Dielectric response of an inhomogeneous quasi-two-dimensional electron gas

F. J. Fernández-Velicia
Departamento de Ciencia de Materiales, Universidad Nacional de Educación a Distancia, Senda del Rey s/n, Madrid, Spain

F. García-Moliner and V. R. Velasco
Instituto de Ciencia de Materiales, Consejo Superior de Investigaciones Científicas, Serrano 123, 28006 Madrid, Spain
(Received 25 April 1995)

The solution of the integral equation required to invert the dielectric function of a confined quasi-two-dimensional electron gas is studied by means of a formal analysis which yields a convergent algorithm. The dielectric function can then be inverted in real space for an arbitrary number of populated subbands and taking into account the effect of intersubband excitations involving empty subbands to any desired degree of accuracy. Plasma modes and screened potential can then be easily studied by using a basis which bears out explicitly the consequences of symmetry in symmetric systems. A model calculation of dynamical screening at frequencies of the order of those of confined polar optical modes in usual GaAs wells indicates that the empty states may play a quite significant role and the screened potential, explicitly obtained in real space, may exhibit a great variety of behaviors: the sign of the potential may change and its magnitude may be either reduced (ordinary screening) or enhanced (antiscreening).

I. INTRODUCTION

A standard self-consistent type of argument leads to a relationship of the form

$$V_e(\mathbf{r}) = \int d\mathbf{r}'\, \epsilon(\mathbf{r},\mathbf{r}')V(\mathbf{r}'), \tag{1}$$

where V_e is the external bare potential, V is the total screened potential, and the dependence on t (or ω) is understood everywhere when not explicitly indicated. The problem is to know V given V_e, which amounts to solving the integral equation (1) for V.

In a homogeneous three-dimensional (3D) electron gas, the problem is trivially solved in full Fourier transform, from $\epsilon(\mathbf{r}-\mathbf{r}';t-t')$ to $\epsilon(\mathbf{K},\omega)$ and the inverse is simply $1/\epsilon(\mathbf{K},\omega)$. In this case, the theoretical effort has been mostly concentrated on trying to improve $\epsilon(\mathbf{K},\omega)$ beyond the random phase approximation (RPA).[1] A different situation arises in a number of systems of considerable physical interest, like δ-doped semiconductors or modulation doped heterojunctions or quantum wells, where ϵ is of the form $\epsilon(\boldsymbol{\rho}-\boldsymbol{\rho}',z,z';t,t')$, $\boldsymbol{\rho}$ is the in plane 2D position vector, and z the coordinate in the perpendicular direction. In this case, one has to deal with an integral equation in the z variable, which cannot be solved by simple Fourier transform.

The usual approach is to take matrix elements of the integral equation between one electron wave function and then find the matrix elements of the screened potential by making often drastic approximations. This paper presents a different approach in which the integral equation itself is explicitly solved without these approximations.

The free particle in plane motion will be described in terms of plane waves with a 2D wave vector $\boldsymbol{\kappa}$, while the quantized motion in the z direction is described by some wave functions $\varphi_n(z)$. In fact, there are the wave functions of the discrete bound states under study—e.g., a quantum well, a modulation doped heterojunction, or a δ-doped system—and those of the continuum of free states above the barriers. For practical purposes, we can embed the system in a sufficiently large quantization box, in the z direction with infinite potential barriers at the extremes. The whole spectrum of electronic states then formally consists of discrete quantized states in the z direction. The formal argument could be equally carried out with a continuous part in the spectrum, but it is simpler this way. Thus, if $\alpha=(\boldsymbol{\kappa},n)$ denotes the set of quantum numbers, the electronic wave functions will be written

$$\psi_\alpha(\mathbf{r}) = S^{-1/2}\exp(i\boldsymbol{\kappa}\cdot\boldsymbol{\rho})\varphi_n(z), \tag{2}$$

where S is a normalization area. Consistent with this, we use the 2D Fourier transform convention,

$$\begin{aligned} f(\boldsymbol{\rho}) &= \sum_{\kappa} f(\kappa)\exp(i\mathbf{Q}\cdot\boldsymbol{\rho}), \\ f(\mathbf{Q}) &= \int_S d^2\boldsymbol{\rho} f(\boldsymbol{\rho})\exp(-i\mathbf{Q}\cdot\boldsymbol{\rho}). \end{aligned} \tag{3}$$

With $(\omega,\mathbf{Q})$ dependence understood, our integral equation reads

$$V_e(z) = \int dz'\, \epsilon(z,z')V(z'). \tag{4}$$

The purpose of this paper is to find the solution

$$V(z) = \int dz'\, \epsilon^{-1}(z,z')V_e(z') \tag{5}$$

under general conditions, as will be presently explained.

Whether for screening or for energy loss calculations, the problem of finding ϵ^{-1} for a quantum well has been treated in various approximations. This is an easy problem if the confined electron gas is described as strictly 2D.[2] The main point of such an approximation is not only that the electronic wave functions are treated as strictly 2D, but also—and more

important, as will be seen later—that the spectrum of states describing the quantized motion in the z direction is reduced to just one single band. The finite spread of the electronic wave functions in the z direction has been taken into account in several calculations of ϵ^{-1}, often based on the assumption that all carriers are confined to the lowest subband and ignoring the rest of the spectrum, that is the empty subbands of electronic states. Two occupied subbands have been studied in some cases, though often with some limitations like the restriction to static screening[3] and also ignoring the empty subbands. One of the most articulate attempts so far appears to be a treatment of a metallic film[4] including up to two occupied subbands and some of the empty ones.

We stress three features common to the usual treatment of the inversion of ϵ for a confined quasi-two-dimensional system, namely: (i) The problem is formulated in terms of matrix elements of the different potentials between electronic wave functions, so that one seeks to obtain the matrix elements of the screened potential in terms of those of the bare potential. (ii) It is usually assumed that only one or two subbands are occupied and, with exceptions like the one noted above,[4] the empty subbands are usually ignored. (iii) Even when the effect of some empty subbands is included, the problem is treated by means of a finite truncation of the spectrum, for which no formal justification is given.

The last question is far from trivial. The problems amounts to solving an integral equation and this ultimately amounts to inverting a matrix—which depends on the basis one chooses. The point is that this matrix *is infinite* and, while one may expect a finite truncation of the basis to yield, in practice, a reasonable approximation, a formal mathematical analysis is, in principle, mandatory when infinities are involved. Such an analysis has been recently given by the authors[5] and, based on this, the inversion of ϵ is formulated here in a way which provides an explicit formula for $\epsilon^{-1}(z,z')$, which can be used, in practice, for any spectrum of electronic wave functions $\varphi_n(z)$ and any number of occupied or empty subbands.

The self-consistent field expression for $\epsilon(z,z')$ is summarized in Sec. II and a suitable set of basis functions is introduced in Sec. III, where the formula for $\epsilon^{-1}(z,z')$ is given in a way that displays the consequences of symmetry for symmetric structures. The main concern here is with a sound method for the inversion of $\epsilon(z,z')$, but the analysis also provides an interesting viewpoint on the collective modes of the confined inhomogeneous electron gas. Plasmons are incidentally discussed in Sec. IV, which is mainly devoted to the discussion of the screened potential for a simple model used to demonstrate the practical use of the method. Some final comments are made in Sec. V.

II. THE RPA DIELECTRIC FUNCTION FOR A QUASI-TWO-DIMENSIONAL INHOMOGENEOUS ELECTRON GAS

We summarize here the standard self-consistent analysis, which allows us to write $\epsilon(\mathbf{r},\mathbf{r}')$ in a suitable form in real space. Let $\psi_\alpha(\mathbf{r})$ denote quite generally the electronic wave functions, not necessarily those of (2). An external potential V_e produces an induced particle density $N_{\rm ind}$ and induced potential $V_{\rm ind}$, with

$$N_{\rm ind}(\mathbf{r})=\sum_{\alpha,\alpha'}\frac{f_\alpha-f_{\alpha'}}{E_\alpha-E_{\alpha'}-\hbar\omega^+}\langle\alpha|V|\alpha'\rangle\psi^*_{\alpha'}(\mathbf{r})\psi_\alpha(\mathbf{r}), \tag{6}$$

where $V=V_e+V_{\rm ind}$, $\omega^+=\lim_{\eta\to0}(\omega+i\eta)$ and the summations are just symbols that may or may not involve integrations. $N_{\rm ind}$ and $V_{\rm ind}$ are related by the Poisson equation,

$$\nabla^2V_{\rm ind}(\mathbf{r})=-\frac{4\pi e^2}{\epsilon_x}N_{\rm ind}(\mathbf{r}), \tag{7}$$

with appropriate boundary conditions. Here, we assume simply a background dielectric constant ϵ_x. However, in view of the interest of this analysis for the study of heterostructures, some remarks will be made later on the treatment of dielectric discontinuities. The purpose, for the moment, is to focus on the essential problem—the inversion of ϵ—without paying attention to inessential details.

Let $G(\mathbf{r},\mathbf{r}')$ be the electrostatic response Green function defined by

$$\nabla^2_{\mathbf{r}}G(\mathbf{r},\mathbf{r}')=-4\pi\delta(\mathbf{r}-\mathbf{r}'), \tag{8}$$

with the same boundary conditions as (7). Then

$$V_{\rm ind}(\mathbf{r})=\int\mathbf{d}^3\mathbf{r}'G(\mathbf{r},\mathbf{r}')N_{\rm ind}(\mathbf{r}'). \tag{9}$$

Writing down explicitly the matrix elements $\langle\alpha|V|\alpha'\rangle$ as integrals and expressing V_e as $V-V_{\rm ind}$, from (6)–(9), one finds Eq. (1) with

$$\epsilon(\mathbf{r},\mathbf{r}')=\delta(\mathbf{r}-\mathbf{r}')-\frac{e^2}{\epsilon_x}\sum_{\alpha,\alpha'}\frac{f_\alpha-f_{\alpha'}}{E_\alpha-E_{\alpha'}-\hbar\omega^+}\times\left[\int\mathbf{d}^3\mathbf{r}''G(\mathbf{r},\mathbf{r}'')\psi_\alpha(\mathbf{r}'')\psi^*_{\alpha'}(\mathbf{r}'')\right]\times[\psi^*_\alpha(\mathbf{r}')\psi_{\alpha'}(\mathbf{r}')]. \tag{10}$$

For a 3D homogeneous electron gas, this yields the dielectric function in real space as $\epsilon(\mathbf{r}-\mathbf{r}')$ and its 3D Fourier transform yields the well known RPA formula for $\epsilon(\mathbf{K},\omega)$. For the quasi-two-dimensional system with the wave functions of (2), this yields

$$\epsilon(\boldsymbol{\rho}-\boldsymbol{\rho}';z,z')=\delta(\boldsymbol{\rho}-\boldsymbol{\rho}')\delta(z-z')-\frac{e^2}{\epsilon_xS}\sum_{\substack{\boldsymbol{\kappa},\boldsymbol{\kappa}'\\n,n'}}\frac{f_{\boldsymbol{\kappa},n}-f_{\boldsymbol{\kappa}',n'}}{E_{\boldsymbol{\kappa},n}-E_{\boldsymbol{\kappa}',n'}-\hbar\omega^+}\times\left[\int dz''\varphi_n(z'')\varphi^*_{n'}(z'')\int\mathbf{d}^2\boldsymbol{\rho}''G(\boldsymbol{\rho}-\boldsymbol{\rho}'';z,z'')e^{i(\boldsymbol{\kappa}-\boldsymbol{\kappa}')\cdot\boldsymbol{\rho}''}\right]\times[\varphi^*_n(z')\varphi_{n'}(z')e^{-i(\boldsymbol{\kappa}-\boldsymbol{\kappa}')\cdot\boldsymbol{\rho}'}], \tag{11}$$

whence the 2D Fourier transform, from $\boldsymbol{\rho}-\boldsymbol{\rho}'$ to $\mathbf{Q}$:

$$\epsilon(z,z')=\delta(z-z')-\sum_{n,n'}\left[\int dz''G(z,z')\varphi_n(z'')\varphi_{n'}^*(z'')\right]\times X_{n,n'}[\varphi_n^*(z')\varphi_{n'}(z')], \tag{12}$$

where

$$G(z,z')=\int \mathbf{d}^2(\boldsymbol{\rho}-\boldsymbol{\rho}'')G(\boldsymbol{\rho}-\boldsymbol{\rho}'';z,z')e^{i\mathbf{Q}\cdot(\boldsymbol{\rho}-\boldsymbol{\rho}'')}, \tag{13}$$

and $X_{n,n'}$ is the 2D polarizability element,

$$X_{n,n'}(\mathbf{Q},\omega)=\frac{e^2}{\epsilon_x S}\sum_{\boldsymbol{\kappa}}\frac{f_{\boldsymbol{\kappa}+\mathbf{Q},n'}-f_{\boldsymbol{\kappa},n}}{E_{\boldsymbol{\kappa}+\mathbf{Q},n'}-E_{\boldsymbol{\kappa},n}+\hbar\omega^+}. \tag{14}$$

Our problem is to find the inverse $\epsilon^{-1}(z,z')$ satisfying the condition

$$\int dz''\epsilon(z,z'')\epsilon^{-1}(z'',z')=\delta(z-z'), \tag{15}$$

from which we obtain the solution (5) for the screened potential.

III. THE INVERSE DIELECTRIC FUNCTION

We define the functions

$$S_{n,n'}(z)=\varphi_n^*(z)\varphi_{n'}(z);$$

$$L_{n,n'}(z)=\int dz'G(z,z')S_{n,n'}(z'). \tag{16}$$

Then,

$$\epsilon(z,z')=\delta(z-z')-\sum_{n,n'}L_{n,n'}^*(z)X_{n,n'}S_{n,n'}(z'). \tag{17}$$

It will be presently shown that ϵ^{-1} is of the form

$$\epsilon^{-1}(z,z')=\delta(z-z')+\sum_{n,n'}L_{n,n'}^*(z)M_{n,n'}S_{n,n'}(z'). \tag{18}$$

The $(\omega,\mathbf{Q})$ dependence is in $M_{nn,'}$. Thus, given an external potential $V_e(z')$, the induced potential at z is

$$V_{\text{ind}}(z)=\sum_{n,n'}L_{n,n'}^*(z)M_{n,n'}\int dz'S_{n,n'}(z')V_e(z'). \tag{19}$$

In a typical quantum well problem, $V_e(z')$—e.g., the potential of a confined optical mode—is characteristically confined within the bounded domain of the quantum well, which is also where the amplitudes $\varphi_n(z)$ of the lower part of the spectrum are mostly confined, while due to the long range nature of the Coulomb interaction, $V_{\text{ind}}(z)$ may reach out to longer distances, which is the feature distinguishing $L_{n,n'}(z)$ from $S_{n,n'}(z')$. We shall denote these as long (L) and short (S) range functions, respectively. To simplify the presentation, we shall assume that the φ_n are real, although there is no difficulty in dealing with complex eigenfunctions. Then $L_{n,n'}=L_{n',n}$ and $S_{n,n'}=S_{n',n}$ are also real and the sum in (17) takes the form

$$\sum_{n\geqslant n'}L_{n,n'}(z)P_{n,n'}S_{n,n'}(z), \tag{20}$$

where we have defined

$$X_{n,n}\equiv P_{n,n};\quad X_{n,n'}+X_{n',n}=P_{n,n'}\quad (n\neq n'). \tag{21}$$

We also assume degenerate statistics. This is not a requirement of the formal basis of the analysis, but the presentation is again simpler and, in fact, the confined quasi-two-dimensional electron gases that one encounters, in practice, are often degenerate to a good approximation. Then the $P_{n,n'}$ vanish when both subbands (n and n') are empty and the matrix of nonvanishing elements $P_{n,n'}$ $(n\geqslant n')$, with k subbands occupied, is

$$\begin{array}{cccccccc}
(1,1) & & & & & & & \\
(2,1) & (2,2) & & & & & & \\
(3,1) & (3,2) & (3,3) & & & & & \\
\cdots & \cdots & \cdots & \cdots & & & & \\
\cdots & \cdots & \cdots & \cdots & \cdots & & & \\
(k,1) & (k,2) & (k,3) & \cdots & (k,k) & & & \\
(k+1,1) & (k+1,2) & (k+1,3) & \cdots & (k+1,k) & 0 & & \\
(k+2,1) & (k+2,2) & (k+2,3) & \cdots & (k+2,k) & 0 & 0 & \\
(k+3,1) & (k+3,2) & (k+3,3) & \cdots & (k+3,k) & 0 & 0 & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots
\end{array}. \tag{22}$$

Now start with the main diagonal and define a relabeling according to the assignment

$$\begin{array}{ccccccc}(n,n') & (1,1) & (2,2) & (3,3) & \cdots & (k,k)\\ \mu & 1 & 2 & 3 & \cdots & k.\end{array}$$

Then take the second diagonal $(n+1,n)$ and assign

$$\begin{array}{ccccccc}(n,n') & (2,1) & (3,2) & (4,3) & \cdots & (k+1,k)\\ \mu & k+1 & k+2 & k+3 & \cdots & 2k.\end{array}$$

The following diagonal $(n+2,n)$ would have the assignment

$$\begin{array}{ccccccc}(n,n') & (3,1) & (4,2) & \cdots & \cdots & (k+2,k)\\ \mu & 2k+1 & 2k+2 & \cdots & \cdots & 3k,\end{array}$$

and so on. From each diagonal, we have k terms and each diagonal is characterized by an increasing value of $(n-n')$. We then case (18) as

$$\epsilon(z,z')=\delta(z-z')+\sum_{\mu} L_{\mu}(z)P_{\mu}S_{\mu}(z'). \tag{23}$$

Now, given (23) we seek the inverse,

$$\epsilon^{-1}(z,z')=\delta(z-z')+\sum_{\mu,\nu} L_{\mu}(z)M_{\mu,\nu}S_{\nu}(z'). \tag{24}$$

The point is that μ and ν run from 1 to infinity and finding the (infinite) matrix of elements $M_{\mu\nu}$ amounts to inverting an infinite matrix, and this should not be treated lightly. One may hope that only a limited number of spectral terms will yield, in practice, a reasonable approximation, but this is not a mathematically valid argument. The usual procedure, which consists in truncating the spectrum of electronic subbands up to a finite (and usually rather small) finite number, may or may not work reasonably in practice, but the presence of infinities requires, in principle, first a formal proof of the existence of the inverse and then the establishment of an algorithm, which generates a sequence of solutions converging to the exact answer. Physically speaking, the presence of intersubband excitations in the polarizability elements means that the empty part of the spectrum plays a role and the assumption, implicit in the usual calculations, that higher order terms will be negligible on account of increasing energy denominators may or may not be very accurate, depending on the numerics of the situation, as $\hbar\omega$ may be of the order of some $E_{n',\boldsymbol{\kappa}+\mathbf{Q}}-E_{n,\boldsymbol{\kappa}}$. Only a proper calculation can tell and the point is that it is the establishment of a sufficiently powerful method of calculation that really requires a formal analysis. Here we shall outline the main points of such an analysis[5] applied to the specific case of the dielectric function.

We define the matrix β of elements

$$\beta_{\mu\nu}=\int dz S_{\mu}(z)L_{\nu}(z). \tag{25}$$

Then, from the condition (15) we have, in compact notation,

$$-\mathbf{L}(z)\cdot\mathbf{P}\cdot\mathbf{S}(z')+\mathbf{L}(z)\cdot\mathbf{M}\cdot\mathbf{S}(z')-\mathbf{L}(z)\cdot\mathbf{P}\cdot\boldsymbol{\beta}\cdot\mathbf{M}\cdot\mathbf{S}(z')=0, \tag{26}$$

where $\mathbf{P}$ is the diagonal matrix of elements $P_{\mu}\delta_{\mu\nu}$, $\mathbf{L}$ is the vector of components L_{μ}, and $\mathbf{S}$ the vector of components S_{ν}. All the L_{μ} and S_{ν} are linearly independent, and thus they form a dual basis in a space of infinite dimensions. We represent the problem in this basis, which is obviously intrinsically more adequate to describe the short range and long range aspects of the physical problem under study than the basis of the electronic wave functions φ_n. Then the problem is to solve the matrix equation,

$$(\mathbf{I}-\mathbf{P}\cdot\boldsymbol{\beta})\cdot\mathbf{M}=\mathbf{P}. \tag{27}$$

Formally, the solution is simply

$$\mathbf{M}=(\mathbf{I}-\mathbf{P}\cdot\boldsymbol{\beta})^{-1}\cdot\mathbf{P}, \tag{28}$$

but this is just what requires the inversion of an infinite matrix.

Let us define successive truncations of (22), and hence of the sum in (24), by means of a *cut index c* defined as follows: For k occupied subbands, the first cut corresponds to $c=k$ and includes the first k diagonals of (22), the next cut has cut index $c=k+1$ and includes the first $k+1$ diagonals, and so on. Then c determines the maximum value $N=c\times k$ attained by μ or ν and the matrices in (27) and (28) are of order N. The problem is (i) to prove the existence of the inverse of the infinite matrix and (ii) to prove that the successive solutions for successive values of N converge to the exact answer as $N\to\infty$. This can be proved by applying the mathematical analysis of Ref. 5.

What matters is the asymptotic behavior of the terms in the series (24) for large values of μ. Irrespective of the details of the confined system under study, the electronic wave functions φ_n for large energies tend asymptotically to the sine functions of the confining box and then it is easily seen that for large n and n' the $X_{n,n'}$, when they do not vanish—see (22)—decrease for large n,n' as the inverse of $(n+n')(n-n')$ and then the corresponding $L_{n,n'}$ decrease like $1/n^2$. Now, in this asymptotic range, both n and n' are proportional to μ, and so are, therefore, $(n+n')$ and $(n-n')$. Thus, for large μ, the terms in the series (24) decrease like $1/\mu^4$ and the series is absolutely and uniformly convergent. Under these circumstances, the conditions established in Ref. 5 hold, which among other things guarantees that the terms in the series can be rearranged to suit convenience, as we have done by choosing the relabeling assignment $(n,n')\to\mu$, as explained after (22).

Moreover, a very useful algorithm was described in Ref. 5 by starting from the well known fact that any given matrix $\mathbf{a}$ can be factorized as the product

$$\mathbf{a}=\mathbf{p}\cdot\mathbf{t} \tag{29}$$

of a unitary matrix $\mathbf{p}$ and a triangular matrix $\mathbf{t}$. Direct, *nonrecursive* formulas were given to obtain the elements of $\mathbf{p}$ and $\mathbf{t}$ from the given elements of $\mathbf{a}$. Thus, the problem of inverting $\mathbf{a}$, since $\mathbf{p}^{-1}$ is simply $\mathbf{p}^{\dagger}$, is reduced to the much simpler problem of inverting the triangular matrix $\mathbf{t}$, where all the elements below the main diagonal vanish. The procedure is very efficient on account of the nonrecursive character of the formulas yielding directly $\mathbf{p}$ and $\mathbf{t}$.

Having obtained the matrix $\mathbf{M}$ to the desired degree of accuracy, we can obtain explicitly the screened potential,

$$V(z)=V_e(z)+\sum_{\mu} L_{\mu}(z)M_{\mu\nu}\int dz' S_{\nu}(z')V_e(z'). \tag{30}$$

This formula is quite general, but many systems of physical interest, like symmetric modulation doped quantum wells or δ-doped systems, have mirror image symmetry. Then all the φ_n have definite parity and hence also all the S_μ and L_μ, so by suitable rearrangement these functions can be arranged in symmetric (S) and antisymmetric (A) sets, which are mutually orthogonal and the scalar products of (25) vanish for all μ, ν when S_μ and L_ν have different parity. The space of the problem is then factorized in mutually orthogonal S and A subspaces and β is factorized in blocks in the form

$$\boldsymbol{\beta}=\begin{Vmatrix} \boldsymbol{\beta}_S & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\beta}_A \end{Vmatrix}. \tag{31}$$

Since **P** is diagonal, all matrices involved—including **M**—are likewise factorized at all successive stages of approximation corresponding to successive truncations of (22). For instance, for two occupied subbands the successive label assignments for the S and A blocks are

	S		A	
$c=2$	n,n'	μ^s	n,n	μ^A
	1,1	1	2,1	1
	2,2	2	3,2	2,
$c=3$	1,1	1	2,1	1
	2,2	2	3,2	2
	3,1	3		
	4.2	4		
$c=4$	1,1	1	2,1	1
	2,2	2	3,2	2
	3,1	3	4,1	3
	4.2	4	5,2	4.

(32)

Each approximation includes intersubband excitations up to a maximum intersubband difference $(n-n')$ for all occupied subbands and this determines the corresponding matrices $\boldsymbol{\beta}_s$ and $\boldsymbol{\beta}_A$, which may or may not be of the same order. Both the formal mathematical analysis and the practical algorithm hold separately in the S and A subspaces. This allows us to cast the formula (30) in a very transparent way as

$$V(z)=V_e(z)+\sum_{\mu} L_{S,\mu}(z)\sum_{\nu} M_{S,\mu\nu}\int dz' S_{S,\nu}(z')V_e(z') + \sum_{\mu} L_{A,\mu}(z)\sum_{\nu} M_{A,\mu\nu}\int dz' S_{A,\nu}(z')V_e(z'), \tag{33}$$

which displays the obvious fact that a S/A external potential is only screened by the S/A part of ϵ^{-1}, the integrals of the other parity being identically nil.

The importance of the choice of basis was stressed in Ref. 4 from a practical point of view, although the calculations appear to have been carried out in terms of matrix elements of the potentials between the φ_n. The study of ϵ^{-1} closest in spirit to the present analysis was given in Ref. 6 for an idealized model of a superlattice involving only two bands from the start. This was later applied to another idealized model of a quantum well,[7] again with only two subbands. Functions equivalent to our S_μ are introduced and others equivalent to our L_μ can be recognized in some integrands. The present analysis differs in some important respects, namely: (i) the explicit identification of the L_μ and S_μ as long range and short range basis functions, (ii) the consequent formulation in a way which explicitly displays the role of symmetry, (iii) the study of the successive truncations of the spectrum of electronic states and the convergence for $N\to\infty$, and (iv) the establishment of a practical algorithm with which various situations and approximations can be readily studied. The relevance of these remarks will be demonstrated by the examples studied in the following section.

IV. PLASMONS AND SCREENED POTENTIAL

Although the main purpose of this paper is to discuss the inversion of ϵ, it is also interesting to see briefly how can one view the plasmons of the confined quasi-two-dimensional electron gas in the light of the present analysis. For this purpose, it is convenient to write (28) in the alternative form,

$$M=\mathbf{d}\cdot(\mathbf{d}-\mathbf{P})^{-1}\cdot\mathbf{P} \quad (\mathbf{d}\equiv\beta^{-1}), \tag{34}$$

so the matrix **a** to be inverted is $\mathbf{d}-\mathbf{P}$ and the plasmons are the roots of

$$\det|\mathbf{d}-\mathbf{P}|=0. \tag{35}$$

It is easily seen that in the limit $N\to\infty$, the matrix **d** represents the δ function $\delta(z-z')$ in the dual basis $\{L_\mu(z),S_\nu(z')\}$, but even then **d** is *not* diagonal, although there is one $\mathbf{d}_S$ and one $\mathbf{d}_A$ for the case of symmetry just discussed. Since the matrix **p** of (29) is unitary, its determinant cannot vanish and (35) is equivalent to

$$\det|\mathbf{t}|=\Pi_\mu t_{\mu\mu}=0. \tag{36}$$

So the plasmons are in the zeros of the $t_{\mu\mu}$—or in the poles of the $t_{\mu\mu}^{-1}$. This statement requires interpretation.

It was proved in Ref. 5 that (i) the $t_{\mu\mu}$ are real and nonnegative numbers and (ii) the necessary and sufficient condition for the vanishing of $t_{\mu\mu}$ is that the μth column of **a** is a linear combination of the preceding $\mu-1$ columns. Suppose, for instance, a symmetric system with one occupied subband ($k=1$). We start by considering the vanishing of t_{11}. From the formulas derived in Ref. 5 we obtain

$$t_{11}=\left\{|d_{11}-P_1|^2+\sum_{\mu>1}|d_{\mu1}|^2\right\}^{1/2}, \tag{37}$$

which agrees with (i). Clearly this cannot vanish *unless we make the lowest approximation* (cut $c=1$) in which the spectrum is drastically truncated to just one subband. Then, $t_{11}=0$ is equivalent to

$$d_{11}-P_1=0, \tag{38}$$

which is the standard result in the customary ''electric quantum limit'' for the intraband plasmon. The next approximation ($c=2$) leaves the S part as it is and introduces the first

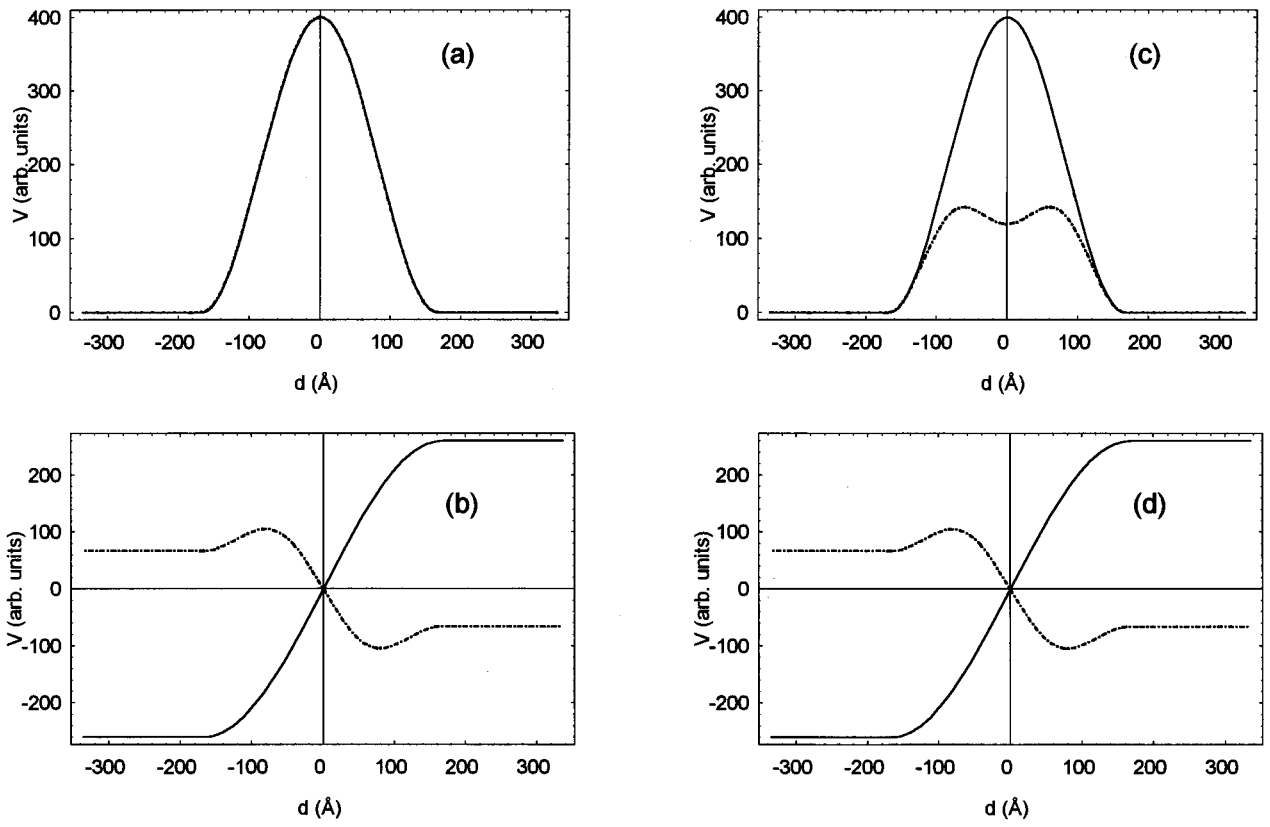

FIG. 1. One occupied subband. In this and all subsequent figures, external (full line) and screened (dotted line) potentials are plotted in arbitrary units against position in Å relative to the center of the well for $Q=0$ and $\omega=290$ cm^{-1}. Here, $L_w=336$ Å; $k=1$; $E_F=15.8$ meV (corresponding to an average particle density $\bar{n}=8.9\times10^{16}$ cm^{-3}); $c=2$ for (a) symmetric (S), (b) antisymmetric (A) and $c=3$ for (c) S, (d) A.

term of the A part, for which $\mu=1$ corresponds to $n=2$, $n'=1$. The same argument holds leading to (37) and (38), only that d_{11} and P_1 now have a different meaning and (38) yields the first interband A plasmon. In the next approximation ($c=3$), the A part remains unchanged, while the S part is increased so that $\mu=2$ corresponds to $n=3$, $n'=1$. Then, $t_{22}=0$ implies that the second column of the matrix $\mathbf{d}-\mathbf{P}$ is proportional to the first one, corresponding to

$$\begin{vmatrix} d_{11}-P_1 & d_{12} \\ d_{21} & d_{22}-P_2 \end{vmatrix} = 0. \tag{39}$$

This yields (a). The intraband plasmon with the first interband corrections coming from interband excitations between $n=1$ and $n=3$—of the same parity—and (b) the first interband S plasmon, and so on. In practice, it is useful to work from the $t_{\mu\mu}$. Since one is in any case interested in obtaining these objects for the inversion of ϵ, one has incidentally a neat way of obtaining the plasmons of the system in a way that displays the role of the symmetry.

Moreover, this also provides an insight into the situation when more than one subband is occupied. Modulation doped quantum wells may easily have two[8] or three[9] populated subbands and δ-doped systems may easily have a few more.[10] Sometimes their individual populations are experimentally determined from Shubnikov–de Haas measurements.[10,11] The lowest plasmon branch—with $\omega\to0$ as $Q\to0$—then splits in k branches if k is the number of occupied subbands.[9] Assume, for instance, $k=2$ and consider the successive approximations described in (32). For $c=2$, the two intraband modes are the roots of a 2×2 determinant, formally like (39), but with a different meaning of the μ labels now in correspondence with (32). This reveals that the main coupling between the (1,1) and (2,2) intraband modes is *not dynamical*, as no interband excitations are involved, but rather a consequence of the spectrum of electronic states and is simply due to the fact that d_{12} and d_{21}, which only depend on Q but not on ω, do not vanish. Of course, dynamical corrections are to be expected, but in higher approximations. Indeed, for $c=3$, the S part leads to a 4×4 secular determinant, which yields (i) the two intraband modes with (3,1) and (4,2) interband corrections, involving excitations between subbands of the same symmetry and (ii) the two symmetric interband modes, all four being coupled, while the A part leads to a 2×2 determinant yielding the two antisymmetric interband modes, (2,1) and (3,2). The analysis can be easily carried out for higher values of k and provides an insight into the meaning of the terms involved in the successive approximations.

Now, our main purpose is to study the screened potential in real space as a consequence of the role played by the various terms entering the dynamical dielectric response of the confined electron gas. In order to demonstrate these effects, it suffices to consider an electron gas confined by an infinite square well. Width and Fermi energy were initially

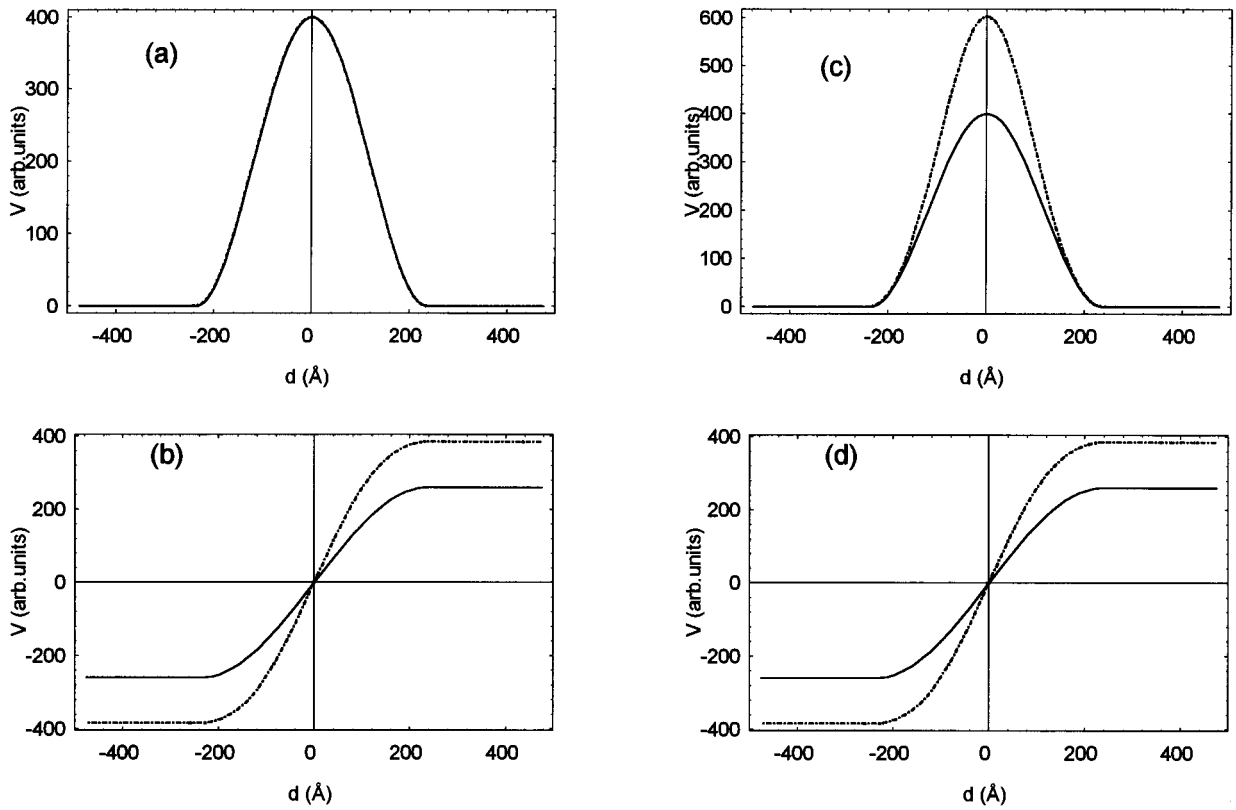

FIG. 2. One occupied subband. $L_w=476$ Å; $k=1$; $E_F=4.9$ meV ($\bar{n}=1.4\times10^{16}$ cm^{-3}); $c=2$ for (a) S, (b) A and $c=3$ for (c) S, (d) A.

chosen so that areal particle densities and energy levels in the lower part of the spectrum are representative of typical experimental GaAs quantum wells. The bare potential to be screened was chosen to imitate, both in frequency and space-dependent amplitude, the first two optical phonon modes at $Q=0$ of the same type of quantum well structures. A detailed study of the screened potential in real space shows that this exhibits a remarkable degree of variety.

In Fig. 1(a), with one occupied subband and $c=2$, which amounts to the usual electric quantum limit *for the S subspace*, a symmetric V_e is not screened at all, because the only intervening polarization term in this approximation is P_1, corresponding to $X_{1,1}$ and it is easily seen that in the limit $Q\rightarrow0$, the $X_{n,n'}$ depend on $n-n'$ and vanish for $n=n'$. Thus, this approximation is totally inadequate to describe the *dynamic* screening of symmetric potentials at low Q, which happens to be typically the range of (long) wavelengths of predominant experimental interest. The antisymmetric potential is screened, because for $c=2$ the polarizability P_2, corresponding to $n=2$, $n'=1$ intervenes in the A part and this does not vanish. It is curious that the screening not only reduces the magnitude of V_e, but also changes its sign. This means that an electron driven by the screened potential of an A mode would, under these circumstances, oscillate completely out of phase with respect to the phonon wave. In order to obtain any screening at all for the S part, we must take at least the next approximation $c=3$ [Fig. 1(b)] in which the S part includes $\mu=2$, corresponding to $n=3$, $n'=1$, for which $P_\mu\neq0$. The A part does not change, since it includes no new terms, but the changes from the usual electric quantum limit approximation ($c=1$) to the next two approximations are drastic: first ($c=2$) an A potential is screened—and may even have its sign changed, as demonstrated in this example and then ($c=3$) an S potential is also screened, in this case quite strongly, even displaying a small depression in the middle region, where the external potential has its maximum. We note the energies involved: $E_1=5$ meV, $E_F=15.8$ meV, $E_2=20$ meV, $E_3=45$ meV, and $\hbar\omega=40$ meV. Some small numerical changes can be appreciated for higher approximations, up to $c=8$, though no new feature appears, and from $c=8$ to $c=12$ there is no appreciable change. The algorithm employed in the calculation, with the choice of basis functions and the arrangement of successive terms as explained above, converges very fast. In this example, $c=3$ constitutes already a quite good approximation for the screening.

Similar considerations apply to the example of Fig. 2, corresponding to a wider well and also with one occupied subband. For $c=2$, only the A potential is screened, for $c=3$, the S potential is also screened and this is a fairly good approximation altogether, with small numerical changes from $c=3$ to $c=8$ and no appreciable change from then on. However, the striking feature is that the external potential is not actually screened but *antiscreened*, that is $|V|>|V_e|$.

Antiscreening and change of sign may combine and result in a remarkable variety of behaviors of the screened potential, depending on the parameters of each case. For the well described in Fig. 3(a) with two occupied subbands, the approximation $c=4$ which, as seen in (32), includes up to $n=5$, is not yet very accurate. The screened S potential then only displays a small depression in the middle. Appreciable changes take place from $c=4$ to $c=8$, involving up to $n=9$,

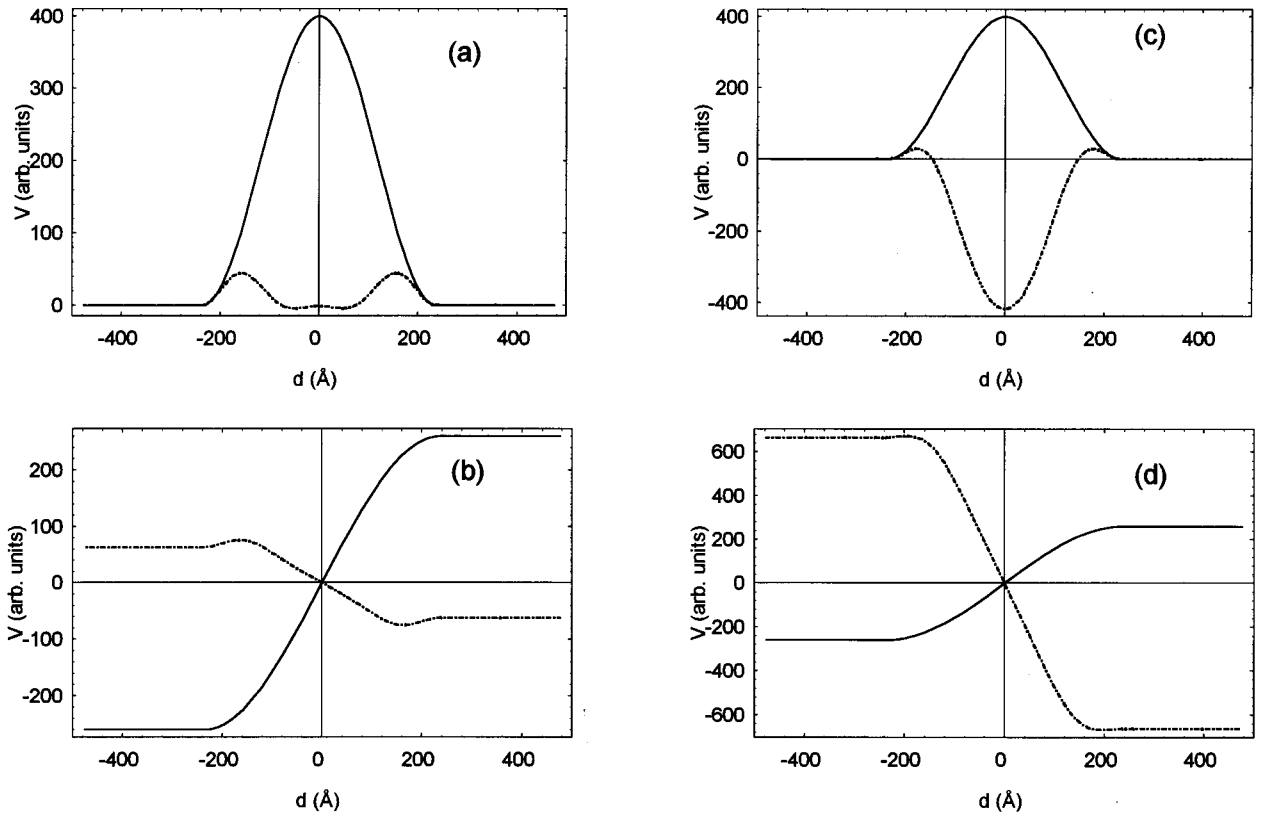

FIG. 3. Two occupied subbands. L_w=476 Å; $k=2$; $c=12$. (a) S and (b) A: E_F=21.5 meV ($\bar{n}=1.1\times10^{17}$ cm^{-3}); (c) S and (d) A: E_F=11 meV ($\bar{n}=5.0\times10^{16}$ cm^{-3}).

when the S potential is strongly screened, the depression in the middle is then significant and a tendency to a change of sign is beginning to emerge near the middle. The results are appreciably the same from then on. The figure shows the actual results for $c=12$. The A potential is also reduced in magnitude, while its sign is changed. Now, these results are for an electron population corresponding to an average concentration of 1.1×10^{17} cm^{-3}. If the Fermi level is lowered so the average concentration decreases to 5×10^{16} cm^{-3} [Fig. 3(b)] then the A potential displays both, change of sign and substantial decrease in magnitude, while the S potential displays change of sign and the beginning of an increase in magnitude in the middle region. Thus, something which emerges in the variety of behaviors which one can encounter is that the tendency to antiscreening—meaning by this the enhancement of the magnitude of the potential, irrespective of the sign—is different for S or A external potentials.

Substantial changes may also occur for very small changes of the parameters. Figure 4 shows the results with $c=10$, when convergence has been more than amply achieved, for the same well with one occupied subband and two values of E_F. In Fig. 4(a), with an average electron concentration of $\bar{n}=3.8\times10^{16}$ cm^{-3}, the S potential behaves qualitatively as in Fig. 3(b) but maintains $|V|<|V_e|$, while the A potential exhibits very strong antiscreening combined with change of sign. On lowering $\bar{n}$ only to 3.1×10^{16} cm^{-3} [Fig. 4(b)] the screened potential for the S case changes only a little, but the screened A potential changes drastically: now it has the same sign as V_e and shows still fairly strong antiscreening, though not nearly as much as in Fig. 4(a).

A study of various different cases indicated that the tendency to antiscreening appears as L_w increases or E_F decreases, that is, when the average electron density $\bar{n}$ altogether decreases. In all cases, antiscreening appears for $\bar{n}$ below 5×10^{16} cm^{-3}. In fact, although this does not appear to be mentioned in the literature on confined quasi-two-dimensional systems, dynamical antiscreening of optical polar mode potentials was found in bulk semiconductors for low electron concentrations.[12] What we find here is the same phenomenon together with further features characteristic of confined quasi-two-dimensional systems. The resulting pattern of behavior is remarkably diverse with a variety that only explicit real space calculations can reveal.

V. FINAL CONSIDERATIONS

We have presented an algorithm for the inversion of ϵ in real space for confined quasi-two-dimensional systems, which (i) is based on a rigorous mathematical analysis of the solution of the integral equation and (ii) converges to the exact answer. The dual basis of long range and short range functions is consistent with the physics of the dielectric response and yields a representation, which displays in a transparent way the role of symmetry and of the empty states, both for the study of the plasmons of the inhomogeneous quasi-two-dimensional electron gas and of the screened potential. Real space calculations of the screened potential can then be readily carried out. From model calculations at frequencies of typical confined polar optical modes in GaAs wells, we find that the dynamical screening shows a great

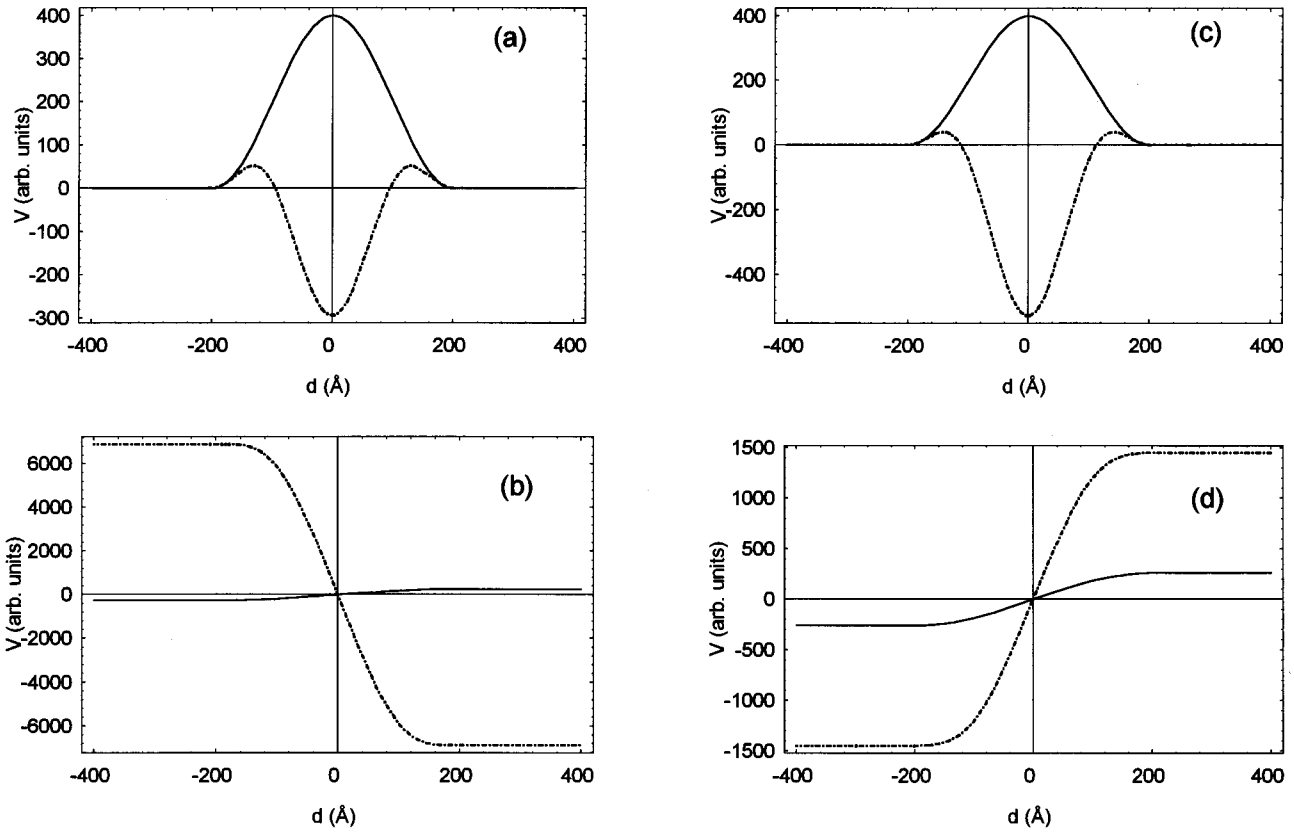

FIG. 4. One occupied subband. L_w=400 Å; k=1; c=10. (a) S and (b) A: E_F=9 meV ($\bar{n}$=3.8×10^{16} cm^{-3}); (c) S and (d) A: E_F=8 meV ($\bar{n}$=3.1×10^{16} cm^{-3}).

variety of behaviors: The sign of the screened potential V may be equal or opposite to that of the external potential V_e, symmetric or antisymmetric potentials may be screened differently, and the magnitude of V_e may be reduced (screening) or enhanced (antiscreening). The antiscreening effect appears at low average electron concentrations, as in early work on bulk polar semiconductors, and appears to have remained unnoticed so far for confined systems. The consequent enhancement of the strength of the electron-phonon interaction is bound to have significant phenomenological implications.

The role of the empty states appears to be also significant and deserves some comment. We have seen in the example of Fig. 3, with two occupied subbands, that for an accurate answer we must include up to $n=9$. Then the energy level E_9 is about 200 meV above the bottom of the well. Now, in Si δ-doped systems in a GaAs matrix, the bottom of the conduction band may be typically about 120–140 meV above the lowest bound state and in GaAs quantum wells with $Al_xGa_{(1-x)}As$ barriers the conduction band well depth may range from about 70 meV (for $x=0.1$) to about 230 meV (for $x=0.3$). Thus, in a calculation for a realistic system, it would not be surprising to find that empty levels, which still have a significant influence, are invading the range of extended states above the well barriers, where the density of states is large. These are only plausible suggestions, which in any case would have to be borne out by a proper calculation. We also note that, in general, one would not expect anything special for static screening, irrespective of how accurate the calculation may be. The striking features demonstrated by the above examples are characteristic of dynamical screening and, among other factors, this may depend quite significantly on the actual values of ω.

The model calculations are presented here to demonstrate that the practical use of the method can equally be carried out with more realistic wave functions $\varphi_n(z)$ corresponding to real systems. Only the screening has been studied here, but an efficient method to obtain $\epsilon^{-1}(z,z')$ opens the way to realistic studies of interesting physical properties like power loss or quasiparticle properties, as indicated in Ref. 6.

Some comments are in order concerning the limitations of the present work. First, the background dielectric constant has been assumed to be the same everywhere, but if desired, the effects of dielectric discontinuities can be included by simply redefining the differential equation (8) in an appropriate manner. The electrostatic response Green function can then be easily obtained by standard matching.[13] This would only remove the factor ϵ_x from the denominator in (10) and change the result of evaluating the long range functions $L_{n,n'}$ of (16), while the problem of the integral equation to be solved is again the same. This is a very simple matter. A substantially more serious question concerns the limitation of the analysis to the random phase approximation. Due to the combined effects of inhomogeneity and nonlocality, together with the multiband structure of the spectrum, to extend the theory beyond the RPA is by no means trivial. In fact, there is still a substantial amount of literature dealing with this type of problems in the RPA. The latest study of normal

modes in a thin metal film is based on this approximation[4] and so is the latest study of a semiconductor superlattice.[14] Work on the extension of the theory beyond the RPA is in progress in our laboratory and the preliminary indications suggest that improved formulas will differ essentially in substantial modifications of the quasi-two-dimensional polarizability elements, while the mathematical structure of the dielectric function to be inverted remains the same. Thus, the corresponding integral equation can be solved by using the method presented here. The real purpose of this paper is not to analyze a concrete physical system, which might require a more advanced formulation, but to present a method which has a firm mathematical foundation and can be useful to invert the dielectric function in real space in a way in which one can easily see the effects of symmetry and of the intersubband excitations. Finally, the case of confinement in one dimension has been discussed here, but the mathematical analysis on which the method is based[5] applies equally to any inhomogeneous system and the same approach can be used to study superlattices, quantum wires, or quantum dots.

ACKNOWLEDGMENTS

The authors are grateful to H. Rodríguez-Coppola for stimulating discussions on the normal modes and power loss of the confined quasi-two-dimensional electron gas. This work was partly supported by the Spanish Dirección General de Investigación Científica y Técnica under Grant No. PB93-1251.

[1] D. G. Mahan, *Many-Particle Physics* (Plenum Press, New York, 1990).

[2] F. Maldague, Surf. Sci. **73**, 296 (1978); T. Ando, A. B. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982); S. Das Sarma and B. A. Mason, Phys. Rev. B **31**, 5536 (1985); G. Gumbs, Solid State Commun. **62**, 365 (1987).

[3] S. Mori and T. Ando, Phys. Rev. B **19**, 6433 (1979); X. L. Lei, J. Phys. C **18**, L593 (1985).

[4] W. H. Backes, F. M. Peeters, F. Brosens, and J. T. Devreese, Phys. Rev. B **45**, 8437 (1992).

[5] J. Fernández-Velicia, F. García-Moliner, and V. R. Velasco, J. Phys. A **28**, 391 (1995).

[6] R. D. King-Smith and J. C. Inkson, Phys. Rev. B **33**, 5489 (1986).

[7] J. A. White and J. C. Inkson, Phys. Scr. **T35**, 116 (1991).

[8] H. L. Störmer, A. Pinczuk, A. C. Gossard, and W. Wiegmann, Appl. Phys. Lett. **38**, 691 (1981); J. Shah, A. Pinczuk, A. C. Gossard, and W. Wiegmann, Phys. Rev. Lett. **54**, 2045 (1985); K. Inoue, H. Sakaki, J. Yoshino, and Y. Yoshikoka, Appl. Phys. Lett. **46**, 973 (1985); K. Inoue and H. Sakaki, Jpn. J. Appl. Phys. **23**, L61 (1984); V. M. S. Gomes, A. S. Chaves, J. R. Leite, and J. M. Worlock, Phys. Rev. B **35**, 3984 (1987); S. Sasa, J. Saito, K. Nanbu, T. Ishiwaka, and S. Hiyamizu, Jpn. J. Appl. Phys. **23**, L573 (1984).

[9] G. Fasol, R. D. King-Smith, D. Richards, U. Ekenberg, N. Mestres, and K. Ploog, Phys. Rev. B **39**, 12 695 (1989).

[10] A. Zrenner, F. Koch, and K. Ploog, Surf. Sci. **196**, 671 (1988).

[11] J. J. Harris, J. M. Lagemaat, S. J. Bettersby, C. M. Hellon, C. T. Foxon, and D. E. Lacklison, Semicond. Sci. Technol. **3**, 773 (1988).

[12] S. Doniach, Proc. Phys. Soc. (London) **73**, 849 (1959); H. Ehrenreich, J. Phys. Chem. Solids **9**, 129 (1959).

[13] F. García-Moliner and V. R. Velasco, *Theory of Single and Multiple Interfaces* (World Scientific, Singapore, 1992).

[14] V. U. Nazarov, Phys. Rev. B **49**, 17 342 (1994).