Optimization methods and their use in low-energy electron-diffraction calculations

H. Over
Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-1000 Berlin 33, Germany

U. Ketterl
Sektion Numerik, Fachbereich Mathematik, Technische Universität Berlin, D-1000 Berlin 12, Germany

W. Moritz
Institut für Kristallographie und Mineralogie, Universität München, D-8000 München 2, Germany

G. Ertl
Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, D-1000 Berlin 33, Germany
(Received 22 May 1992; revised manuscript received 4 August 1992)

The speed of automatic optimization procedures used in surface structure determination by low-energy electron diffraction can be greatly enhanced by the use of linear approximations in the calculation of scattering amplitudes. It is shown how linear approximations can be used in the calculation of derivatives of intensities which are required in the least-squares optimization method. The derivatives with respect to structural and nonstructural parameters are calculated applying a combination of analytic and numerical methods in connection with approximations of the sum over lattice points in the angular momentum representation. Special cases for different structural and nonstructural parameters and simplifications for special geometries are discussed. The computational effort becomes nearly independent of the number of free parameters and enables the analysis of complex surface structures.

I. INTRODUCTION

The application of optimization techniques and fast calculational procedures is essential for the future development of low-energy electron diffraction (LEED). The simultaneous optimization of a large number of parameters makes much more complex structures accessible to an analysis, in contrast to the usual method (grid search) of approaching the R-factor minimum by independent variation of all parameters. In principle, various kinds of optimization schemes are applicable in LEED, but until recently only a few different approaches have been proposed. These include a variant of the method of steepest descent,[1] a combination of tensor LEED (Ref. 2) with the gradient method,[3] so-called direct methods,[4] and least-squares optimization procedures.[5–7] In general, those techniques which apply the derivatives of a fitting function are faster than search procedures. The calculation of derivatives, however, can become rather time consuming because an analytic solution is not, in general, applicable. If one avoids the calculation of derivatives by applying other optimization algorithms this may simplify the computer program but does not improve the speed of the calculation because in any case the calculation of LEED intensities is necessary for several points in parameter space to determine the gradient of the R-factor function.

A very efficient optimization procedure is the so-called expansion method which is routinely applied in many fields, e.g., in x-ray structure determination. This method has the advantage of converging rapidly near a minimum of the fitting function in the parameter space, but, on the other hand, converges only slowly far from a minimum. Therefore we use a slightly different method proposed by Marquardt[8] which combines the advantages of the gradient method far from the optimum and the expansion method near the optimum. Its use in connection with LEED has been described previously[5–7] and it has been applied successfully in a number of structural analyses.[9–11]

The optimization procedure requires the knowledge of the derivatives of a fitting function with respect to the parameters which have to be optimized. As a fitting function frequently an R factor is used which is a complicated function of experimental and theoretical I/V curves and involves integrations over the energy range. The derivatives of the R factor, therefore, cannot be analytically calculated. In the least-squares optimization method (the Marquardt method is a variant of that method) the mean-square deviation between experimental and theoretical intensities is taken as the fitting function. Its derivatives are simply calculated from the derivatives of the intensities. Therefore only these are required. The purpose of this paper is to present methods to calculate derivatives of intensities with respect to structural and nonstructural parameters with high efficiency. The intensity derivatives are intended here to be used with the least-squares optimization method but could be used with other optimization techniques as well.

The optimization algorithm does not require an analytic solution for the derivatives of the intensities and the simplest way is to calculate the derivatives numerically as has been done in previous calculations.[5,6] In this case the

computational effort scales linearly with the number of free parameters. With an increasing number of free parameters the calculation of derivatives soon becomes the most time-consuming step in the analysis. It is, therefore, highly desirable to improve the efficiency in calculating derivatives. One approach, proposed by Rous *et al.*,[2] is the so-called tensor LEED method. This approach has been published for the use with the renormalized forward scattering (RFS) method.[12] The combination of the tensor LEED approach with direct methods and with optimization procedures[13–15] has been successfully applied for several structures leading to an efficient reduction of the computational effort. The calculation of derivatives described here differs from the tensor LEED method in that multiple-scattering effects are fully included and further that nonstructural parameters are also treated. The nonstructural parameters considered here are the Debye temperature and an occupation factor to be used within the average t-matrix approximation (ATA). Another difference to the tensor LEED method is that only the first derivative is calculated and higher-order effects, which are included in the second step of tensor LEED, are not considered. We use here the layer-doubling scheme for interlayer multiple scattering but the calculation can be easily extended for the RFS scheme.

The paper is organized as follows. In Sec. II A we describe the general scheme for the calculation of derivatives for structural parameters in the multiple-scattering formalism. In Sec. II B, it is shown how the calculation of the derivatives of the layer-scattering amplitudes is carried out. It is also discussed how further simplification brought about by approximations in performing the lattice sum can be used. In Sec. II C we describe the implementation in the layer-doubling scheme. The treatment of nonstructural parameters is described in Sec. II D. Simplifications for special cases are discussed in Sec. II E. We compare the different methods and discuss the results in the last section.

II. DESCRIPTION OF THE APPROXIMATION SCHEME

A. Calculation of derivatives of multiple-scattering amplitudes

In the least-squares optimization scheme, the minimum of a fitting function in the parameter space is determined from its derivatives using a linear expansion of the fitting function. Nonlinearities are overcome by iterating the process. As a fitting function usually the mean-square deviation between measured and calculated intensities is used, and the calculation of its derivations requires the calculation of the derivatives of the intensities with respect to the variable parameters. The least-squares optimization procedure is, of course, not limited to the use of this fitting function and the Y function, as defined for the r factor proposed by Pendry,[16] could be taken as well. We will describe in the following the theory to calculate the derivatives of the intensities which are required in any case.

The simplest approach is to expand the intensities in a Taylor series around a reference structure and to calculate the derivative numerically. It is best to use the derivative of the amplitudes

$$\frac{\partial A(\mathbf{p})}{\partial p_j} \approx \frac{A(\mathbf{p}+\delta\mathbf{p}_j)-A(\mathbf{p})}{\delta p_j} \tag{1}$$

and to calculate the intensities for a new set of parameters $\mathbf{p}+\delta\mathbf{p}$ by

$$I(\mathbf{p}+\delta\mathbf{p}_j)=|A(\mathbf{p})|^2+\frac{\partial A(\mathbf{p})}{\partial p_j}A^*(\mathbf{p})\delta\mathbf{p}_j + \frac{\partial A^*(\mathbf{p})}{\partial p_j}A(\mathbf{p})\delta\mathbf{p}_j+\left|\frac{\partial A(\mathbf{p})}{\partial p_j}\right|^2\delta p_j^2 \tag{2}$$

which slightly improves the quality of the linear expansion. If the changes of the parameters are larger than a certain limit, the full dynamical calculation has to be repeated in the iteration process. The vector $\mathbf{p}=(p_1,p_2,\ldots,p_N)$ denotes the set of structural or nonstructural parameters to be optimized.

The numerical calculation requires the solution of the multiple-scattering formalism for each parameter increment, and hence the computational effort increases approximately linearly with the number of free parameters. Improvements can be made by deriving analytic expressions and separating those parts which are common to different parameters so that they have to be calculated only once.

A complete analytic expression for the derivatives is not useful because its calculation would not reduce the computational effort. We therefore first separate the matrix inversion occurring in the self-consistent solution of the multiple-scattering problem. This occurs in different formulations of the multiple-scattering theory either in direct space with a spherical-wave expansion or in reciprocal space with a plane-wave expansion. In both formulations, the size of the system of linear equations can become very large for complex structures. The dimensions of the matrices are governed by the number of spherical waves for calculating the layer-scattering matrices and by the number of plane waves in the layer-doubling scheme. Although a fast computation is possible using array processors, the calculation of the matrix elements and the inversion of matrices are the most time-consuming processes. Simplifications can be introduced in two different steps. At first, the derivative can be calculated by expanding the inverse of the matrix $([1-S(\mathbf{p})])$ in a Taylor series around a reference structure given by the set of parameters $\mathbf{p}_0$,

$$(1-S(\mathbf{p}_0+\delta\mathbf{p}))^{-1}=(1-S(\mathbf{p}_0))^{-1} +(1-S(\mathbf{p}_0))^{-1} \times[S(\mathbf{p}_0+\delta\mathbf{p})-S(\mathbf{p}_0)] \times(1-S(\mathbf{p}_0))^{-1}+\cdots . \tag{3}$$

The inverse matrix $(1-S(\mathbf{p}_0))^{-1}$ or its solution vectors, respectively, can be stored in order to use it for derivatives of all parameters. The main effect of using Eq. (3) is that the matrix inversion is replaced by a matrix

multiplication and furthermore that the calculation of the increment $S(\mathbf{p}_0+\delta\mathbf{p})-S(\mathbf{p}_0)$ can be simplified in angular-momentum space as well as in reciprocal space as will be outlined below.

The numerical calculation of matrix elements $[S(\mathbf{p}_0+\delta\mathbf{p})-S(\mathbf{p}_0)]$ for small increments $\delta\mathbf{p}$ occurs in both the plane-wave and spherical-wave representations, which are both used in the multiple-scattering scheme. At first, reflection and transmission matrices are calculated for monatomic or composite layers. In case the layer spacings are small, several subplanes have to be treated as a composite layer. This step of the calculation is done in a spherical-wave expansion and the atomic positions are required in polar coordinates. The layer-scattering matrices are then combined by the layer-doubling scheme or the RFS method.[17] This part is done in a plane-wave expansion. Both parts will be discussed separately.

B. Derivatives of layer scattering matrices of composite layers

The variation of the atomic positions in a composite layer leads to a variation of its reflection and transmission matrices. The reflection and transmission matrices for a composite layer are given as follows by[18]

$$M_{g'g}=\sum_{lm\nu}\sum_{l'm'\mu}\frac{8\pi^2 i}{|\mathbf{k}|\,A\,|\mathbf{k}_{g'z}|}\{i^{l}(-1)^{m}Y_{l,m}(\Omega(\mathbf{k}_{g'}))e^{i\mathbf{k}_{g'}\mathbf{d}_\nu}\times[1-X]^{-1}_{lm\nu,l'm'\mu}t_l^{(\mu)}i^{-l'}\times Y_{l',-m'}(\Omega(\mathbf{k}_g))e^{-i\mathbf{k}_g\mathbf{d}_\mu}\}\ , \tag{4}$$

where

$$X_{lml'm'}=\sum_{l''m''}c(lm,l'-m',l''m'')F_{l''m''}\ , \tag{5}$$

$$c(lm,l'-m',l''m'')=\int Y_{lm}(\Omega)Y_{l'-m'}(\Omega)Y_{l''m''}(\Omega)d\Omega\ , \tag{6}$$

$$F_{l''m''}(\mathbf{k}_0)=\sum_{\mathbf{P}}e^{i\mathbf{k}_0\cdot\mathbf{P}}i^{l''}h_{l''}(|\mathbf{k}_0|\cdot|\mathbf{P}+\mathbf{d}_\mu-\mathbf{d}_\nu|)\times Y_{l''-m''}(\Omega(\mathbf{P}+\mathbf{d}_\mu-\mathbf{d}_\nu))=\sum_{\mathbf{P}}f^{\nu\mu}_{l''m''}(\mathbf{P}+\mathbf{d}_\mu-\mathbf{d}_\nu)\ , \tag{7}$$

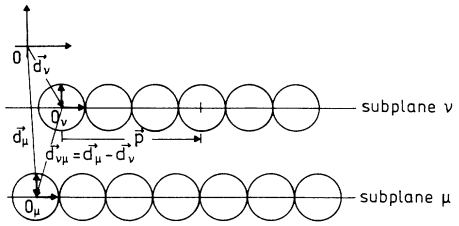

FIG. 1. Illustration of atomic positions in a composite layer as it is used in Eqs. (4)–(7) for calculating the scattering matrices.

$h_l(z)$ are the spherical Hankel functions of the first kind, $\mathbf{d}_\mu-\mathbf{d}_\nu$ is the translation vector between origins of the subplanes μ and ν, $\mathbf{P}$ are the lattice vectors within a subplane, and $\mathbf{k}_0$ is the wave vector of the incident wave field, see Fig. 1.

The matrix elements $M_{g'g}$ are reflection and transmission coefficients describing the amplitude of the wave diffracted from the direction $\mathbf{k}_g$ into the direction $\mathbf{k}_{g'}$. Most of the computing time is spent with the calculation of the elements of the matrix X which is dominated by the calculation of the lattice sums $F_{l''m''}(\mathbf{k}_0)$.

We define vectors as

$$A_{\nu,l,m}(\mathbf{k}_{g'})=i^{l}(-1)^{m}Y_{l-m}(\Omega(\mathbf{k}_{g'}))e^{-i\mathbf{k}_{g'}\mathbf{d}_\nu}\ , \tag{8}$$

$$B_{\mu,l,m}(\mathbf{k}_g)=i^{-l}Y_{l,m}(\Omega(\mathbf{k}_g))e^{i\mathbf{k}_g\mathbf{d}_\mu}t_l^{(\mu)}\ , \tag{9}$$

$$Z_{\nu,l,m}=\sum_{\mu,l'm'}(1-X)^{-1}_{\nu lm,\mu l'm'}B_{\mu l'm'}\ , \tag{10}$$

and obtain for the scattering matrices

$$M_{g',g}=\sum_{\nu\mu}A_\nu[1-X]^{-1}_{\nu\mu}B_\mu=\sum_{\nu}A_\nu Z_\nu\ , \tag{11}$$

where the indices l,m and the arguments $\mathbf{k}_g$ have been dropped for convenience.

We expand the difference $\delta M_{g'g}=M_{g'g}(\mathbf{p}_0+\delta\mathbf{p})-M_{g'g}(\mathbf{p}_0)$ linearly in terms of $\delta\mathbf{p}$,

$$\begin{aligned}\delta M_{g'g}&=\sum_{\nu,\mu}A_\nu(\mathbf{p}_0+\delta\mathbf{p})[1-X(\mathbf{p}_0+\delta\mathbf{p})]^{-1}B_\mu(\mathbf{p}_0+\delta\mathbf{p})-\sum_{\nu,\mu}A_\nu(\mathbf{p}_0)[1-X(\mathbf{p}_0)]^{-1}B_\mu(\mathbf{p}_0)\\&\approx\sum_{\nu,\mu}\left[A_\nu(\mathbf{p}_0+\delta\mathbf{p})[1-X(\mathbf{p}_0)]^{-1}_{\nu,\mu}\frac{\partial B_\mu}{\partial\mathbf{p}}\delta\mathbf{p}+\frac{\partial A_\nu}{\partial\mathbf{p}}\delta\mathbf{p}([1-X(\mathbf{p})])^{-1}_{\nu,\mu}B_\mu(\mathbf{p})\right]\\&\quad+\sum_{\nu,\mu}\sum_{\nu'\mu'}A_\nu(\mathbf{p}_0+\delta\mathbf{p})[1-X(\mathbf{p}_0)]^{-1}_{\nu\nu'}\delta X_{\nu'\mu'}[1-X(\mathbf{p}_0)]_{\mu'\mu}B_\mu(\mathbf{p}_0+\delta\mathbf{p})\ .\end{aligned} \tag{12}$$

The derivative of the reflection and transmission coefficients thus consists of two parts. The first contribution arises from the geometrical phase occurring in the vectors A_ν and B_μ, and the second contribution arises from intralayer multiple scattering. Both parts can be treated separately and are of significant different importance as will be discussed below.

Equation (12) enables a very efficient calculation. The

vectors $A_\nu(\mathbf{p})$, $B_\mu(\mathbf{p})$, and $Z_\mu(\mathbf{p})$, respectively, as well as the spherical harmonics, have been calculated for the reference structure anyway. For each parameter increment only the matrix elements $\delta X=([X(\mathbf{p}_0+\delta\mathbf{p})-X(\mathbf{p}_0)])$ are required. In special cases, as will be discussed in Sec. II E, the calculation of δX can be skipped.

The lattice sum in Eq. (7) runs over all lattice points within a limiting radius given by the convergence criterion

$$-\mathrm{Im}(k_z)\cdot|\mathbf{P}|\leq\ln\varepsilon\ . \tag{13}$$

Setting $\varepsilon=0.001$ leads usually to a radius $R_{\max}$ of about 10–15 interatomic distances. If the interlayer spacing $\mathbf{d}_\mu-\mathbf{d}_\nu$ is varied by an increment $\delta\mathbf{d}_{\nu\mu}$, this changes the phase factors, the spherical Hankel functions, and the spherical harmonics. An analytic calculation of the derivatives would not really improve the speed of the calculation. It is therefore more advantageous to make use of the asymptotic behavior of the derivative of the spherical Hankel functions for large arguments

$$\frac{d}{dz}h_l(z)\approx ih_l(z)\ \text{for}\ z\gg1 \tag{14}$$

with

$$z=|\mathbf{k}||\mathbf{P}+\mathbf{d}_\mu-\mathbf{d}_\nu|\ .$$

We may consider the case that the position of atom μ has to be optimized. The spherical harmonics do not change much with the increment $\delta\mathbf{d}_{\nu\mu}$ at large polar angles, i.e., at large vectors $\mathbf{P}$ (see Fig. 1). For values of $|\mathbf{P}|$ larger than a limiting radius R_0, the contribution of the spherical harmonics to the lattice sum F_{lm} associated with the increment $\delta\mathbf{d}_{\nu\mu}$ can be neglected. Therefore the lattice sum is split into two parts, a near-neighborhood part which needs to be recalculated for the incremented parameter $d_{\nu\mu}=d_{\nu\mu}+\delta d_{\nu\mu}$ and a second part which by Eq. (14) is easily obtained from the corresponding part of the reference structure. It is advantageous to distinguish displacements parallel and vertical to the plane. The separate treatment of $\mathbf{d}_\parallel$ and $\mathbf{d}_\perp$ takes into account that displacements vertical to the surface do not influence the interatomic distances significantly for large $\mathbf{P}$ in contrast to lateral displacements; hence $\varepsilon_\parallel$ takes the values zero and unity for vertical and lateral displacements, respectively. We obtain for the increment of the lattice sum,

$$\begin{aligned}\delta F_{lm}^{\nu\mu}=\sum_{\mathbf{P}}f_{lm}^{\nu\mu}(\mathbf{P}+\mathbf{d}_\mu-\mathbf{d}_\nu+\delta\mathbf{d}_{\nu\mu})-f_{lm}^{\nu\mu}(\mathbf{P}+\mathbf{d}_\mu-\mathbf{d}_\nu)\approx\varepsilon_\parallel\sum_{|\mathbf{P}|>R_0}i^{l+1}e^{i\mathbf{k}_0\mathbf{P}}|\mathbf{k}_0|(|\mathbf{d}_\mu-\mathbf{d}_\nu+\delta\mathbf{d}_{\nu\mu}+\mathbf{P}|-|\mathbf{d}_\mu-\mathbf{d}_\nu+\mathbf{P}|)\\ \times h_l(|\mathbf{k}_0||\mathbf{d}_\mu-\mathbf{d}_\mu-\mathbf{d}_\nu+\mathbf{P}|)Y_{l-m}(\Omega(\mathbf{d}_\mu-\mathbf{d}_\nu+\mathbf{P}))\\ +\sum_{|\mathbf{P}|\leq R_0}f_{lm}^{\nu\mu}(\mathbf{P}+\mathbf{d}_\mu-\mathbf{d}_\nu+\delta\mathbf{d}_{\nu\mu})-f_{lm}^{\nu\mu}(\mathbf{P}+\mathbf{d}_\mu-\mathbf{d}_\nu)\ .\end{aligned} \tag{15}$$

Utilizing the symmetry properties of the spherical harmonics, the first term in Eq. (15) can be rearranged as

$$\begin{aligned}&\varepsilon_\parallel\sum_{|\mathbf{P}|>R_0}i^{l+1}|\mathbf{k}_0|(|\mathbf{d}_\mu-\mathbf{d}_\nu+\delta\mathbf{d}_{\nu\mu}+\mathbf{P}|-|\mathbf{d}_\mu-\mathbf{d}_\nu+\mathbf{P}|)h_l(|\mathbf{k}_0||\mathbf{d}_\mu-\mathbf{d}_\nu+\mathbf{P}|)Y_{l-m}(\Omega(\mathbf{d}_\mu-\mathbf{d}_\nu+\mathbf{P}))\\ &\approx\varepsilon_\parallel\sum_{|\mathbf{P}|>R_0}i^{l+1}|\mathbf{k}_0||\delta\mathbf{d}_{\nu\mu}|h_l(|\mathbf{k}_0||\mathbf{d}_\mu-\mathbf{d}_\nu+\mathbf{P}|)[Y_{l-m}(\Omega(\mathbf{d}_\mu-\mathbf{d}_\nu+\mathbf{P}))-Y_{l-m}(\mathbf{d}_\mu-\mathbf{d}_\nu-\mathbf{P})]\cos\varphi(\mathbf{P})\end{aligned} \tag{16}$$

with $\cos\varphi(\mathbf{P})=\mathbf{P}\delta\mathbf{d}_{\nu\mu}/|\mathbf{P}||\delta\mathbf{d}_{\nu\mu}|$.

In the derivation given above we have neglected the differential of the spherical harmonics due to its small influence on the lattice sum. That this contribution vanishes for large $\mathbf{P}$ may be seen also from the following argument. Because of the properties of the spherical harmonics at $\theta\approx\pi/2$,

$$\begin{aligned}Y_{l-m}(90°,\varphi)&=a_{l|m|}P_l^{|m|}(0)e^{im\varphi}\\ &=\begin{Bmatrix}0, & l-|m|\ \text{odd}\\ c_{l,|m|}, & l-|m|\ \text{even}\end{Bmatrix}e^{-im\varphi}\ .\end{aligned} \tag{17}$$

The first term in Eq. (16) vanishes for $l+m$ odd and large vectors $\mathbf{P}$. The term with $l+m$ even vanishes also for $m\neq0$ because of the diagonal dominance of the matrix X which works remarkably well at normal incidence,[19]

$$\delta X_{lm\nu;l'm'\mu}\approx\delta X_{lm\nu;l'm'\mu}\delta_{mm'}\ . \tag{18}$$

Finally the matrix δX is given by

$$\begin{aligned}\delta X&=X(\mathbf{P}_0+\delta\mathbf{P})-X(\mathbf{P}_0)\\ &\approx X^{(|\mathbf{P}|<R_0)}(\mathbf{P}_0+\delta P)-X^{(|\mathbf{P}|<R_0)}(\mathbf{P}_0)\ .\end{aligned} \tag{19}$$

A large reduction of the computational effort arises from the restriction of the lattice sum to $|\mathbf{P}|<R_0$, R_0 suitable. It should be noted that the number of lattice points scales quadratic with R_0. The latter can be chosen for 3–5 interatomic distances. The main multiple-scattering contribution arises from near-neighbor contributions. However, in order to keep the error in the intensities smaller than 0.001, usually 10–15 interatomic distances ($R_{\max}$) need to be included. For the calculation of derivatives this precision is not required because a linear approximation is not precise anyway. A cutoff at about 3–5 interatomic distances (R_0) is well justified.

The value for R_0 is determined by the validity of the asymptotic expansion of the derivative of the spherical Bessel functions. It generally depends on the quantum numbers l and m. However, we use an average value of about 3–5 interatomic distances independent of l and m. This results in an accuracy of the derivatives of 0.2–0.4 % which seems to be highly sufficient and leads to a reduction of the computing time for the matrix X by a factor of 4.

C. Implementation in the layer-doubling scheme

In the second step of the multiple-scattering calculation the scattering matrices of all (possibly composite) layers are stacked by the layer-doubling method[17] which treats the multiple scattering between layers exactly by matrix inversion, or alternatively by the RFS scheme which corresponds to an iterative solution of the matrix inversion.[20] We will treat here only the layer-doubling scheme because of its wider range of applicability. Let r_a, t_a and r_b, t_b be the reflection and transmission matrices for two layers a and b. The matrices r^{-+}, t^{++}, etc. are given by Eq. (4) choosing the proper directions of the incoming and outgoing waves. The reflection matrix for the stack of two layers is given by

$$R^{-+} = r_a^{-+} + t_a^{--} P^- r_b^{-+} P^+ \times (1 - r_a^{-+} P^- r_b^{-+} P^+)^{-1} t_a^{++} , \quad (20)$$

where r^{-+} describes the reflection of an incoming wave $e^{i\mathbf{k}_g^+ \mathbf{r}}$ into an outgoing wave $e^{i\mathbf{k}_g^- \mathbf{r}}$ and the plane-wave propagators $P^\pm$ are defined by

$$P^\pm = e^{\pm i \mathbf{k}_g^\pm \mathbf{d}} , \quad (21)$$

$\mathbf{d}$ is the distance vector between the two layers.

Applying the layer-doubling method for calculation of derivatives two different cases have to be distinguished. First, a variation in the layer-scattering matrices may occur due to an increment in the atomic coordinates or in the atomic scattering factors of one of the two composite layers which are coupled by layer doubling,

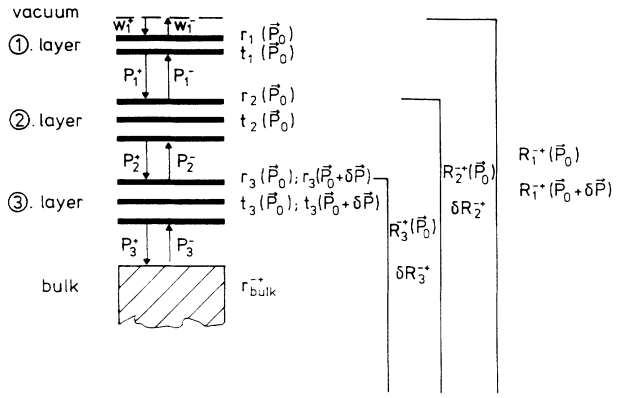

FIG. 2. Diagram of the layer-doubling approach for the case of a parameter variation in the third layer. The scattering matrices to be required in the approximation scheme are indicated.

$$\delta r^{-+} = r^{-+}(\mathbf{p}_0 + \delta\mathbf{p}) - r^{-+}(\mathbf{p}_0) ,$$
$$\delta t^{\pm\pm} = t^{\pm\pm}(\mathbf{p}_0 + \delta\mathbf{p}) - t^{\pm\pm}(\mathbf{p}_0) , \quad (22)$$

and second, by an increment of the interlayer distances which changes the interlayer propagators

$$\delta P^{-+} = P(1 \pm i \mathbf{k}_g \delta\mathbf{d}) . \quad (23)$$

In both cases the linear expansion in Eq. (3) can be applied to calculate δR^{-+}.

Let us first consider the case that an increment occurs in the parameter of the layer a leading to an increment δr_a and δt_a. Using the abbreviation

$$Q = P^- r_b^{-+} P^+ [1 - r_a^{+-}(\mathbf{p}_0) P^- r_b^{-+} P^+]^{-1} \quad (24)$$

and neglecting quadratic terms we obtain for the increment of the reflection matrix of the stack of two layers

$$\delta R^{-+} = \delta r_a^{-+} + t_a^{--}(\mathbf{p}_0) Q \delta r_a^{-+} Q t_a^{++}(\mathbf{p}_0) + \delta t_a^{--} Q t_a^{++}(\mathbf{p}_0) + t_a^{--}(\mathbf{p}_0) Q \delta t_a^{++} . \quad (25)$$

For a variation of layer b and using

$$W = [1 - r_a^{+-}(\mathbf{p}_0) P^- r_b^{-+} P^+]^{-1} \quad (26)$$

we obtain

$$\delta R^{-+} = t_a^{--} P^- r_b^{-+}(\mathbf{p}_0) P^+ W r_a^{+-} P^- \delta r_b^{-+} P^+ W t_a^{++} + t_a^{--} P^- \delta r_b^{-+} P^+ W t_a^{++}$$
$$= t_a^{--} P^- [r_b^{-+}(\mathbf{p}_0) P^+ W r_a^{+-} P^- + 1] \delta r_b^{-+} P^+ W t_a^{++} . \quad (27)$$

Similar formulas apply for the transmission matrices.

By varying the interlayer distance $\mathbf{d}$ we obtain

$$\delta R^{-+} = t_a^{--}([1 + P^- r_b^{-+} P^+ (1 - r_a^{+-} P^- r_b^{-+} P^+)^{-1} r_a^{+-}]) \times [\delta P^- r_b^{-+} P^+ + P^- r_b^{-+} \delta P^+](1 - r_a^{+-} P^- r_b^{-+} P^+)^{-1} t_a^{++} . \quad (28)$$

In both cases the term $(1 - r_a^{+-} P^- r_b^{-+} P^+)^{-1}$ in Eqs. (25), (27), and (28) can be calculated and stored prior to incrementing the structural parameter. As an example, in Fig. 2 the consecutive stacking of scattering matrices for the case of a parameter variation in the third composite layer is illustrated. The scattering matrices to be required in the approximation scheme are indicated. In the first step the third layer is combined with the bulk resulting in the reflection matrix of that compound configuration. The difference δR_3^{-+} has to be transported to the topmost layer via δR_2^{-+}.

It is advantageous to stack the layers starting from the topmost layer because then only multiplications of matrices with one vector appear.

D. Derivatives with respect to nonstructural parameters

The nonstructural parameters which are subject to optimization are the following: (i) the inner potential with a real and imaginary part and its energy dependence; (ii) the thermal vibrations entering the theory via the Debye temperature; (iii) an occupation factor. The latter occurs, e.g., in binary alloys or adsorbate layers with mixed occupation of adsorbate sites.

The optimization of the inner potential causes no problems. At normal or near-normal incidence, the refraction of the incident wave can be neglected and the calculated I/V curves are fitted to the experimental ones by shifting the energy scale. There is no need to calculate the derivative of the theoretical intensities with respect to the inner potential. It is completely equivalent to shift the experimental energy scale and to take the derivative of the experimental intensities without much computational effort. It corresponds to use $\partial I_{\rm ex}/\partial E$ instead of $\partial I_{\rm th}/\partial V_0$ in the optimization procedure. A variation of the energy dependence of the inner potential corresponds to a nonlinear expansion or compression of the energy scale and the corresponding differential is given by $(\partial I_{\rm ex}/\partial E)(\partial E/\partial p)$ where p is a parameter of the energy dependence of the inner potential, e.g., the frequently used form is a square-root dependence of V_0 on the energy $V_0=-\text{const}+p/E$ for values of E above ca. 50 eV. The imaginary part of the inner potential and its energy dependence have usually little influence on the structural result and these parameters are not considered here for optimization.

The treatment of temperature effects in the dynamical calculation using a Debye temperature θ_D and temperature-dependent phase shifts[17] allows a rapid calculation of derivatives with respect to this parameter. A variation of the thermal vibrations at the temperature T leaves the geometry unchanged and affects only the single-scattering t matrices via temperature-dependent phase shifts,[17,21]

$$t_l(\theta_D,T)=\frac{1}{i|\mathbf{k}_0|}e^{i\delta_l(\theta_D,T)}\sin\delta_l(\theta_D,T)\ . \tag{29}$$

The variation of layer-scattering matrices by a variation of the Debye temperature $\delta\theta_D$ is obtained starting from Eq. (4) and neglecting second-order terms. Using the vectors defined in Eqs. (8)–(10) we obtain

$$\begin{aligned}\delta M_{g'g}\approx&\sum_{\nu\mu}A_\nu(1-X(\theta_D))^{-1}\delta B_\mu\\&+\sum_{\nu\mu}A_\nu(1-X(\theta_D))^{-1}\delta X(1-X(\theta_D))^{-1}B_\mu(\theta_D)\ ,\end{aligned} \tag{30}$$

where

$$\begin{aligned}\delta B_\mu&=B_\mu(\theta_D+\delta\theta_D)-B(\theta_D)\ ,\\ \delta X&=X(\theta_D+\delta\theta_D)-X(\theta_D)\ .\end{aligned} \tag{31}$$

A statistical occupation of lattice sites or adsorbate sites can be introduced in the multiple-scattering formalism using the ATA approximation.[22] Its successful application in LEED structure analysis has been demonstrated in a number of surfaces of binary alloys.[23] The average t matrix for a binary alloy of component A and B is given by

$$t_l(c)=ct_{l,A}+(1-c)t_{l,B}\ , \tag{32}$$

where c is the concentration of element A and $(1-c)$ is the concentration of element B. An increment in the concentration δc leads to

$$\begin{aligned}\delta M_{g'g}=&\sum_{\nu,\mu}A_\nu(1-X(c_0))^{-1}(t_A-t_B)B_\mu\delta c\\&+\sum_{\nu,\mu}A_\nu(1-X(c_0))^{-1}(t_A-t_B)\delta cX(c_0)\\&\times([1-X(c_0)])^{-1}B_\mu\ .\end{aligned} \tag{33}$$

The increment $\delta M_{g'g}$ of the layer-scattering matrices can be inserted in the layer-doubling scheme as described in Sec. II C. The optimization of nonstructural parameters is a relatively fast process, because the time-consuming recalculation of electron propagator matrices is not required.

E. Simplifications for special cases

In many cases the existence of an adsorbate layer or a reconstructed surface layer causes slight distortions in the underlying substrate layers. The distortions may involve lateral and vertical displacements with vertical displacements being usually more influential on the LEED I/V data than the lateral displacements. This is not only a consequence of the backscattering geometry used in LEED. Vertical displacements occur frequently because there is more freedom for the atoms to move in a direction normal to the surface than parallel to the surface. Lateral displacements usually influence the bond lengths more than vertical displacements and in most cases adsorbate layers induce slight rumpling in the uppermost substrate layers. In these cases the calculation of the derivatives with respect to vertical displacements in a slightly buckled substrate layer becomes necessary. Corresponding calculations are very time consuming, because, on the one hand, the full overlayer unit cell has to be considered and, on the other hand, the substrate layers are usually densely packed resulting in a large number of atoms in the unit cell.

Simplifications can be made by taking advantage of the fact that the bond lengths change very little in a slightly buckled layer. Consequently it is sufficient to neglect the second term in Eq. (12) and to use only the change of the phase factors in the first part. The derivative of the scattering amplitudes is then given by

$$\begin{aligned}\frac{\partial M_{g'g}}{\partial P_j}\approx\sum_{\nu,\mu}\Bigg[&A_\nu(\mathbf{P})(1-X)^{-1}_{\nu\mu}\frac{\partial B_\mu}{\partial P_j}\\&+\frac{\partial A_\nu}{\partial P_j}(1-X)^{-1}_{\nu\mu}B_\mu(\mathbf{P})\Bigg]\ .\end{aligned} \tag{34}$$

This corresponds to a quasikinematic approximation because only the geometric phase is considered and the

multiple-scattering part is neglected.

We compare this approximation with a full dynamical calculation in Fig. 3(b). Obviously, there exist only very slight differences to the full dynamical calculation. As the derivatives are required to determine the direction of the next parameter increment and to estimate its magnitude, it is not important to calculate the derivatives more precisely. By the iteration process in the optimization scheme the full dynamical calculation is repeated in each step anyway.

As expected, we have found close agreement between full dynamical ("exact") and quasikinematical calculated derivatives of intensities only for a variation of vertical parameters. Lateral parameters, e.g., a pairing of neighboring atoms, exhibit a significant poorer correspondence between exact and quasikinematical derivatives, as shown in Fig. 4(b), upper part. Applying that simple approximation, the automatic structure refinement leads to an unreliable new structure and the iterative process does not converge to the optimum structure. However, this problem can easily be circumvented by including the multiple-scattering part according to Eq. (12). The high accuracy of derivatives with respect to lateral parameters obtained by this extended version is demonstrated in Fig. 4(b), lower part.

Summarizing, we can state that the derivatives of the LEED intensities with respect to vertical displacements in a slightly buckled layer is well approximated by a "kinematic" calculation while the variation of lateral structural parameters should be either performed in a full dynamical calculation or an extended approximation scheme described in Sec. II B.

It should be noted also that the quasikinematic approximation for the calculation of derivatives is valid only for slightly buckled layers. In the case of composite layers, where the interatomic distances equal nearly the distance between subplanes, the multiple-scattering contribution cannot be neglected.

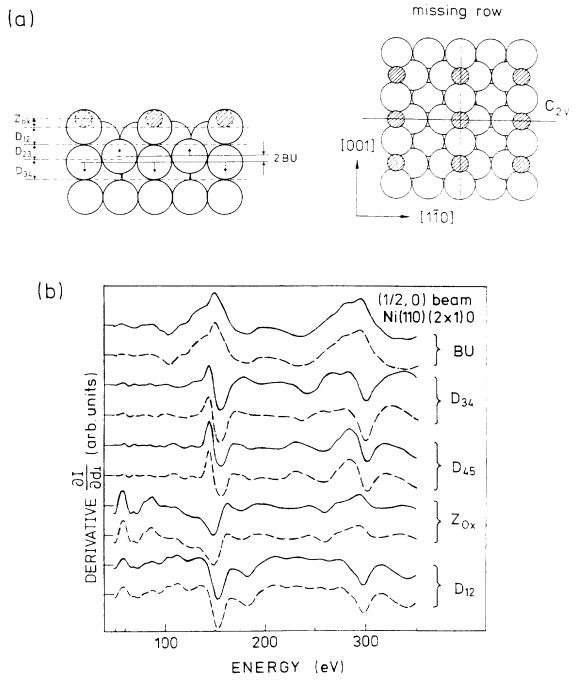

FIG. 3. (a) Structure model of O/Ni(110)-(2×1) and the different vertical displacements used in the calculation of the derivatives of intensities with respect to the parameters specified in the figure. (b) Comparison of the quasikinematically approximate (full line) and full dynamical calculation (dashed line) of derivatives.

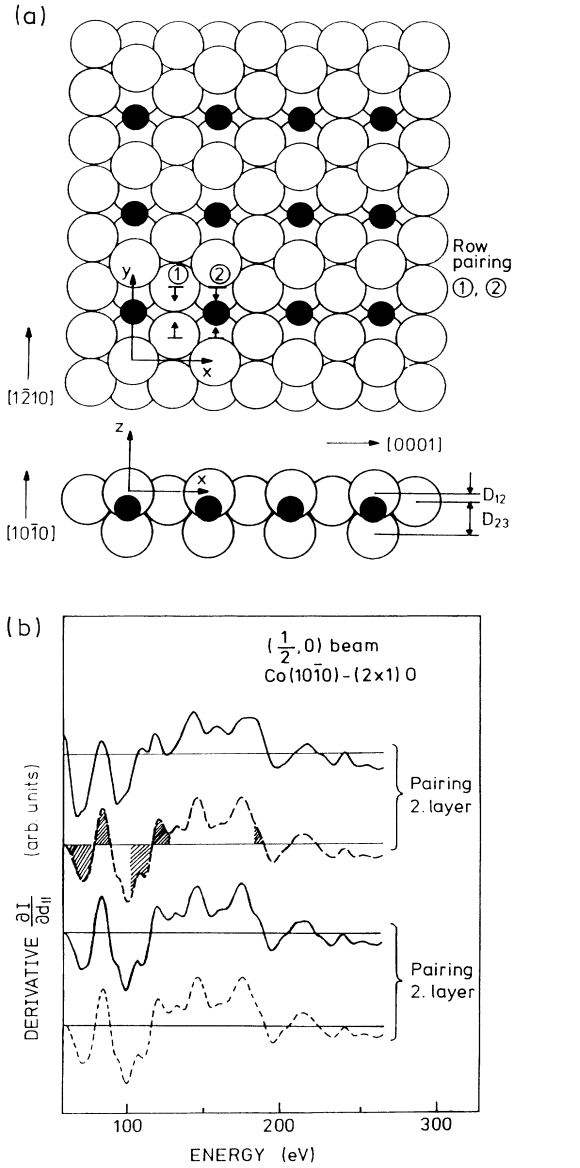

FIG. 4. Illustration of the limiting radius for a linear expansion of the intensities. (a) A structure model of O/Co(10$\bar{1}$0)-(2×2) and the different horizontal displacements used in the calculation of the derivatives. (b) Upper part: Comparison of the quasikinematically approximate (full line) and full dynamical calculation (dashed line) of derivatives. Lower part: Comparison of the approximate (including the multiple scattering part full line) and full dynamical calculation (dashed line) of derivatives.

III. DISCUSSION

The approximate calculation of derivatives of the intensities with respect to structural and nonstructural parameters provides a very efficient way to use optimization procedures in connection with structure determination by LEED. A large part of the initial calculation for the reference structure can be stored and used again for the incremented structure thus causing the total calculational effort to increase only weakly with the number of free parameters. The above methods therefore allow us to determine structures of complexity up to now not accessible by LEED.

A further acceleration of this algorithm can be achieved by using a reduced basis of wave functions in the spherical-wave expansion as well as in the plane-wave representation. We found that no loss of accuracy occurs (in the derivatives) reducing the number of angular-momentum components by about 40% by using, e.g., $l_{max}=6$ instead of $l_{max}=8$ in the full calculation. This may not be generally valid for high-Z atoms. Because the calculation in angular-momentum space scales with $(l_{max}+1)^3$ this part of the calculation is speeded up considerably. A similar reduction is also possible for the set of plane waves in the layer-doubling scheme.

The reduction in computing time achieved by applying linear approximations in the calculation of derivatives depends, of course, on the number of free parameters. In the case of O/Ni(110)-(2×1), with only six free parameters in the three uppermost layers, the full numerical calculation of derivatives for seven iterations, 21 energy points in 15-eV steps, 10 phase shifts, and 97 symmetrically independent beams amounts to about 5000 s on a CRAY-YMP. Using the approximations described above this time reduces to about 600 s. The save in computational effort becomes larger with an increasing number of free parameters. The number of free parameters is not the limiting factor for the complexity of a structure. The limits are mainly set by the number of symmetrically independent atoms in the unit cell of a composite layer for which a full dynamical calculation can be performed in reasonable computing times. It is further necessary to find a reasonable model with at least some agreement between experimental and theoretical I/V curves which can be used as a start model for the optimization procedure.

The calculation of derivatives of the intensity function is independent of the optimization procedure actually used for structure refinement. In the optimization scheme described previously,[6] the mean-square deviation between experimental and theoretical intensities are minimized. It has been found in a number of previous studies that this criterion is sufficient to localize the optimum structure, with the advantage that intensities need only be calculated in large steps on the energy scale thus reducing the calculational effort by an additional factor of 3–5. In all cases where a comparison could be made, even in the cases where the final agreement was marginal, the same structural parameters were found as determined from a grid search using Pendry's R factor R_P.[16] It remains, however, to be investigated in detail which fitting function is best suited for application with LEED. The comparison of relative intensities certainly gives the absolute height of maxima in the I/V curves a weight which may not be appropriate. In the multiple-scattering theory the thermal vibrations are not correctly treated, and from the experience of x-ray structure determination one may conclude that this is the cause for the major discrepancies between theory and experiment in the present state of the analysis. Additional uncertainties occur in the measurement. Effects from misalignment of the surface normal, energy dependence of the detector function, or inhomogeneities of the display screen, etc., may influence the absolute height of maxima but to a lesser extent the position of maxima on the energy scale.

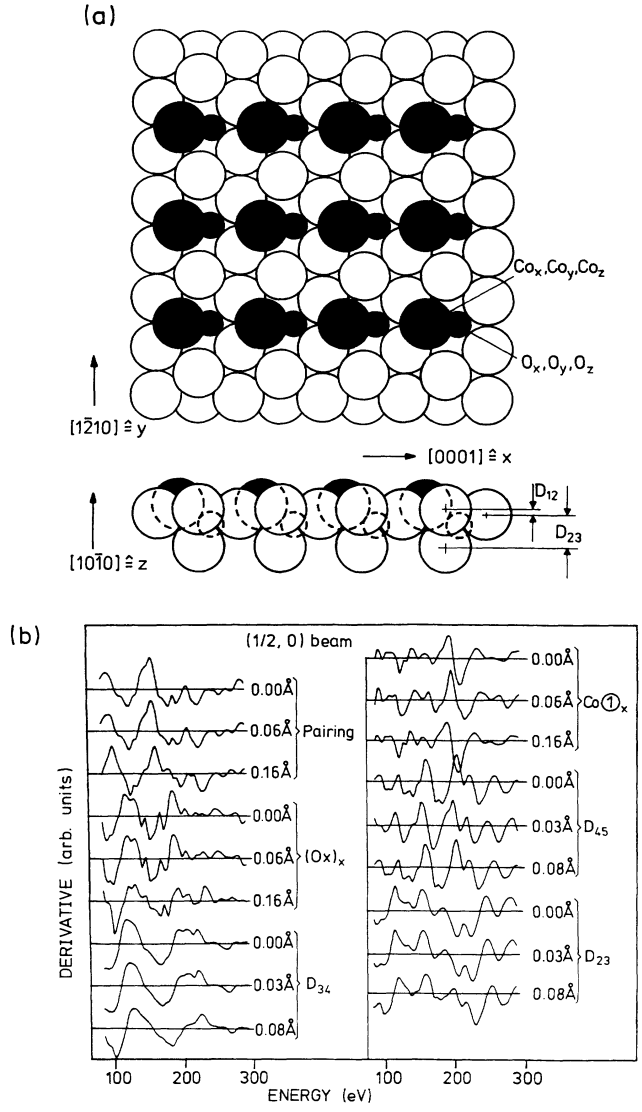

FIG. 5. (a) Structure model of O/Co(10$\bar{1}$0)-(2×1) and the different parameters used in the calculation of the derivatives. (b) Derivatives of LEED intensities with respect to different horizontal and vertical parameters for the reference structure (0.0 Å) and model structures with small increments in parameters (0.06 Å, 0.16 Å for lateral parameter and 0.03 Å, 0.08 Å for vertical parameter).

Therefore it may be more appropriate to use a different fitting function which weights the position of peaks as it does the Y function proposed by Pendry.[4,16] The Y function can likewise be implemented in the least-squares algorithm. The only disadvantage is that the step width on the energy scale has to be chosen on a correspondingly small scale to allow the calculation of the derivative of the I/V curve with respect to the energy required by the Y function. The relative intensities, on the other hand, are certainly more important than the peak position in determining vibrational amplitude, and in case the theory can be developed further towards a more precise calculation of temperature factors the comparison of intensities will be more useful.

The numerical calculation of derivatives as outlined above is related to the tensor LEED method proposed by Pendry and co-workers[2] but uses a different mathematical approach. Here only a linear approximation is used and the nonlinearity of the intensity function is overcome by an iterative solution in the fit procedure. A new full dynamical calculation is usually required at each iteration step.

A further improvement of the optimization scheme can be achieved by calculating the LEED intensities around the reference structure by a linear expansion utilizing the corresponding derivatives with respect to the parameters to be optimized. The reliability of this procedure is demonstrated in Fig. 5, where the derivatives of different vertical and horizontal displacements are shown for various model structures (0.0 indicates the reference structure). The derivatives do not change significantly for model structures with parameter variations around the reference structure within the limits of 0.06 Å for lateral and 0.03 Å for vertical displacements. Thus a linear expansion of LEED intensities should be applicable within these limits. Thus a new full dynamical calculation is only required at each second or third iteration step.

In the tensor LEED method the main calculational effort is put in calculating the tensor and saving computing time by calculating all structures around the reference structure with the once calculated tensor. The convergence radius of this method is probably larger than that of the least-squares optimization. A detailed study of the optimization procedures using different fitting functions and a comparison with the tensor LEED method is in progress.

ACKNOWLEDGMENTS

Part of this work was supported by SFB 338 of the Deutsche Forschungsgemeinschaft. The authors are grateful to M. A. Van Hove and M. Gierer for helpful discussions.

[1]P. G. Powell and V. E. de Carvalho, Surf. Sci. **187**, 175 (1987).
[2]P. J. Rous, J. B. Pendry, D. K. Saldin, K. Heinz, K. Müller, and N. Bickel, Phys. Rev. Lett. **57**, 2951 (1986).
[3]P. Rous, M. A. Van Hove, and G. A. Somorjai, Surf. Sci. **226**, 15 (1990).
[4]J. B. Pendry and K. Heinz, Surf. Sci. **230**, 137 (1990).
[5]G. Kleinle, W. Moritz, D. L. Adams, and G. Ertl, Surf. Sci. **219**, L637 (1989).
[6]G. Kleinle, W. Moritz, and G. Ertl, Surf. Sci. **238**, 119 (1990).
[7]W. Moritz, H. Over, G. Kleinle, and G. Ertl, in *The Structure of Surfaces III,* edited by S. Y. Tong, M. A. Van Hove, X. Xide, and K. Takayanagi (Springer, Berlin, 1991).
[8]D. W. Marquardt, J. Soc. Inc. Appl. Math. **11**, 431 (1963).
[9]H. Over, G. Kleinle, G. Ertl, W. Moritz, K.-H. Ernst, H. Wohlgemuth, K. Christmann, and E. Schwarz, Surf. Sci. Lett. **254**, L469 (1991).
[10]H. Over, H. Bludau, M. Skottke-Klein, W. Moritz, and G. Ertl, Phys. Rev. B **45**, 8638 (1992); **46**, 4360 (1992).
[11]Y. Gauthier, R. Baudoing-Savois, and W. Moritz, Phys. Rev. B **44**, 12 977 (1991).
[12]P. J. Rous and J. B. Pendry, Comput. Phys. Commun. **54**, 135 (1989); **54**, 157 (1989).
[13]J. B. Pendry, K. Heinz, and W. Oed, Phys. Rev. Lett. **61**, 175 (1987).
[14]J. B. Pendry and K. Heinz, Surf. Sci. **230**, 137 (1990).
[15]P. J. Rous, D. Jetz, D. G. Kelly, R. Q. Hwang, M. A. Van Hove, and G. A. Somorjai, in *The Structure of Surfaces III* (Ref. 7).
[16]J. B. Pendry, J. Phys. C **13**, 937 (1980).
[17]J. B. Pendry, *Low Energy Electron Diffraction* (Academic, New York, 1974).
[18]D. J. Titterington and C. G. Kinniburgh, Comput. Phys. Commun. **20**, 237 (1980).
[19]H. D. Shih and S. W. Tam, in *Determination of Surface Structure by LEED,* edited by P. M. Marcus and F. Jona (Plenum, New York, 1984).
[20]D. L. Adams, J. Phys. C **14**, 789 (1981).
[21]C. B. Duke and G. E. Lavamore, Phys. Rev. B **2**, 4765 (1970); **2**, 4783 (1970).
[22]F. Jona, K. O. Legg, H. D. Shih, D. W. Jepsen, and P. M. Marcus, Phys. Rev. Lett. **40**, 1466 (1978).
[23]Y. Gauthier and R. L. Baudoing, in *Surface Segregation and Relative Phenomena,* edited by P. A. Dowben and A. Miller (CRC, Boca Raton, FL, 1990).