}_{\mathrm{I},D}^{--}=(\mathbf{A}^{-})^{-1}, \tag{7}$$

$$\mathbf{t}_{\mathrm{I},D}^{+-}=(\mathbf{B}^{-})(\mathbf{A}^{-})^{-1}. \tag{8}$$

An efficient technique to control the numerical instabilities inherent in the computation of these transfer matrices is presented in Ref. [10]. Reference [11] contains techniques enabling one to take advantage of any symmetry present in the system.

C. Matching between the two representations by nonsquare transfer matrices

The fact that the basis states $\Psi_i^{\mathrm{D},\pm}$ and $\Psi_i^{\mathrm{III},\pm}$ are not in equal number does not prevent us from constructing solutions similar to Eqs. (3) and (4), namely,

$$\overline{\Psi}_j^{+\overset{z=D}{=}}\sum_i A_{i,j}^{+}\Psi_i^{D,+}+\sum_i B_{i,j}^{+}\Psi_i^{D,-\overset{z\geqslant D}{=}}\Psi_j^{\mathrm{III},+}, \tag{9}$$

$$\overline{\Psi}_j^{-\overset{z=D}{=}}\Psi_j^{D,-\overset{z\geqslant D}{=}}\sum_i A_{i,j}^{-}\Psi_i^{\mathrm{III},-}+\sum_i B_{i,j}^{-}\Psi_i^{\mathrm{III},+}. \tag{10}$$

The first set of solutions describes the projection of each state $\Psi_j^{\mathrm{III},+}$ on the states $\Psi_i^{D,\pm}$ at $z=D$, while the second set contains the projection of each state $\Psi_j^{D,-}$ on the states $\Psi_i^{\mathrm{III},\pm}$.

Let us write all these solutions in the compact form

$$(\ldots,\overline{\Psi}_j^{+},\ldots)\overset{z=D}{=}(\ldots,\Psi_i^{D,+},\ldots)\mathbf{A}^{+}+(\ldots,\Psi_{si}^{D,-},\ldots)\mathbf{B}^{+}\overset{z\geqslant D}{=}(\ldots,\Psi_j^{\mathrm{III},+},\ldots), \tag{11}$$

$$(\ldots,\overline{\Psi}_j^{-},\ldots)\overset{z=D}{=}(\ldots,\Psi_j^{D,-},\ldots)\overset{z\geqslant D}{=}(\ldots,\Psi_i^{\mathrm{III},-},\ldots)\mathbf{A}^{-}+(\ldots,\Psi_i^{\mathrm{III},+},\ldots)\mathbf{B}^{-}. \tag{12}$$

Since there are not as many basis states in region III as at $z=D$, the two matrices $\mathbf{A}^{+}$ and $\mathbf{B}^{+}$ have more rows than columns and the two matrices $\mathbf{A}^{-}$ and $\mathbf{B}^{-}$ fewer rows than columns. The relations (5)–(8) need to be generalized to take account of this nonsquare shape.

According to the generalized inverse theory [12], every matrix $\mathbf{A}$ $(m\times n)$ of rank r can be written as $\mathbf{A}=\mathbf{WH}$, where $\mathbf{W}$ and $\mathbf{H}$ are two matrices respectively $(m\times r)$ and $(r\times n)$ of rank r. The associated generalized inverse is a matrix $\mathbf{A}^{\#}$ $(n\times m)$ defined by $\mathbf{A}^{\#}=\mathbf{H}^{\dagger}(\mathbf{HH}^{\dagger})^{-1}(\mathbf{W}^{\dagger}\mathbf{W})^{-1}\mathbf{W}^{\dagger}$, where † stands for the transpose complex conjugate operation.

If the basis states $\Psi_j^{D,\pm}$ and $\Psi_j^{\mathrm{III},\pm}$ are orthonormal, the rank of the two matrices $\mathbf{A}^{\pm}$ is the number of states in region III and the generalized inverses of $\mathbf{A}^{+}$ and $\mathbf{A}^{-}$ are in this situation

$$(\mathbf{A}^{+})^{\#}=[(\mathbf{A}^{+})^{\dagger}(\mathbf{A}^{+})]^{-1}(\mathbf{A}^{+})^{\dagger}, \tag{13}$$

$$(\mathbf{A}^{-})^{\#}=(\mathbf{A}^{-})^{\dagger}[(\mathbf{A}^{-})(\mathbf{A}^{-})^{\dagger}]^{-1}. \tag{14}$$

It is easily checked that

$$(\mathbf{A}^{+})^{\#}(\mathbf{A}^{+})=\mathbf{I}, \tag{15}$$

$$(\mathbf{A}^{-})(\mathbf{A}^{-})^{\#}=\mathbf{I}. \tag{16}$$

However, we should be aware that in general

$$(\mathbf{A}^{+})(\mathbf{A}^{+})^{\#}\neq\mathbf{I}, \tag{17}$$

$$(\mathbf{A}^{-})^{\#}(\mathbf{A}^{-})\neq\mathbf{I}. \tag{18}$$

These inequalities come basically from the fact that an irrecoverable projection occurs while applying a transformation that is represented by a matrix with fewer rows than columns. In the situation considered, this irrecoverable projection is encountered with the part of the wave function that is incident on the side of the circular aperture, since the wave function expansion used in region III is limited to its behavior inside the circular aperture. For our technique to be consistent, these projected incoming states have to be properly reflected into their original region.

To see how to generalize the relations (5)–(8), let us multiply to the right the solutions (11) by $(\mathbf{A}^{+})^{\#}$ and the solutions (12) by $(\mathbf{A}^{-})^{\#}$. Considering the relations (15)–(18), one finds

$$(\ldots,\Psi_j^{+},\ldots)\overset{z=D}{=}(\ldots,\Psi_i^{D,+},\ldots)(\mathbf{A}^{+})(\mathbf{A}^{+})^{\#}+(\ldots,\Psi_i^{D,-},\ldots)(\mathbf{B}^{+})(\mathbf{A}^{+})^{\#}\overset{z\geqslant D}{=}(\ldots,\Psi_j^{\mathrm{III},+},\ldots)(\mathbf{A}^{+})^{\#}, \tag{19}$$

$$(\ldots,\Psi_j^{-},\ldots)\overset{z=D}{=}(\ldots,\Psi_j^{D,-},\ldots)(\mathbf{A}^{-})^{\#}\overset{z\geqslant D}{=}(\ldots,\Psi_i^{\mathrm{III},-},\ldots)\mathbf{I}+(\ldots,\Psi_i^{\mathrm{III},+},\ldots)(\mathbf{B}^{-})(\mathbf{A}^{-})^{\#}. \tag{20}$$

The solutions (19) are not in the appropriate form due to the factor $\mathbf{A}^{+}(\mathbf{A}^{+})^{\#}$, which does not simplify like $\mathbf{I}$. Considering previous comments about this inequality, the solution to this problem consists in enforcing explicitly a reflection of the projected incoming states $(\ldots,\Psi_i^{D,+},\ldots)[\mathbf{I}-(\mathbf{A}^{+})\times(\mathbf{A}^{+})^{\#}]$, i.e., posing the following relations:

$$(\ldots,\Psi_j^{D},\ldots)\overset{z=D}{=}(\ldots,\Psi_i^{D,+},\ldots)[\mathbf{I}-(\mathbf{A}^{+})(\mathbf{A}^{+})^{\#}]+(\ldots,\Psi_i^{D,-},\ldots)\mathbf{D}^{-+}\times[\mathbf{I}-(\mathbf{A}^{+})(\mathbf{A}^{+})^{\#}]\overset{z\geqslant D}{=}(\ldots,\Psi_j^{\mathrm{III},+},\ldots)\mathbf{O}. \tag{21}$$

The matrix $\mathbf{D}^{-+}$ ensures the continuity of the wave function at the reflection point $z=D$. It must be chosen so that

$$(\ldots,\Psi_i^{D,+},\ldots)+(\ldots,\Psi_i^{D,-},\ldots)\mathbf{D}^{-+}\overset{z=D}{=}0. \tag{22}$$

For example, if the states $\Psi_i^{D,\pm}$ are given by $\Psi_i^{D,\pm}=e^{\pm ik_{z,i}z}$, $\mathbf{D}^{-+}$ is defined by $D_{i,j}^{-+}=-e^{i2k_{z,i}D}\delta_{i,j}$.

In the same way, we can define the set of solutions

$$(\ldots,\Psi_j^{\mathrm{III}},\ldots)\overset{z=D}{=}(\ldots,\Psi_j^{D,-},\ldots)\mathbf{O}\overset{z\geqslant D}{=}(\ldots,\Psi_i^{\mathrm{III},-},\ldots)[\mathbf{I}-(\mathbf{A}^{-})(\mathbf{A}^{-})^{\#}]+(\ldots,\Psi_i^{\mathrm{III},+},\ldots)\times\mathbf{D}^{+-}[\mathbf{I}-(\mathbf{A}^{-})(\mathbf{A}^{-})^{\#}], \tag{23}$$

where the matrix $\mathbf{D}^{+-}$ is defined by

$$(\ldots,\Psi_i^{\mathrm{III},-},\ldots)+(\ldots,\Psi_i^{\mathrm{III},+},\ldots)\mathbf{D}^{+-}\overset{z=D}{=}0. \tag{24}$$

This set of solutions is trivially verified due to the relation (16).

By adding the solutions (21) to (19) and (23) to (20), one finds two sets of solutions that are in the appropriate form:

$$(\ldots,\Psi_j^+,\ldots)\overset{z=D}{=}(\ldots,\Psi_i^{D,+},\ldots)\mathbf{I}+(\ldots,\Psi_i^{D,-},\ldots)\mathbf{t}_{D,\mathrm{III}}^{-+}\overset{z\geqslant D}{=}(\ldots,\Psi_j^{\mathrm{III},+},\ldots)\mathbf{t}_{D,\mathrm{III}}^{++},\quad(25)$$

$$(\ldots,\Psi_j^-,\ldots)\overset{z=D}{=}(\ldots,\Psi_j^{D,-},\ldots)\mathbf{t}_{D,\mathrm{III}}^{--}\overset{z\geqslant D}{=}(\ldots,\Psi_i^{\mathrm{III},-},\ldots)\mathbf{I}+(\ldots,\Psi_i^{\mathrm{III},+},\ldots)\mathbf{t}_{D,\mathrm{III}}^{+-}.\quad(26)$$

In conclusion, we can write the relations (5)–(8) in the more general form

$$\mathbf{t}_{D,\mathrm{III}}^{++}=(\mathbf{A}^+)^{\#},\quad(27)$$

$$\mathbf{t}_{D,\mathrm{III}}^{-+}=(\mathbf{B}^+)(\mathbf{A}^+)^{\#}+\mathbf{D}^{-+}[\mathbf{I}-(\mathbf{A}^+)(\mathbf{A}^+)^{\#}],\quad(28)$$

$$\mathbf{t}_{D,\mathrm{III}}^{--}=(\mathbf{A}^-)^{\#},\quad(29)$$

$$\mathbf{t}_{D,\mathrm{III}}^{+-}=(\mathbf{B}^-)(\mathbf{A}^-)^{\#}+\mathbf{D}^{+-}[\mathbf{I}-(\mathbf{A}^-)(\mathbf{A}^-)^{\#}],\quad(30)$$

where $\mathbf{D}^{-+}$ and $\mathbf{D}^{+-}$ are two matrices that ensure the wave function continuity at $z=D$ and the generalized inverse $(\mathbf{A}^{\pm})^{\#}$ of a given matrix $(\mathbf{A}^{\pm})$ is computed by

$$(\mathbf{A}^{\pm})^{\#}=[(\mathbf{A}^{\pm})^{\dagger}(\mathbf{A}^{\pm})]^{-1}(\mathbf{A}^{\pm})^{\dagger},\quad(31)$$

when $(\mathbf{A}^{\pm})$ has more rows than columns or

$$(\mathbf{A}^{\pm})^{\#}=(\mathbf{A}^{\pm})^{\dagger}[(\mathbf{A}^{\pm})(\mathbf{A}^{\pm})^{\dagger}]^{-1}\quad(32)$$

in the other case.

D. Final solution by a transfer-matrix combination

In the final step, the four matrices $\mathbf{t}_{\mathrm{I},D}^{\pm,\pm}$ have to be combined with the four matrices $\mathbf{t}_{D,\mathrm{III}}^{\pm,\pm}$ in order to obtain the transfer matrices that contain the coefficients of the solutions

$$\Psi_j^{+}\overset{z\leqslant 0}{=}\Psi_j^{\mathrm{I},+}+\sum_i(t_{\mathrm{I,III}}^{-+})_{i,j}\Psi_i^{\mathrm{I},-}\overset{z\geqslant D}{=}\sum_i(t_{\mathrm{I,III}}^{++})_{i,j}\Psi_i^{\mathrm{III},+},\quad(33)$$

$$\Psi_j^{-}\overset{z\leqslant 0}{=}\sum_i(t_{\mathrm{I,III}}^{--})_{i,j}\Psi_i^{\mathrm{I},-}\overset{z\geqslant D}{=}\Psi_j^{\mathrm{III},-}+\sum_i(t_{\mathrm{I,III}}^{+-})_{i,j}\Psi_i^{\mathrm{III},+}\quad(34)$$

corresponding to a single incoming state in region I or region III. These four transfer matrices $\mathbf{t}_{\mathrm{I,III}}^{\pm\pm}$ are obtained by using formulas developed by Pendry [13,14] in dynamic low energy electron diffraction simulations:

$$\mathbf{t}_{\mathrm{I,III}}^{++}=\mathbf{t}_{D,\mathrm{III}}^{++}[\mathbf{I}-\mathbf{t}_{\mathrm{I},D}^{+-}\mathbf{t}_{D,\mathrm{III}}^{-+}]^{-1}\mathbf{t}_{\mathrm{I},D}^{++},\quad(35)$$

$$\mathbf{t}_{\mathrm{I,III}}^{-+}=\mathbf{t}_{\mathrm{I},D}^{-+}+\mathbf{t}_{\mathrm{I},D}^{--}\mathbf{t}_{D,\mathrm{III}}^{-+}[\mathbf{I}-\mathbf{t}_{\mathrm{I},D}^{+-}\mathbf{t}_{D,\mathrm{III}}^{-+}]^{-1}\mathbf{t}_{\mathrm{I},D}^{++},\quad(36)$$

$$\mathbf{t}_{\mathrm{I,III}}^{--}=\mathbf{t}_{\mathrm{I},D}^{--}[\mathbf{I}-\mathbf{t}_{D,\mathrm{III}}^{-+}\mathbf{t}_{\mathrm{I},D}^{+-}]^{-1}\mathbf{t}_{D,\mathrm{III}}^{--},\quad(37)$$

$$\mathbf{t}_{\mathrm{I,III}}^{+-}=\mathbf{t}_{D,\mathrm{III}}^{+-}+\mathbf{t}_{D,\mathrm{III}}^{++}\mathbf{t}_{\mathrm{I},D}^{+-}[\mathbf{I}-\mathbf{t}_{D,\mathrm{III}}^{-+}\mathbf{t}_{\mathrm{I},D}^{+-}]^{-1}\mathbf{t}_{D,\mathrm{III}}^{--}.\quad(38)$$

III. APPLICATION

A. Preliminaries

In order to illustrate this theory, let us consider electronic field emission from a metallic tip and the diffraction of the extracted beam by a three-dimensional circular aperture facing the emitter. The extraction field results from the application of a potential bias V established between the support of the tip and a conducting grid at a distance D, which supports the circular aperture. It is convenient to assume the axial direction z to be an n-fold symmetry axis and to use polar coordinates in the plane normal to the symmetry axis (i.e., ϕ for the azimuthal angle and ρ for the radial distance to the axis).

Region I (i.e., the metallic support of the tip) is assumed to be a Sommerfeld metal, delimited by the plane $z=0$ and characterized by empirical values of W (work function) and E_F (Fermi energy). The potential energy in region III ($D\leqslant z\leqslant D+D'$, with D' the length of the circular aperture) is set conventionally to 0 inside the aperture ($\rho\leqslant R'$) and ∞ outside ($\rho>R'$). The potential energy in region IV (i.e., the open space beyond the conducting grid $z>D+D'$) takes the same constant value 0. The value of the potential energy in region I is then $V_{met}=eV-W-E_F$. With these assumptions, region II is the only diffusive part of the problem and, the Schrödinger equation being linear, the transfer-matrix methodology can be applied.

B. Wave function expansion in regions I and III

The wave function is expanded along basis functions ψ that contain the ϕ and ρ dependences. The set of these functions is forced to be enumerable, by specifying that the scattering electron remains localized inside a cylinder with radius R [6] in regions I and II. The fact that R differs from R' in region III is responsible for the non-square shape of the transfer matrices.

The basis states $\Psi^{\mathrm{I},\pm}$, $\Psi^{D,\pm}$ and $\Psi^{\mathrm{III},\pm}$ introduced in the previous section to describe the wave function in regions I, and III then take the specific forms

$$\Psi_{(m,j)}^{\mathrm{I},\pm}=e^{\pm i\sqrt{(2m/\hbar^2)(E-V_{met})-k_{m,j}^2}z}\psi_{(m,j)}(\rho,\phi),\quad(39)$$

$$\Psi_{(m,j)}^{D,\pm}=e^{\pm i\sqrt{(2m/\hbar^2)E-k_{m,j}^2}z}\psi_{(m,j)}(\rho,\phi),\quad(40)$$

$$\Psi_{(m,j)}^{\mathrm{III},\pm}=e^{\pm i\sqrt{(2m/\hbar^2)E-k_{m,j}'^2}z}\psi'_{(m,j)}(\rho,\phi),\quad(41)$$

with

$$\psi_{(m,j)}(\rho,\phi)=\frac{J_m(k_{m,j}\rho)e^{im\phi}}{\sqrt{2\pi\int_0^R\rho[J_m(k_{m,j}\rho)]^2d\rho}},\quad(42)$$

$$\psi'_{(m,j)}(\rho,\phi)=\frac{J_m(k'_{m,j}\rho)e^{im\phi}}{\sqrt{2\pi\int_0^{R'}\rho[J_m(k'_{m,j}\rho)]^2d\rho}}.\quad(43)$$

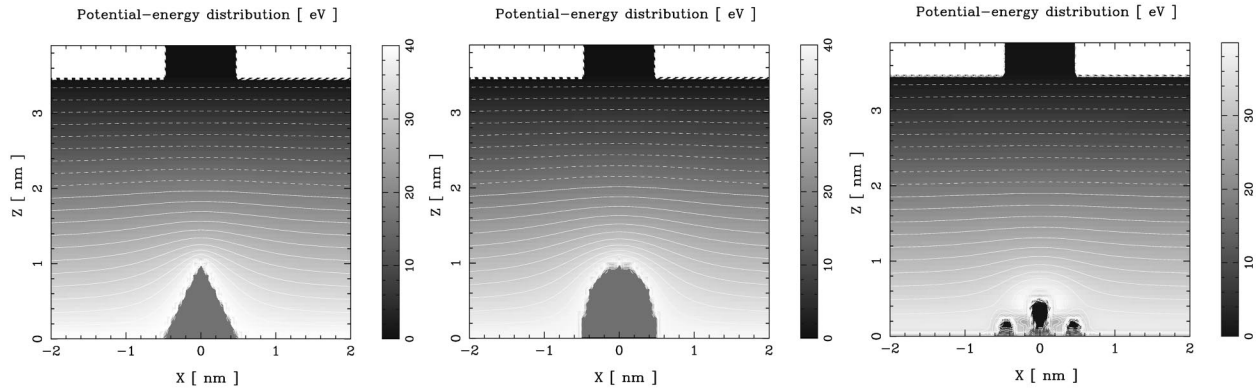

FIG. 2. Real part of the potential-energy distribution (in eV) in the xz plane. A 40 V bias is applied over the 3.5 nm separation between the metal surface and the conducting grid. This grid supports a circular aperture with a radius of 0.5 nm and length of 0.4 nm. The tip is represented by a conical emitter with a height and base radius of 1 and 0.5 nm (left), an elliptical emitter whose main axes in the z and ρ directions are 1 and 0.5 nm (center), and an atomic protrusion (right).

All functions involved in these expressions have a pair of subscripts (m,j). The radial wave vectors $k_{m,j}$ in regions I and II, where the wave function is mainly confined around the central z axis, are solutions of $J_m(k_{m,j}R)=0$, while the radial wave vectors $k'_{m,j}$ in region III, where the wave function can take significant values at $\rho=R'$, are solutions of $J'_m(k'_{m,j}R')=0$.

C. Propagation through regions II and IV

To propagate the solutions $\overline{\Psi}^{\pm}_{(m,j)}$ through region II, we use the following expression:

$$\overline{\Psi}^{\pm}_{(m,j)}=\sum_{m,j}\Phi_{(m,j)}(z)\psi_{(m,j)}(\rho,\phi), \tag{44}$$

where the z dependence is contained in the coefficients $\Phi_{(m,j)}(z)$ of the expansion. The techniques given in Refs. [10,15,16] to deal with the stationary three-dimensional Schrödinger equation are used to propagate these coefficients between $z=0$ and $z=D$.

The propagation to a distant screen in region IV is achieved within the Green's function formalism (see Ref. [17]) and the Neumann boundary conditions [18], which have as formal consequence the cancelation the z component of the current density on the aperture border at $z=D+D'$. A proper renormalization of the Green's functions is achieved to enforce current conservation in region IV.

D. Characterization of the physical system

Let us consider an electric bias of 40 V and a metal-grid distance of 3.5 nm. The bulk of the metal is a jellium characterized by a Fermi energy of 19.1 eV and a work function of 4.5 eV (values for tungsten). The emitter is represented by a conical or elliptical extension of the supporting jellium, with various dimensions. A representation of the tip by an atomic protrusion is also considered.

In the two first representations, the tip is a conical or elliptical metallic medium characterized by an infinite dielectric constant and the same Fermi energy and work function as in the bulk. In the case of a conical emitter, the height and base radius take, respectively, the three following pairs of values: 0.6 and 0.6 nm, 1 and 0.5 nm, 1.5 and 0.75 nm. In the case of an elliptical emitter, the main axes in the z and ρ directions take respectively the same pairs of values. The corresponding potential-energy distributions in region II are computed by the relaxation techniques given in Ref. [15].

In the last representation, the atomic protrusion consists of four atomic layers whose distance is the 0.091 nm separation encountered between two adjacent atomic layers in the bulk of the metal in the $\langle 111\rangle$ direction. The first atomic layer is made of twelve atoms at 0.013 nm from the metallic surface, the second layer of seven atoms is at $z=0.104$ nm, the third of three atoms at 0.195 nm and the last atom is moved from its crystallographic position to $z=0.289$ nm (see Ref. [19]). Each atom is represented by a dipole. Following references [15,20], the polarizability of each dipole is an average, weighted by the number of neighboring atoms, of the polarizability of a neutral isolated atom [21] $\alpha_0/(4\pi\epsilon_0)=7$ Å^3 and the polarizability in the bulk [22] at the Fermi frequency $\alpha_{bulk}/(4\pi\epsilon_0)=1.49$ Å^3. The corresponding potential-energy distribution in region II is computed by using techniques given in Ref. [16]. The result is similar to that obtained with an elliptical tip, whose main axes are equal to 0.6 nm.

With a purely real potential, the electrons that fall beside the circular aperture do not disappear from the simulation, as they would disappear from consideration in reality. They are reflected several times between the aperture border and the metal surface, to be finally reflected by the cylinder boundary and come back to the aperture, where they interfere with the original beam. To cope with this situation, the potential energy is given a negative imaginary component, in a part of the system that is not met by the incident beam so only the electrons that are reflected by the aperture border are absorbed. This imaginary component is zero below $\rho=2.5$ nm, grows as a square function between $\rho=2.5$ nm and 3.5 nm from 0 to -1 eV, and keeps this last value for larger values of ρ.

A vertical section of the real part of the potential-energy distribution corresponding to the various representations of the tip is given in Fig. 2. The aperture in region III has a radius of 0.5 nm and a length of 0.4 nm.

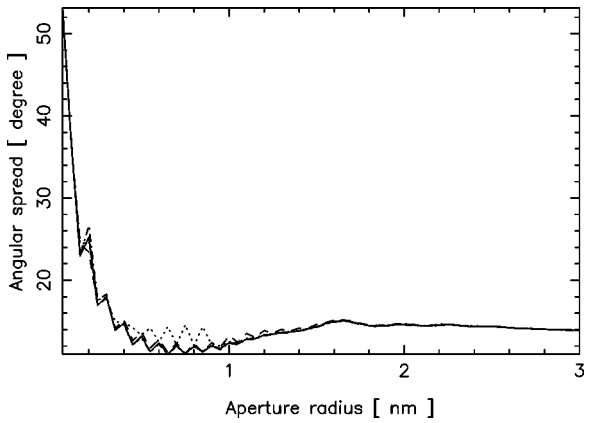

FIG. 3. Angular spread of the diffracted electronic beam as a function of the aperture radius. The length of the aperture is 0.4 nm. The basis states are defined by $\Delta E=0$ and $R=6$ nm (dotted), $\Delta E=2$ eV and $R=6$ nm (full line), $\Delta E=4$ eV and $R=6$ nm (dashed), and $\Delta E=2$ eV and $R=12$ nm (dot-dashed).

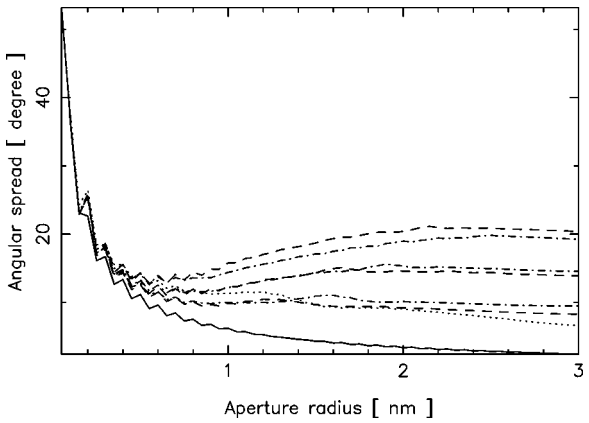

FIG. 4. Angular spread of the diffracted electronic beam as a function of the aperture radius. The full curve corresponds to a single incident plane wave, the dashed curves to a conical tip, the dot-dashed curves to an elliptical tip, and the dotted curve to an atomic structure. The height and base radius corresponding to the three curves associated with the conical/elliptical emitter are respectively 0.6 and 0.6 nm, 1 and 0.5 nm, 1.5 and 0.75 nm (upward).

E. Angular spread as a function of the aperture radius

The present application aims at studying the angular spread $\Delta\theta$ of the diffracted electronic beam as a function of the aperture radius R', for different shapes and dimensions of the emitter. The total current density in region IV results from the current densities associated with each solution Ψ^+. These contributions are weighted according to the corresponding density of states in the metal (see Ref. [17]). For a given total current density distribution, the angular spread $\Delta\theta$ is computed according to

$$\Delta\theta \underset{r\to\infty}{=} \frac{\int_0^{2\pi} d\phi \int_0^{\pi/2} d\theta |\theta| J_r(r,\theta,\phi)}{\int_0^{2\pi} d\phi \int_0^{\pi/2} d\theta J_r(r,\theta,\phi)}. \tag{45}$$

To first control the stability of the results toward the representation of the wave function, we used four different base for computating the dependence of $\Delta\theta$ on the aperture radius R'. The results are gathered in Fig. 3. The tip is described in these simulations by a cone with a height and base diameter of 1 nm (see left part of Fig. 2). The width D' of the circular aperture is 0.4 nm. The simulations are performed by considering m values ranging from -3 to 3. The basis states are limited by the condition $k_{m,j}, k'_{m,j} \leqslant \sqrt{(2m/\hbar^2)(E+\Delta E)}$, where ΔE is 0 eV in the first representation, 2 eV in the second, and 4 eV in the third. The confinement radius R in regions I and II is 6 nm in these three representations. In the last representation, ΔE and R are respectively 2 eV and 12 nm.

The four results are in good agreement, except for the first representation (the dotted line) corresponding to $\Delta E=0$. This is due to the fact that the evanescent states introduced in the circular aperture and in the nearby region by taking $\Delta E>0$ keep significant values over distances of the order of D' and improve the description of the lateral behavior of the wave function. The (dashed) curve corresponding to $\Delta E=4$ eV presents some discrepancies compared to those obtained with $\Delta E=2$ eV, due to the instabilities inherent in the consideration of evanescent states. In the remaining part of this paper, values of $\Delta E=2$ eV and $R=6$ nm are used.

With this choice of representation, we then compared the results obtained by considering conical or elliptical tips, with the previously given dimensions, and the atomic structure. These results are gathered in Fig. 4. The full curve corresponds to a single incident plane wave in the circular aperture, propagating in the z direction. The dashed curves correspond to the conical tips, the dot-dashed curves to the elliptical tips, and the dotted curve to the atomic structure. The three curves associated with either a conical or elliptical tip correspond (upward) to increasing tip dimensions. The curves consist typically of three parts: (i) a first decreasing part associated with Fraunhofer diffraction; (ii) a second growing part associated with Fresnel diffraction (for tips larger than 1 nm); (iii) a third constant part related to the finite size of the beam (for all tips).

As expected for Fraunhofer diffraction, the first part of the figure decreases with increasing aperture radius R'. All curves fit perfectly below $R'=0.2$ nm (the wavelength λ in the aperture being equal to 0.206 nm), since details of the emitter structure are no longer relevant down to this reduced length scale.

The second part of Fig. 4 is dominated by Fresnel diffraction. Fresnel diffraction is encountered when the spherical shape of the incident wave prevails over the dimensions of the diffractive structure. The diffraction pattern is then correlated with the spatial distribution of this structure, which explains the increase of the angular spread with the aperture radius.

The transition between Fraunhofer and Fresnel diffraction occurs when the distance d between the virtual projection point and the sample is given by $d=a^2/\lambda$, where $a=2R'$ is the typical dimension of the diffractive structure and λ the relevant electronic wavelength. By considering this transition to occur for a radius $R'=0.55$ nm in the case of a 1.5 nm long tip, we find the virtual projection point to lie at $z=-2.3$ nm. In the case of a 1 nm long tip, the transition occurs for a radius $R'=0.75$ nm and the corresponding vir-

tual projection point lies at $z=-7.4$ nm. These results are of the same order as experimental values [9] and show how the position of the virtual projection point is sensitive to the potential-energy distribution surrounding the emitter.

In the last part of Fig. 4, the aperture dimensions tend to exceed the size of the electronic beam when incident on the extraction grid, so the angular spread tends to the value that would be encountered without any aperture. This value is obtained for $R'=2$ nm, which corresponds to the size of the electronic beam in the plane $z=D$.

The angular spread grows with the height of the tip. This is consistent with the initial angular spread increasing as the top of the emitter becomes sharper [23], but is mainly due to the fact that the distance at which the electrons encounter a nonradial electric field (tending to reduce the angular spread) decreases as the length of the tip increases.

The results are rather insensitive to the conical or elliptical shape of the tip. In these two model situations, the tip has an axial symmetry so the sphericity of the electronic wave is not altered by atomic corrugation. The large values of the virtual projection point distance and the electronic wavelength (0.28 nm) in the tip, which forces the incident electronic states to be evanescent when encountering the top of the conical tip, explain also why the results obtained with both tip shapes are comparable.

The result obtained with the four-layer atomic structure is close to those obtained with the conical/elliptical tip with 0.6 nm length. This is because the surrounding potential-energy distributions are similar. These results show the necessity (in a context where the metallic support is limited by a plane) of improving the description of the atomic structure of the emitter to account for the occurrence of Fresnel diffraction with nanotips. This could be achieved by considering more than four atomic layers or by describing the metallic support by a hemispherical boundary.

Finally, the oscillations of all curves arise approximately at the same values of the aperture radius. They are related to the long-range behavior of the wave function at $\theta=\pi/2$. Their magnitude tends to increase (essentially in Fraunhofer conditions) and the curves tend to grow (essentially in Fresnel conditions) with the length D' of the aperture. This behavior, illustrated in Fig. 5, is related to the wave function taking larger values at the border of the aperture (compared with those taken at the center) as the length of the aperture increases.

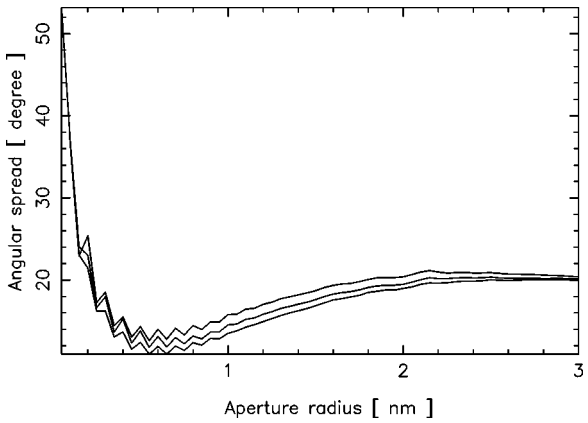

FIG. 5. Angular spread of the diffracted electronic beam as a function of the aperture radius. The curves correspond (upward) to an aperture width of 1, 2, and 4 nm. The height and base radius of the conical emitter are respectively 1.5 and 0.75 nm.

IV. CONCLUSION

An extension of the transfer-matrix methodology that allows the consideration of nonsquare matrices was presented. This extension is useful in deriving solutions whose representation in the two regions surrounding the diffusive part of the system requires different numbers of basis states.

For pedagogical reasons, the method was used to compute the angular spread of a field-emitted electronic beam, after it is diffracted by a three-dimensional circular aperture. The scattering computation relies on the three-dimensional Schrödinger equation and requires techniques adapted to nonsquare matrices, since the confinement in the aperture causes the number of basis states in this region to differ from that used in the preceding regions.

The results are well explained in terms of Fraunhofer and Fresnel diffraction. The radius where the transition between these two types of diffraction occurs gives the position of the virtual projection point. This position turns out to be very sensitive to the tip dimensions and to the orientation of the surrounding electric field. In the conditions of this paper (40 V bias and flat metallic support), Fresnel diffraction was not encountered with tips smaller than 1 nm.

Typical future applications of nonsquare transfer matrices will be the study of connections between optical fibers, carbon nanotubes, or other systems with different sections or physical properties requiring different numbers of basis states for the expansion of the solution. Non-square transfer matrices are also suited for the study of electronic circuits with several input or output lines.

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.

[1] P.St.J. Russel, T.A. Birks, and F.D. Lloyds-Lucas, in *Confined Electrons and Photons*, edited by E. Burstein and C. Weisbuch (Plenum, New York, 1995).

[2] W.D. Sheng and J.B. Xia, J. Phys.: Condens. Matter **8**, 3635 (1994).

[3] J.-P. Vigneron, I. Derycke, T. Laloyaux, P. Lambin, and A.A. Lucas, Scanning Microsc. Suppl. **7**, 261 (1993).

[4] A.J. Ward and J.B. Pendry, J. Mod. Opt. **44**, 1703 (1997).

[5] T.M. Kalotas and A.R. Lee, Am. J. Phys. **59**, 48 (1991).

[6] T. Laloyaux, A.A. Lucas, J.-P. Vigneron, P. Lambin, and Morawitz, J. Microsc. **152**, 53 (1988).

[7] T. Laloyaux, I. Derycke, J.-P. Vigneron, and A.A. Lucas Phys.

Rev. B **47**, 7508 (1993).
[8] G. Binning and H. Rohrer, Helv. Phys. Acta **55**, 726 (1982).
[9] V.T. Binh, V. Semet, and N. Garcia, Ultramicroscopy **58**, 307 (1995).
[10] A. Mayer and J.-P. Vigneron, Phys. Rev. E **59**, 4659 (1999).
[11] A. Mayer and J.-P. Vigneron, Phys. Rev. E **60**, 7533 (1999).
[12] A. BenIsrael and T.N.E. Greville, *Generalized Inverses. Theory and Applications.* (Robert E. Krieger Publishing Company, New York, 1980).
[13] J.B. Pendry, J. Mod. Opt. **41**, 209 (1994).
[14] J.B. Pendry and A. MacKinnon, Phys. Rev. Lett. **69**, 2272 (1992).
[15] A. Mayer and J.-P. Vigneron, Phys. Rev. B **56**, 12 599 (1997).
[16] A. Mayer and J.-P. Vigneron, J. Phys.: Condens. Matter **10**, 869 (1998).
[17] A. Mayer and J.-P. Vigneron, Phys. Rev. B **60**, 2875 (1999).
[18] J.D. Jackson *Classical Electrodynamics*, 2nd ed. (Wiley, New York, 1962), p. 427.
[19] A. Mayer and J.-P. Vigneron, Ultramicroscopy **79**, 35 (1999).
[20] M. Devel, C. Girard, and C. Joachim, Phys. Rev. B **53**, 13 159 (1996).
[21] R. Rodney and T.P. Russel, At. Data **3**, 195 (1971).
[22] E.D. Palik, *Handbook of Optical Constants of Solids* (Academic, London, 1985).
[23] H. De Raedt and K. Michielsen, *Nanosources and Manipulation of Atoms under High Fields and Temperatures: Applications* (Kluwer Academic, Dordrecht, 1993), p. 45.