Nearsightedness of electronic matter in one dimension

E. Prodan
PRISM, Princeton University, Princeton, New Jersey 08544, USA
(Received 9 June 2005; revised manuscript received 17 January 2006; published 14 February 2006)

The concept of nearsightedeness of electronic matter (NEM) was introduced by Kohn in 1996 as the physical principle underlining Yang's electronic structure algorithm of divide and conquer. It describes the fact that, for fixed chemical potential, local electronic properties at a point r, like the density $n(r)$, depend significantly on the external potential v only at nearby points. Beyond a distance R, changes Δv of that potential, no matter how large, have limited effects on local electronic properties, which tend to zero as functions of R. This remains true even if the changes in the external potential completely surround the point r. NEM can be quantitatively characterized by the nearsightedness range $\mathsf{R}(r,\Delta n)$, defined as the smallest distance from r beyond which any change of the external potential produces a density change, at r, smaller than a given Δn. The present paper gives a detailed analysis of NEM for periodic metals and insulators in one dimension and includes sharp, explicit estimates of the nearsightedness range. Since NEM involves arbitrary changes of the external potential, strong, even qualitative changes can occur in the system, such as the quantization of the energy bands or the filling of the insulating gap of an insulator with continuum spectrum. In spite of such drastic changes, we show that Δv has only a limited effect on the density, which can be quantified in terms of simple parameters of the unperturbed system.

DOI: 10.1103/PhysRevB.73.085108 PACS number(s): 71.10.−w, 71.15.−m

I. INTRODUCTION

This paper is based on a preliminary remark by Kohn[1] about a general concept called "nearsightedness of electronic matter" (NEM) and on a recent short report by Prodan and Kohn[2] (PK), which amplified that remark in various aspects but did not include detailed proofs. In the present paper, we select that part of PK dealing with noninteracting one-dimensional (1D) electrons and provide a full discussion, including detailed proofs. Future publications will amplify other sections of PK.

By "electronic matter" we understand a system of many electrons with significant wave-function overlap, in equilibrium under the action of a given external potential $v(r)$, at a given and fixed temperature T. We shall consider the change of a local electronic property, like the electron density $n(r)$, under the action of an arbitrarily strong potential perturbation $w(r')$, which vanishes inside a specified sphere $|r-r'|=\mathsf{R}$. Note that we allow situations when the perturbation completely surrounds the point r. NEM states that the resulting density change at r, $\Delta n(r,\mathsf{R})$, is bounded by a function $\overline{\Delta n}(r,\mathsf{R})$,

$$\Delta n(r,\mathsf{R}) \leqslant \overline{\Delta n}(r,\mathsf{R}), \tag{1}$$

independent of the amplitude or shape of $w(r')$, and that

$$\overline{\Delta n}(r,\mathsf{R}) \to 0 \quad \text{monotonically as } \mathsf{R} \to \infty, \tag{2}$$

provided only that the chemical potential μ is held fixed.

The essence of NEM is contained in Eq. (1). Although it may not look very special, consider our perturbing potential $w(r')$, confined outside the sphere of radius R. Common sense says that, as we increase its strength, we need to increase the radius R if we want to maintain its effect at the center of the sphere below a certain level. Common sense also says that we need to increase R to infinity as we make $w(r')$ stronger and stronger. In reality, this is not so: if the chemical potential is kept fixed, these effects will saturate and, in fact, no matter what $w(r')$ we put outside the sphere of radius R, they cannot exceed a certain upper bound, which we will determine in this paper.

For a given Δn, we can solve for R in $\overline{\Delta n}(r,\mathsf{R})=\Delta n$ and define the nearsightedness range $\mathsf{R}(r,\Delta n)$. The significance of $\mathsf{R}(r,\Delta n)$ is the following: no perturbation beyond $\mathsf{R}(r,\Delta n)$, of arbitrary shape and amplitude, can produce a density change at r larger than Δn. $\mathsf{R}(r,\Delta n)$ provides a simple, quantitative measure of nearsightedness.

The above formulation of NEM often reminds people of Thomas-Fermi screening, sometimes even when we discuss insulators. However, let us point out a few facts. If one puts a charge inside the uniform charged electron gas and calculates the density response, one will find that the Thomas-Fermi exponential screening is valid only very near the impurity. Further away from the impurity, one will see Friedel oscillations, decaying as an inverse power law.[3] These oscillations are not negligible. In fact, the Friedel oscillations were observed experimentally, though not directly, one year after their theoretical prediction.[4] However, this was not realized until much later, when Kohn made the connection between the two results.[5,6] He showed that there is a big discrepancy between the prediction of the Thomas-Fermi screening theory and these experimental results and that a self-consistent calculation along Friedel's lines of the density change due to impurities in copper brings the theory and experiment to almost perfect agreement. The picture that emerged was that, in the asymptotic region, the screening only renormalizes the amplitude of the Friedel oscillations. The whole issue was considered very important at that time, because it clearly demonstrated the existence of a sharp Fermi surface in real metals. We also like to mention one very well-known fact in surface physics, where self-

consistent calculations of metallic surfaces showed that the effective potential goes to the bulk value extremely fast, typically within one or two layers. However, the density oscillations extend much further into the bulk and they can be viewed as the Friedel oscillations generated by the screened surface potential. With these being said, we hope that the reader will dissociate, right from the begenning, NEM from the Thomas-Fermi exponential screening and the nearsightedness range from the Thomas-Fermi screening length.

Quantum gases display nonlocal density responses to local perturbations because of two factors. First, the effect of any local perturbation propagates to further distances through interparticle interactions. This effect can be regarded as classical, since it manifests, in the same way, in classical gases. Second, there is a purely quantum effect that stems directly from the uncertainty principle. This paper is concerned with this purely quantum effect, so it neglects the interparticle interation effects entirely.

The fact that NEM exists for noninteracting systems is extremely important. To understand why, let us go back in time and recall that, at the beginning of the electronic structure calculations, when the exact diagonalization was the method of choice, people were facing the so-called "exponential wall" when trying to extend the calculations to larger systems: because the number of operations in such calculations scales exponentially with the number of atoms N, their applicability was and is still limited to systems containg a few tens of atoms. Density functional theory[7,8] (DFT) provided a powerful alternative: because the number of operations in DFT calculations scales as N^3, we can now solve the electronic structure for systems containing hundreds of atoms. However, electronic structure calculations for biological and nano systems, or for extremely complex materials, involve thousands of atoms. At this scale, we start feeling the "N^3 wall." *Ab initio* quantum calculations for such complex systems will require a new generation of DFT algorithms, scaling linearly with the number of atoms. It is now generally accepted that NEM is the physical basis for these algorithms.[9]

The quest for linearly scaling algorithms was initiated by Yang, who was the first to argue that $O(N)$ algorithms are possible.[10] The algorithm proposed by Yang is known by the name of divide and conquer (DC). There are now several reviews on linear scaling electronic structure calculations. We will mention here the one by Goedecker[11] and the one by Wu and Jayanthi,[12] which, at the time of their publication, gave an exhaustive discussion of $O(N)$ methodologies. If we examine carefully these methodologies, they are all based on the same original idea, namely, gluing together calculations done for smaller systems. What is different is the representation and the way the size of these smaller systems (the truncation) is determined. For example, the real space approaches will use the decay of the density matrix while the localized basis set approaches will use the overlap of these functions to judge how large these subsystems should be.

Let us focus on the original implementation given by Yang.[10] Consider a self-consistent, finite-temperature DFT iteration process, for a large quantum system. Each iterative step consists in calculating the density of a noninteracting electron gas in equilibrium under a given effective potential (known from the previous iteration). In traditional approaches, this requires a number of operations that scales as N^3. The DC algorithm, if it works as it is supposed to, requires a number of operations that scales linearly with N. It goes like this. The large system is divided into nonoverlaping subregions, which are then surrounded with buffer zones. The global chemical potential μ is fixed and the orbitals are calculated and populated according to the Fermi-Dirac distribution at μ and T, for each subregion+buffer zone. The density in the buffer zones is discarded so, at this point, one has calculated a density for each subregion and, by putting together all these subdensities, one can construct the global density. The charge neutrality condition is then checked for this global density and the chemical potential is adjusted, if necessary. Note that charge neutrality must be satisfied by the entire system, not by each subsystem. In this way we have completed the DFT iteration step in a number of operations that scales linearly with the size of the system. Now, the question is, how accurate is this algorithm? In fact, the most important question is, can we obtain arbitrarily small accuracy with this algorithm? To answer, we need to compare the density calculated for a subregion+buffer zone and the density calculated for the entire system at once, at the same chemical potential μ. Now one can see why NEM can be regarded as the basis for this algorithm: the artificial termination at the outer boundary of the buffer zone, no matter how it is done, represents the change in the effective potential in our formulation of NEM. For example, such changes of the effective potential occur when one calculates the density matrix and ignores the points outside a subregion, or when one calculates the density and ignores the elements of a localized basis set that are centered outside a subregion. Since we have no control on how the effective potential is modified by such truncations, Eq. (1) is paramount: it tells us that the effects of any artificial termination cannot exceed an upper bound. This upper bound is an intrinsic characteristic of the system and is independent of the method of termination. Equation (2) tells us that, if we take the buffer zones large enough, the difference between the subdensity and the real density can be made smaller than any desired accuracy.

Examples and estimates of the width of the buffer zones, together with a discussion of how to optimize this algorithm in 1D, 2D, and 3D and how the CPU time scales with the desired accuracy can be found in Ref. 2. We also wish to mention that DC has been recently implemented in systems containing as many as 65 000 atoms and shown that it can generate electronic structures of the same quality as a traditional approach will do for a system of, let us say, ten atoms.[13] The tests performed in this numerical work agree qualitatively with our theoretical predictions.

There is another important issue related to DC. The ground energy in DFT is given by

$$E = \sum_j \epsilon_j + E_{xc}[n] - \int v_{xc}(r)n(r) - \frac{1}{2}\int \frac{n(r)n(r')}{|r-r'|}. \quad (3)$$

All the above terms can be calculated directly from the density, except the first one. However, this term is just the integral of the energy density,

$$\epsilon(r) = \sum_{\epsilon_i \leqslant \mu} \epsilon_i |\psi_i(r)|^2, \tag{4}$$

and we will show that $\epsilon(r)$ is also nearsighted. As a consequence, within the DC algorithm, $\epsilon(r)$ can be calculated with arbitrary precision, like the density.

The goal of this paper is twofold. We want to prove NEM, i.e., Eqs. (1) and (2), for a simple system, which is the 1D noninteracting, spin-$\frac{1}{2}$ fermion, in periodic potentials at $T=0^+$. We also want to show that one can obtain exact estimates of the nearsightedness range, which is extremely important for DC. The case of 1D noninteracting electrons is important for several reasons. In spite of its simplicity, it captures all the important aspects of nearsightedness. This allows for a thorough investigation of NEM, while keeping the technical aspects at a reasonable level. The 1D noninteracting case is relevant for linear molecular chains when treated within DFT.

II. THE STRATEGY

We shall first develop general tools that will allow us to compute, for arbitrary perturbations, the asymptotic behavior of the density change at $T=0^+$,

$$\Delta n(x) = 2 \sum_{\epsilon_i \leqslant \mu} |\psi_i(x)|^2 - 2 \sum_{\epsilon_i^0 \leqslant \mu} |\psi_i^0(x)|^2, \tag{5}$$

where $\psi_i^0(x), \epsilon_i^0$ and $\psi_i(x), \epsilon_i$ are the wave functions and the corresponding energies of the unperturbed and perturbed systems, respectively. The factor 2 in front of the sums comes from the spin. The above expression is not very useful when dealing with the asymptotic behavior of $\Delta n(x)$. Instead, we will work with an integral representation. Why an integral representation? For answer, we point to the theory of special functions, where the functions are most often defined and introduced as infinite series but, with no exception, their asymptotic behavior is derived from equivalent integral representations.

We can obtain an integral representation of $\Delta n(x)$ by using Green's functions. Indeed, if $G_E^0 \equiv (E-H_0)^{-1}$ and $G_E \equiv (E-H)^{-1}$ denote the Green's functions of the unperturbed and perturbed systems, respectively, then

$$\Delta n(x) = \frac{1}{\pi} \int_{\mathcal{C}} (G_E - G_E^0)(x,x) dE, \tag{6}$$

where $\mathcal{C}$ is a contour in the complex energy plane, surrounding the occupied states. This can be seen from the eigenfunction expansions of G_E^0 and G_E and the residue theorem. Now, the eigenfunction expansions of G_E^0 and G_E are, again, not very useful. Instead, we will use the following compact representation:

$$G_E(x,x') = \frac{\psi_<(x_<)\psi_>(x_>)}{W(\psi_<,\psi_>)}, \tag{7}$$

where $x_< = \min(x,x')$ and $x_> = \max(x,x')$; $\psi_<(x)$ and $\psi_>(x)$ are the solutions of the Schrödinger equation at energy E, satisfying the boundary condition to the left and right, respectively, and $W(\psi_<,\psi_>)$ is the Wronskian of the two solutions. For infinite systems, the case considered in this paper, $\psi_<(x)$ and $\psi_>(x)$ are the solutions decaying at $\pm\infty$, respectively. We will always assume that E does not belong to the energy spectrum. When the contour $\mathcal{C}$ intersects the energy spectrum, such as for the case of metals, G_E will have a discontinuity at the point of intersection. Strictly speaking, this point must be excluded from $\mathcal{C}$, which does not change the result of the integration. For all the other points of $\mathcal{C}$, G_E is uniquely defined and given by Eq. (7). Later, we will use the reflection and transmission coefficients to construct extremely simple and compact expressions of the Green's functions [see Eqs. (15), (22), and (61)].

The last step of our strategy will be to identify the special point in the complex energy plane that determines the asymptotic behavior of the integral Eq. (6).

This strategy will require from us to go into the complex energy plane. We will make use of the analytic structure of the Bloch functions and band energies derived in Ref. 14, which is briefly discussed in the next section. These analytic structure results can be generalized to linear molecular chains and even to 3D crystals.[15] Also, the above expression for the Green's function, Eq. (7), can be generalized to linear molecular chains,[15] or to 3D crystals.[16] In fact, the entire strategy can be applied in 2D and 3D, as was already shown in Ref. 2.

III. THE UNPERTURBED SYSTEM

Throughout this paper, $v(x)$ will be taken as a periodic, inversion-symmetric potential. Following is a brief discussion, largely taken from Ref. 14, of the periodic Schrödinger equation.

The solutions of the periodic Schrödinger equation,

$$[-d^2/dx^2 + v(x)]\psi = E\psi, \quad v(x+b) = v(x), \tag{8}$$

are the well-known Bloch functions ψ_k (k is the wave vector) which will be normalized as in Ref. 14. Their fundamental property is $\psi_k(x+b) = e^{ikb}\psi_k(x)$. When dealing with complex values of k, it is much more convenient to work with the variable $\lambda = e^{ikb}$, instead of k. Thus, from now on, we will index the Bloch functions by λ. Their fundamental property becomes

$$\psi_\lambda(x+b) = \lambda \psi_\lambda(x). \tag{9}$$

The parameter λ relates to the energy of ψ_λ through the following equation:

$$\lambda^2 - 2\mu(E)\lambda + 1 = 0, \tag{10}$$

with $\mu(E)$ the Kramers' function.[17] By examining the fundamental property Eq. (9), one can see that the physical states correspond to the case $|\lambda| = 1$ (λ on the unit circle), otherwise $\psi_\lambda(x)$ explodes to either $x=\infty$ or $x=-\infty$. For λ on the unit circle, the solutions behave like waves; thus it is appropriate to use the term Bloch waves. When discussing arbitrary values of λ, however, it is more appropriate to use the term Bloch functions.

The energy spectrum consists of energy bands, indexed here by $n=1,2,\ldots$, which are separated by energy gaps. The

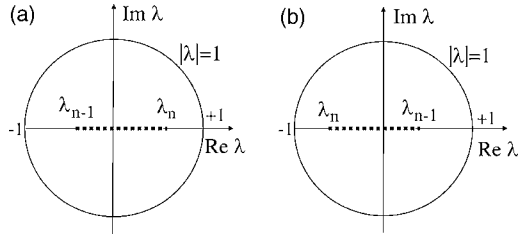

FIG. 1. The analytic domain of $E_n(\lambda)$ and $\psi_{n,\lambda}(x)$ consists of the unit disk except for a branch cut, represented by the dotted line. (a) refers to n even and (b) to n odd.

energy bands can be computed by solving for E in Eq. (10) for all λ on the unit circle. Due to the symmetry $\lambda \to 1/\lambda$ in Eq. (10), we can and shall restrict λ to $|\lambda| \leq 1$, and view $\psi_\lambda(x)$ and $\psi_{1/\lambda}(x)$ as two independent wave functions. For $|\lambda| < 1$, it follows from the fundamental property Eq. (9) that $\psi_\lambda(x)$ decays to zero as $x \to \infty$ and $\psi_{1/\lambda}(x)$ decays to zero as $x \to -\infty$.

If we restrict λ to the unit disk, E uniquely determines λ. The opposite is not true; instead $E(\lambda)$ is a multivalued complex function, with branch points of order 1 at $\lambda_1, \lambda_2, \ldots, \lambda_n$ which are located on the real axis, strictly inside the unit circle and have alternating signs: $0 < (-1)^n \lambda_n < 1$. Their corresponding energies $\tilde{E}_n \equiv E(\lambda_n)$ are also real; $\tilde{E}_n$ is located in the nth gap. Estimates of λ_n in the small-gap and tight-binding limits are given in Appendix A. $E(\lambda)$ can be represented on a Riemann surface with one sheet per band. The nth sheet can be taken as the entire unit disk, except for a cut extending from λ_{n-1} to λ_n, as shown in Fig. 1. Near a branch point, $E(\lambda)$ behaves as the square root,[14]

$$E(\lambda) = \tilde{E}_n + 2\alpha_n(\lambda/\lambda_n - 1)^{1/2} + \cdots . \tag{11}$$

The function $E(\lambda)$ on the nth Riemann sheet will be denoted by $E_n(\lambda)$.

The integral in Eq. (6) will be mapped into the complex λ plane by changing the variable from E to λ. Thus, it will be important to understand how the contour $\mathcal{C}$ looks in this plane. It is more easy to understand how a given contour γ in the λ plane is mapped in the complex energy plane, i.e., to construct the points $E(\lambda)$ when λ sweeps through the points of γ.

The mapping $E_n(\lambda)$, from the unit disk to the complex E plane, is generic in one dimension, in the sense that it is qualitatively independent of the periodic potential. This subject is not discussed in Ref. 14, but it follows from this work. The general picture is as follows. $E_n(\lambda)$ maps the unit disk into a domain $\mathcal{D}_n$ (see Fig. 2). The domains $\mathcal{D}_n$, $n=1,2,\ldots$, are disjoint (with the exception of a possible common boundary) and, all together, they cover the entire complex E plane. Now let us consider several contours. In Fig. 2, we used the same line style for the contour and its image. The unit circle is mapped into a contour that surrounds the nth band, infinitely close. The energy band is represented in Fig. 2 by a thick line. The points $\lambda = \pm 1$ always correspond to the band edges. The origin is mapped at infinity. More exactly, if λ approaches the origin from above (below) the branch cut,

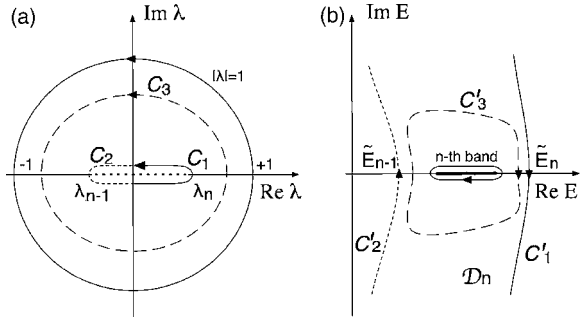

FIG. 2. Illustration of how $E_n(\lambda)$ maps (a) the complex λ plane into (b) the complex E plane (n even).

$E_n(\lambda) \to \pm i\infty$, respectively. We now deform the unit circle until it surrounds the branch cut, infinitely tied. It is instructive to consider an intermediate contour, like $\mathcal{C}_3$ of Fig. 2(a), which is (qualitatively) mapped by $E_n(\lambda)$ into $\mathcal{C}_3'$. As the contour $\mathcal{C}_3$ is continuously shrunk, some points of its image $\mathcal{C}_3'$ move toward $\pm i\infty$, and the limit contour, i.e., the closed contour formed by $\mathcal{C}_1 + \mathcal{C}_2$ in Fig. 2(a), is mapped into the contour $\mathcal{C}_1' + \mathcal{C}_2'$. More exactly, $\mathcal{C}_1$ (the continuous line) is mapped into $\mathcal{C}_1'$ and $\mathcal{C}_2$ (the dashed line) into $\mathcal{C}_2'$. Figure 2(b) displays only finite sectors of $\mathcal{C}_{1,2}'$, which extend from $-i\infty$ to $+i\infty$. The domain $\mathcal{D}_n$ mentioned above lies between the curves $\mathcal{C}_1'$ and $\mathcal{C}_2'$.

The segment from -1 to λ_{n-1} is mapped on the real axis, from the lower edge of the nth band down to $\tilde{E}_{n-1}$. The segment from λ_n to $+1$ is also mapped on the real axis, from $\tilde{E}_n$ down to the upper edge of the nth band. This completes our analysis of the $E_n(\lambda)$ mapping. From the above information, one should be able to construct, qualitatively, the image of any other contour.

We now discuss the analytic structure of the Bloch functions: $\psi_\lambda(x)$ and $\psi_{1/\lambda}(x)$ are multivalued analytic functions of λ, with branch points of order 3 at λ_n. $\psi_\lambda(x)$ will be denoted by $\psi_{n,\lambda}(x)$, for λ on the nth Riemann sheet. Both functions diverge at the branch points as[14]

$$\psi_{n,\lambda}(x) = \frac{u_n(x)e^{-q_n x}}{(\lambda/\lambda_n - 1)^{1/4}} + \cdots \tag{12}$$

and

$$\psi_{n,1/\lambda}(x) = \frac{u_n'(x)e^{q_n x}}{(\lambda/\lambda_n - 1)^{1/4}} + \cdots , \tag{13}$$

where $u_n(x)$ and $u_n'(x)$ are periodic (antiperiodic) functions for n even (odd) and q_n is defined by $|\lambda_n| = e^{-q_n b}$. Using the explicit expressions given in Ref. 14, we can easily calculate the Wronskian of the two independent Bloch functions, which is given by

$$W(\psi_{n,1/\lambda}, \psi_{n,\lambda}) = -\frac{b\lambda}{2\pi}\frac{dE_n(\lambda)}{d\lambda}. \tag{14}$$

Consequently, the Green's function G_E^0 satisfies the identity

$$G_E^0(x,x')\frac{dE}{\pi i} = 2i\psi_{n,1/\lambda}(x_<)\psi_{n,\lambda}(x_>)\frac{d\lambda}{b\lambda}. \tag{15}$$

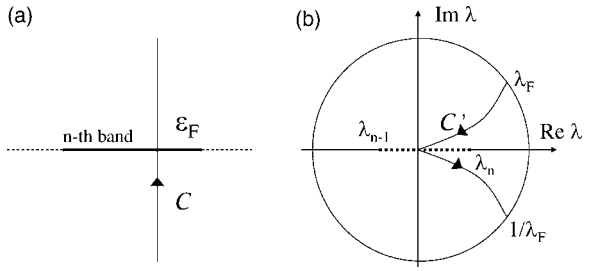

FIG. 3. The contour of integration for metals in (a) complex E plane and (b) complex λ plane. The contour in (a) extends from $-i\infty$ to $+i\infty$.

IV. THE EFFECT OF PERTURBATIONS

A. One-sided perturbations

We consider here perturbations that are either to the left or to the right of the point x, where we measure the density change $\Delta n(x)$. Let us assume that w is to the left of x. For convenience, we choose the origin of x at the right edge of w, so that w is confined in the interval $[-L,0]$, with L arbitrarily large but finite. We calculate the particle and energy density changes at $x>0$.

As already mentioned, the density change $\Delta n(x)$ is given by

$$\Delta n(x) = \frac{1}{\pi i}\int_{\mathcal{C}} [G_E(x,x) - G_E^0(x,x)]dE, \tag{16}$$

where G_E^0 and G_E are the unperturbed and perturbed Green's functions, respectively, and $\mathcal{C}$ is a contour in the complex energy plane, surrounding the eigenvalues below ϵ_F [see, for example, Fig. 3(a)]. Similarly, the change of the energy density is

$$\Delta\epsilon(x) = \frac{1}{\pi i}\int_{\mathcal{C}} E[G_E(x,x) - G_E^0(x,x)]dE. \tag{17}$$

We can focus on $\Delta n(x)$ and give only the final results for $\Delta\epsilon(x)$.

We construct the perturbed Green's function from two independent solutions of the Schrödinger equation:

$$[-d^2/dx^2 + v(x) + w(x)]\psi(x) = E\psi(x). \tag{18}$$

Since $w(x)$ is zero outside the interval $[-L,0]$, any solution of the above equation, outside this interval, is a linear combination of ψ_λ and $\psi_{1/\lambda}$. As already mentioned, we need the solutions decaying to $\mp\infty$, which can be conveniently written in terms of the reflection and transmission coefficients:

$$\psi_{n,\lambda}^{<}(x) = \begin{cases} T_n(\lambda)\psi_{n,1/\lambda}(x), & x<-L, \\ \psi_{n,1/\lambda}(x) + R_n^+(\lambda)\psi_{n,\lambda}(x), & x>0 \end{cases} \tag{19}$$

and

$$\psi_{n,\lambda}^{>}(x) = \begin{cases} \psi_{n,\lambda}(x) + R_n^-(\lambda)\psi_{n,1/\lambda}(x), & x<-L \\ T_n(\lambda)\psi_{n,\lambda}(x), & x>0, \end{cases} \tag{20}$$

where E was taken to be in $\mathcal{D}_n$ (see Fig. 2). One can use the fundamental property of the Bloch functions Eq. (9) to check the behavior of these functions at $\pm\infty$. The Wronskian of the two independent solutions is simply $W(\psi_{n,\lambda}^{<},\psi_{n,\lambda}^{>}) = T_n(\lambda)W(\psi_{n,1/\lambda},\psi_{n,\lambda})$, or

$$W(\psi_{n,\lambda}^{<},\psi_{n,\lambda}^{>}) = -\frac{b}{2\pi}\lambda T_n(\lambda)\frac{dE_n(\lambda)}{d\lambda}, \tag{21}$$

leading to the following useful identity, for $x>0$:

$$(G_E - G_E^0)(x,x)\frac{dE}{\pi i} = 2iR_n^+(\lambda)\psi_{n,\lambda}(x)^2\frac{d\lambda}{b\lambda}. \tag{22}$$

This identity, together with Eqs. (16) and (17), shows that, for $x>0$, $\Delta n(x)$ and $\Delta\epsilon(x)$ are completely determined by the unperturbed wave functions and reflection coefficient, in a simple and universal way.

Since Eq. (22) will be integrated in the complex plane, we need to understand the analytic structure of the reflection coefficient. We can derive the analytic structure of the reflection coefficient from Eq. (22): $R_n^+(\lambda)$ has branch points of order 1 at λ_{n-1} and λ_n. If we go around these branch points, R_n^+ becomes R_{n-1}^+ and R_{n+1}^+, respectively. In other words, R_n^+ are different branches of a multivalued function $R^+(\lambda)$. Near the branch points,

$$R^+(\lambda) = R^+(\lambda_n) + r_n^+(\lambda/\lambda_n - 1)^{1/2} + \cdots . \tag{23}$$

The poles of $R^+(\lambda)$, if any, are mapped by $E(\lambda)$ into the poles of G_E, i.e., the energies of the bound states. For the case considered in this section, the bound states, if any, are always located in the gaps and consequently the poles of $R^+(\lambda)$ are always located on the real axis and away from the branch cuts.

The following is a simple generalization, from the uniform to periodic potentials, of some well-known facts. By evaluating the Wronskian of $\psi_\lambda^{<}(x)$ and $\psi_{1/\lambda}^{<}(x)$ for $x<-L$ and for $x>0$ and equating the two results, one can derive the following identity:

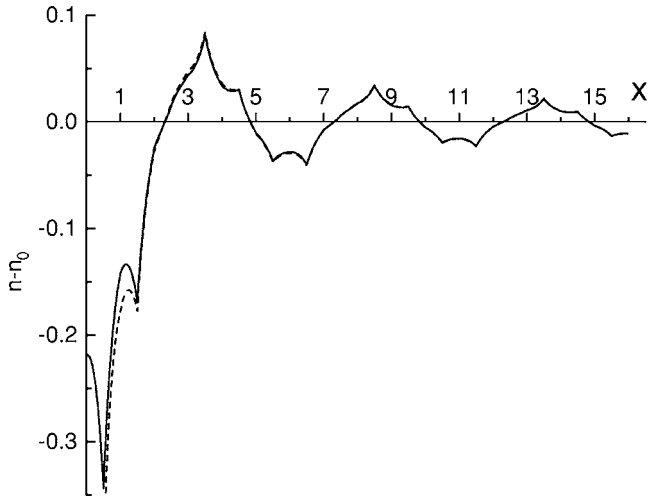

FIG. 4. The exact (solid) and asymptotic (dashed) $\Delta n(x)$ for the model of Eq. (32) ($v_0=-2$, $V_0=10$, $b=1$, first band 20% filled).

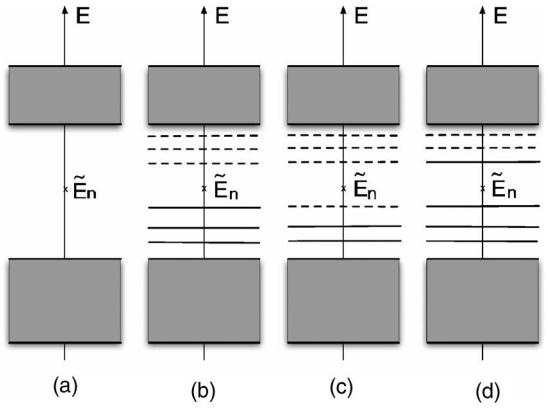

FIG. 5. (a) No bound states are present in the insulating gap. (b) Bound states are present, but all the states below $\widetilde{E}_n$ are occupied (solid lines) and all the states above $\widetilde{E}_n$ are unoccupied (dashed lines). (c) States below $\widetilde{E}_n$ are unoccupied. (d) States above $\widetilde{E}_n$ are occupied.

$$T(\lambda)T(1/\lambda) + R^+(\lambda)R^+(1/\lambda) = 1. \tag{24}$$

For $|\lambda|=1$, this identity becomes $|T(\lambda)|^2+|R^+(\lambda)|^2=1$, showing that $|R^+(\lambda)|\leqslant 1$ for all λ on the unit circle. A similar conclusion holds for $R^-(\lambda)$. In other words, as in the uniform case, the reflection and transmission coefficients cannot exceed unity.

1. Metals

We assume the nth band partially occupied and choose $\mathcal{C}$ in Eq. (16) as in Fig. 3, where λ_F is defined by $\epsilon_F=E_n(\lambda_F)$. We write

$$x = y + mb, \tag{25}$$

with y restricted to the first unit cell. Mapping into the λ plane by using Eq. (22) and recalling the fundamental property of the Bloch functions, for $x>0$, we obtain

$$\Delta n(x) = \frac{2i}{b}\int_{C'} R_n^+(\lambda)\psi_{n,\lambda}(y)^2\lambda^{2m-1}d\lambda. \tag{26}$$

We write $\lambda^{2m-1}=(2m)^{-1}d\lambda^{2m}/d\lambda$ and integrate by parts, which gives

$$\Delta n(x) = 2\,\mathrm{Im}\,\frac{R_n^+(\lambda_F)\psi_{n,\lambda_F}(x)^2}{mb} - \frac{2i}{b}\int_{C'}\frac{\lambda^{2m}}{2m}\frac{d}{d\lambda}[R_n^+(\lambda)\psi_{n,\lambda}(y)^2]d\lambda. \tag{27}$$

The integral is of order $1/x^2$, as can be seen from another integration by parts. We can conclude

$$\Delta n(x) \to \frac{2}{x}\,\mathrm{Im}[R_n^+(\lambda_F)\psi_{n,\lambda_F}(x)^2] \tag{28}$$

for large x. Similarly,

$$\Delta\epsilon(x) \to \frac{2\epsilon_F}{x}\,\mathrm{Im}[R_n^+(\lambda_F)\psi_{n,\lambda_F}(x)^2]. \tag{29}$$

Since $|R_n^+(\lambda_F)|\leqslant 1$, the amplitudes of $\Delta n(x)$ and $\Delta\epsilon(x)$ cannot exceed, in the asymptotic limit, the upper bounds

$$\overline{\Delta n}(x) \to \frac{2}{x}|\psi_{n,\lambda_F}(x)|^2 \tag{30}$$

and

$$\overline{\Delta\epsilon}(x) \to \epsilon_F\overline{\Delta n}(x), \tag{31}$$

independent of the shape and amplitude of $w(x)$. This proves NEM for metals and one-sided perturbations.

A comparison between Eq. (28) and an exact calculation of $\Delta n(x)$ for the perturbed Kronig-Penney model[18]

$$v_t(x) = \begin{cases} v_0\sum_{l=-\infty}^{\infty}\delta\left(x+\frac{b}{2}-lb\right), & x>0 \\ V_0, & x\leqslant 0, \end{cases} \tag{32}$$

is shown in Fig. 4. Note that the asymptotic regime starts from about two lattice constants away from the perturbation.

2. Insulators

We assume the first n bands completely filled. For insulators, the calculations are more involved since the perturbing potential $w(x)$ may generate bound states in the insulating gap ($\equiv$ the gap above the nth band). There can be a discrete or a continuum set of states inside the insulating gap. When the set is discrete, there are four distinct possibilities, as shown in Fig. 5. Let us analyze these cases first.

No bound states in the insulating gap. We can take $\mathcal{C}_1'$ [with opposite orientation; see Fig. 2(b)] as the contour of integration in Eq. (16). Mapping into the complex λ plane and using the fundamental property of the Bloch functions gives

$$\Delta n(x) = \lambda_n^{2m}\frac{2i}{b}\int_{\mathcal{C}_1} R_n^+(\lambda)\psi_{n,\lambda}(y)^2\left(\frac{\lambda}{\lambda_n}\right)^{2m-1}\frac{d\lambda}{\lambda_n}, \tag{33}$$

with $\mathcal{C}_1$ shown in Fig. 2(a) and y and m defined in Eq. (25). The integrand diverges at λ_n as $(\lambda-\lambda_n)^{-1/2}$ but this singularity is integrable. Away from the branch point, the integrand is finite and $(\lambda/\lambda_n)^{2m}$ becomes small as we follow the contour $\mathcal{C}_1$ toward $\lambda=0$. Thus, for large m, the main contribution to the integral comes from the region in the immediate vicinity of the branch point. Expanding the integrand near λ_n and keeping the leading term,[19] we find

$$\Delta n(x) \to R_n^+(\lambda_n)\frac{2}{b}\int_{\mathcal{C}_1}\frac{i(\lambda/\lambda_n)^{2m-1}}{\sqrt{\lambda/\lambda_n-1}}\frac{d\lambda}{\lambda_n}u_n(x)^2e^{-2q_nx}. \tag{34}$$

The integral is equal to $-2B(2m,1/2)$ (B is the Beta function) and behaves asymptotically as $-\sqrt{2\pi/m}$. We conclude that

$$\Delta n(x) \to -2R_n^+(\lambda_n)\left(\frac{2\pi}{xb}\right)^{1/2}u_n(x)^2e^{-2q_nx}. \tag{35}$$

Similarly,

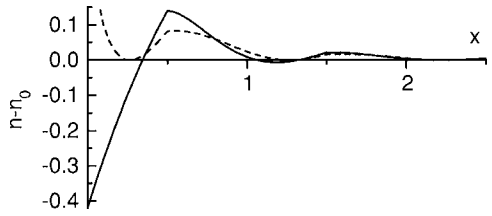

FIG. 6. The exact (solid line) and asymptotic (dashed line) $\Delta n(x)$ for the perturbed Kronig-Penney model Eq. (32) ($v_0=-3$, $b=1$, $V_0=50$, first band completely filled).

$$\Delta\epsilon(x) \to -2R_n^+(\lambda_n)\tilde{E}_n\left(\frac{2\pi}{xb}\right)^{1/2} u_n(x)^2 e^{-2q_n x}. \tag{36}$$

An implementation of Eq. (35) in the perturbed Kronig-Penney model described in Eq. (32) is given in Fig. 6. Note again that the asymptotic regime starts from one or two lattice sites from the perturbation.

To end the proof of nearsightedness, we need to show that $|R^+(\lambda_n)|$ cannot exceed an upper bound. Since the reflection coefficient is not evaluated on the unit circle, the inequality $|R^+(\lambda)| \leqslant 1$ is no longer guaranteed. However, if $h_\lambda^<$ and h_λ denote the logarithmic derivatives at $x=0$ of $\psi_\lambda^<(x)$ defined in Eq. (19) and of $\psi_\lambda(x)$, respectively, then

$$R^+(\lambda) = -\frac{h_\lambda^< - h_{1/\lambda}}{h_\lambda^< - h_\lambda}\frac{\psi_{1/\lambda}(0)}{\psi_\lambda(0)}. \tag{37}$$

As in Ref. 14, we choose the phase of the Bloch functions such that $\psi_\lambda(0)=\psi_{1/\lambda}(0)$, so we can eliminate the last factor in the above expression. For E in the insulating gap, $\psi_\lambda^<(x)$, $\psi_\lambda(x)$, and $\psi_{1/\lambda}(x)$ are real functions and since

$$dh_\lambda^</dE = -\psi_\lambda^<(0)^{-2}\int_{-\infty}^{0}\psi_\lambda^<(x)^2 dx, \tag{38a}$$

$$dh_{1/\lambda}/dE = -\psi_{1/\lambda}(0)^{-2}\int_{-\infty}^{0}\psi_{1/\lambda}(x)^2 dx, \tag{38b}$$

$$dh_\lambda/dE = \psi_\lambda(0)^{-2}\int_{0}^{\infty}\psi_\lambda(x)^2 dx, \tag{38c}$$

it follows that $h_\lambda=-h_{1/\lambda}$ and $h_\lambda^<$ and $h_{1/\lambda}$ are decreasing functions of E. The typical behavior of h_λ and $h_{1/\lambda}$ is shown in Fig. 7. Now, if there are no bound states in the gap, $h_\lambda^<$ and h_λ cannot be equal for any E in the gap. Then, since h_λ increases while $h_\lambda^<$ decreases with E, $h_\lambda^<$ can take values only in the shaded area of Fig. 7, below 0. Consequently, the right side of Eq. (37) is smaller than or equal to 1, i.e.,

$$|R^+(\lambda)| \leqslant 1 \tag{39}$$

remains valid when E is in the insulating gap.

We can then conclude that the amplitudes of $\Delta n(x)$ and $\Delta\epsilon(x)$, in the asymptotic limit, cannot exceed the upper bound

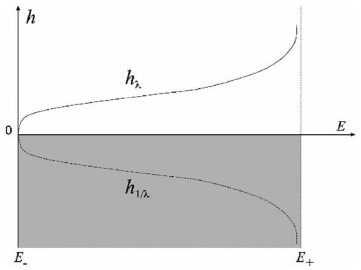

FIG. 7. Typical behavior of h_λ (solid lines) and $h_{1/\lambda}$ (dashed lines). $E_\pm$ denotes the upper (lower) edge of the insulating gap. $h_E^<$ can take values only in the shaded areas.

$$\overline{\Delta n}(x) \to 2\left(\frac{2\pi}{xb}\right)^{1/2} u_n(x)^2 e^{-2q_n x} \tag{40}$$

and

$$\overline{\Delta\epsilon}(x) \to \tilde{E}_n\overline{\Delta n}(x). \tag{41}$$

This completes the proof of NEM for insulators and one-sided perturbing potentials that do not generate bound states in the insulating gap.

Bound states in the insulating gap. We show in Appendix B that if

$$\int w(x)\psi_\pm(x)^2 dx \gtrless 0, \tag{42}$$

where $\psi_\pm(x)$ denotes the Bloch function at the upper (lower) edge of an energy band, then w generates bound states above (below) this band, even for infinitely small coupling constants. Thus, the presence of bound states in the gaps is not a rare occurrence in one dimension.

The asymptotic forms of $\Delta n(x)$ and $\Delta\epsilon(x)$ depend on how the bound states in the insulating gap are occupied. When all bound states below the branch point $\tilde{E}_n$ are occupied and the ones above $\tilde{E}_n$ are unoccupied, i.e., the situation illustrated in Fig. 5(b), the asymptotic behavior Eqs. (35) and (36) of $\Delta n(x)$ and $\Delta\epsilon(x)$ remains unchanged.

Consider now that there are unoccupied bound states below $\tilde{E}_n$, as illustrated in Fig. 5(c). Let φ_0, of energy E_0, be such a state. For $x>0$, φ_0 is equal, up to a factorization constant, to the exponentially decaying Bloch function of energy E_0 [$E_0=E_n(\lambda_0)$]:

$$\varphi_0(x) = \left(\frac{(1-\lambda_0^2)\Lambda}{\int_0^b |\psi_{n,\lambda_0}(x)|^2 dx}\right)^{1/2}\psi_{n,\lambda_0}(x), \tag{43}$$

where

$$\Lambda \equiv \int_0^\infty |\varphi_0(x)|^2 dx \quad (\Lambda \leqslant 1). \tag{44}$$

We have $|\lambda_0|=e^{-q_0 b}$, with q_0 strictly larger than zero, and $\psi_{\lambda_0}(x)=e^{-q_0 x}u_0(x)$, with $u_0(x+b)=(-1)^n u_0(x)$. Since q_0 decreases as E_0 moves away from $\tilde{E}_n$, the first unoccupied state will have the slowest exponential decay, among all unoccu-

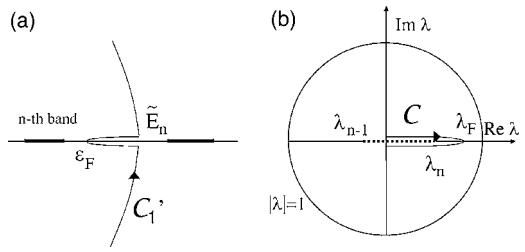

FIG. 8. (a) The contour of integration in the complex E plane ($\mathcal{C}_1'$ was introduced in Fig. 2). (b) The same contour in the complex λ plane.

pied states below $\widetilde{E}_n$. Thus, when the contribution of these states is subtracted from Eq. (35), one finds that the asymptotic form of $\Delta n(x)$ is determined by the first unoccupied bound state:

$$\Delta n(x) \to -\frac{2(1-\lambda_0^2)\Lambda}{\int_0^b |\psi_{n,\lambda_0}(x)|^2 dx} u_0(x)^2 e^{-2q_0 x}, \tag{45}$$

with the index 0 referring to the first unoccupied bound state in the insulating gap. Similarly,

$$\Delta\epsilon(x) \to -\frac{2(1-\lambda_0^2)E_0\Lambda}{\int_0^b |\psi_{n,\lambda_0}(x)|^2 dx} u_0(x)^2 e^{-2q_0 x}. \tag{46}$$

Since $\Lambda \leqslant 1$, the amplitudes of $\Delta n(x)$ and $\Delta\epsilon(x)$ cannot exceed, in the asymptotic limit, the upper bounds

$$\overline{\Delta n}(x) \to \frac{2(1-\lambda_0^2)}{\int_0^b |\psi_{n,\lambda_0}(x)|^2 dx} u_0(x)^2 e^{-2q_0 x} \tag{47}$$

and

$$\overline{\Delta\epsilon}(x) \to E_0 \overline{\Delta n}(x). \tag{48}$$

The results remain the same if, instead of unoccupied bound states below $\widetilde{E}_n$, there are occupied bound states above $\widetilde{E}_n$, as in Fig. 5(d). In this case, the index 0 will refer to the last occupied bound state.

Continuum states in the insulating gap. We consider the case when $w(x)$ fills the entire insulating gap with continuum spectrum, such as when the insulator is in contact with an infinite metal. In this case, $R_n^+(\lambda)$ has a branch cut on the real axis. The states are considered occupied up to the Fermi energy ϵ_F, which is in the insulating gap of the unperturbed insulator. We consider only the generic case when $\epsilon_F \neq \widetilde{E}_n$ and define λ_F by $\epsilon_F = E_n(\lambda_F)$. This λ_F is located strictly inside the unit circle, as opposed to the case of metals. We also define $q_F > 0$ so that $|\lambda_F| = e^{-q_F b}$.

In Eq. (16), we consider the contour of integration shown in Fig. 8(a). Mapping into the complex λ plane and using again the fundamental property of the Bloch functions,

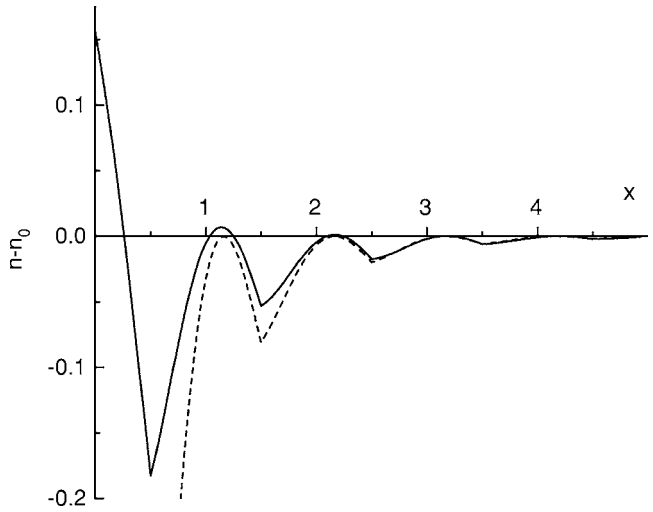

FIG. 9. The exact (solid) and asymptotic (dashed) $\Delta n(x)$ for the model of Eq. (32) ($V_0=-5$, $v_0=-2$, $b=1$, and $\epsilon_F=4$).

$$\Delta n(x) = \lambda_F^{2m} \frac{2i}{b} \int_{\mathcal{C}} R_n^+(\lambda) \psi_{n,\lambda}(y)^2 \left(\frac{\lambda}{\lambda_F}\right)^{2m-1} \frac{d\lambda}{\lambda_F}. \tag{49}$$

With the new variable q defined by $\lambda/\lambda_F = e^{-qb}$, we can write

$$\Delta n(x) = -4\lambda_F^{2m} \operatorname{Im} \int_{q>0} R_n^+(\lambda^+) \psi_{n,\lambda}(y)^2 e^{-2mqb} dq, \tag{50}$$

where $\lambda^+ \equiv \lambda + i0^+$. The asymptotic behavior can be extracted from a simple integration by parts:

$$\Delta n(x) \to -2 \operatorname{Im}[R_n^+(\lambda_F^+)] \frac{\psi_{n,\lambda_F}(x)^2}{x}. \tag{51}$$

By writing $\psi_{n,\lambda_F}(x)$ as $u_F(x)e^{-q_F x}$, with $u_F(x+b) = (-1)^n u_F(x)$, we can conclude that

$$\Delta n(x) \to -2 \operatorname{Im}[R_n^+(\lambda_F^+)] u_F(x)^2 \frac{e^{-2q_F x}}{x}. \tag{52}$$

Similarly,

$$\Delta\epsilon(x) \to -2 \operatorname{Im}[R_n^+(\lambda_F^+)] \epsilon_F u_F(x)^2 \frac{e^{-2q_F x}}{x}. \tag{53}$$

An implementation of Eq. (52) in the perturbed Kronig-Penney model Eq. (32) is shown in Fig. 9. Notice again that the asymptotic regime starts from two lattice constants away from the perturbation.

To end the proof of NEM, we need to give an upper bound on the amplitudes of $\Delta n(x)$ and $\Delta\epsilon(x)$. The imaginary part of $R_n^+(\lambda_F^+)$ is proportional to the local density of states $g(E,x)$ at $E=\epsilon_F$ and $x=0$. Indeed, Eq. (22) provides the following identity:

$$\operatorname{Im}[R_n^+(\lambda_F^+)] \psi_{n,\lambda_F}(x)^2 = \frac{1}{2\pi} \frac{d\epsilon_F}{dq_F} \operatorname{Im}[G_{\epsilon_F + i0}(x,x)], \tag{54}$$

leading to

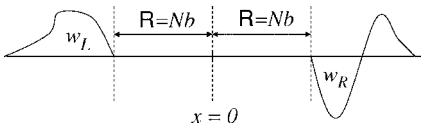

FIG. 10. The case when the point x (near the origin), where we evaluate $\Delta n(x)$ and $\Delta\epsilon(x)$, has perturbing potentials to the left and to the right.

$$\mathrm{Im}[R_n^+(\lambda_F^+)] = \frac{d\epsilon_F/dq_F}{2\psi_{n,\lambda_F}(0)^2} g(\epsilon_F,0). \tag{55}$$

Note that the coefficient in front of $g(\epsilon_F,0)$ is determined by the unperturbed system. If we limit ourselves to the generic case of w's that generate finite densities of states at ϵ_F, then NEM follows from Eqs. (52) and (55). For practical applications, we consider this argument sufficient. However, to achieve a full proof of NEM, we need to consider also the cases when $g(E,0)$ diverges (or becomes extremely large) as $E \to \epsilon_F$. For these special cases, the asymptotic form of Eq. (50) cannot be extracted from a simple integration by parts and $\Delta n(x)$ is no longer given by Eq. (52); its specific functional form will depend on the type of singularity of $g(E,0)$. This special situation will not be discussed here.

B. Two-sided perturbations

We consider here the case when the point x, where we evaluate $\Delta n(x)$ and $\Delta\epsilon(x)$, has perturbing potentials w_L to the left and w_R to the right, as schematically shown in Fig. 10. For one-sided perturbations, $\Delta n(x)$ decays as x moves further and further away from the perturbation and, because of this simple picture, NEM is intuitive and simple to grasp. When left and right perturbing potentials are present, this simple picture is gone: there will be interference terms in $\Delta n(x)$, whose amplitudes remain constant in the region between the two perturbing potentials. In addition, $w_L + w_R$ can induce a strong, qualitative change of the system, namely, the energy bands may become quantized. In spite of all of these, we will show the following. For metals, the interference terms are not negligible, but $\Delta n(x)$ still remains bounded. For insulators, the interference terms are exponentially small and $\Delta n(x)$ is given by the simple superposition of the left and right density changes. Similar conclusions hold for $\Delta\epsilon(x)$.

For convenience, we fix the origin in the middle of the interval that separates the two perturbing potentials and consider the distance R from the origin to the right (left) edge of $w_{L,R}$ to be an integer of b, $\mathsf{R}=Nb$. We are interested in the behavior, for x near the origin, of $\Delta n(x)$ and $\Delta\epsilon(x)$ when $\mathsf{R} \to \infty$. We follow our general strategy and derive first an expression for the Green's function on the interval $[-\mathsf{R},\mathsf{R}]$. We look again for the solutions of the Schrödinger equation

$$[-d^2/dx^2 + v(x) + w_L(x) + w_R(x)]\psi(x) = E\psi(x) \tag{56}$$

that decay at $\mp\infty$. On the intervals from $-\infty$ to the left edge of w_L and from the right edge of w_R to $+\infty$, these solutions can be expressed as in Eqs. (19) and (20), in terms of the total (corresponding to $w_L + w_R$) transmission and reflection coefficients. Then, one can use again the reflection and transmission coefficients to continue these solutions inside the interval $-\mathsf{R} < x < \mathsf{R}$. On this interval, they take the following form:

$$\psi_\lambda^>(x) = \frac{T(\lambda)}{\tilde{T}_R(\lambda)}[\psi_\lambda(x) + \tilde{R}_R^-(\lambda)\psi_{1/\lambda}(x)] \tag{57}$$

and

$$\psi_\lambda^<(x) = \frac{T(\lambda)}{\tilde{T}_L(\lambda)}[\psi_{1/\lambda}(x) + \tilde{R}_L^+(\lambda)\psi_\lambda(x)], \tag{58}$$

where $\tilde{T}_{L,R}(\lambda)$ and $\tilde{R}_{L,R}^\pm(\lambda)$ are the transmission and reflection coefficients of the left (right) potentials, and $T(\lambda)$ is the total transmission coefficient

$$T(\lambda) = \frac{\tilde{T}_L(\lambda)\tilde{T}_R(\lambda)}{1 - \tilde{R}_L^+(\lambda)\tilde{R}_R^+(\lambda)}. \tag{59}$$

All these coefficients depend on R. If $R_{L,R}^\pm(\lambda)$ denote the reflection coefficients when the right (left) edge of $w_{L,R}$ is at $x=0$ [thus $R_{L,R}^\pm(\lambda)$ are independent of R], then Eq. (37) gives

$$\tilde{R}_{L,R}^\pm(\lambda) = \lambda^{2N} R_{L,R}^\pm(\lambda). \tag{60}$$

The Wronskian of the two independent solutions is the same as given in Eq. (21). From the two independent solutions Eqs. (57) and (58) and their Wronskian, we derive, for $-\mathsf{R} < x < \mathsf{R}$, the following identity:

$$[G_E(x,x) - G_E^0(x,x)]\frac{dE}{\pi i} = \frac{2i\lambda^{2N}R_L^+(\lambda)}{1-\lambda^{4N}R_L^+(\lambda)R_R^-(\lambda)}\psi_\lambda(x)^2\frac{d\lambda}{b\lambda} + \frac{2i\lambda^{2N}R_R^-(\lambda)}{1-\lambda^{4N}R_L^+(\lambda)R_R^-(\lambda)}\psi_{1/\lambda}(x)^2\frac{d\lambda}{b\lambda} + \frac{4i\lambda^{4N}R_L^+(\lambda)R_R^-(\lambda)}{1-\lambda^{4N}R_L^+(\lambda)R_R^-(\lambda)}\psi_\lambda(x)\psi_{1/\lambda}(x)\frac{d\lambda}{b\lambda}. \tag{61}$$

We now can use Eqs. (16) and (17) to find $\Delta n(x)$ and $\Delta\epsilon(x)$, for x near the origin. By using the fundamental property of the Bloch functions, we can understand the behavior of each term in the above identity. For λ not on the unit circle, as is the case in our integrals, the first term decays exponentially as x moves to the right, the second term decays exponentially

as x moves to the left, but the third term is periodic, with period b. Fortunately, the amplitude of this term becomes smaller and smaller as the two perturbing potentials are moved apart.

From Eq. (61), one can easily obtain the energy spectrum of the perturbed system. So let us discuss first how the simultaneous presence of $w_{L,R}$ affects the energy spectrum. We are interested in the last occupied band (indexed by n), so, from now on, λ is considered on the nth Riemann sheet. The energies of the discrete states, if any, are given by the poles of the Green's function. From the identity Eq. (61), one can see that these poles correspond to those λ satisfying the equation

$$R_L^+(\lambda)R_R^-(\lambda) = 1/\lambda^{4N}. \tag{62}$$

Since the discrete state energies are real, the solutions of Eq. (62) are always located on the unit circle or on the real axis, away from the branch cuts. Inside the unit disk we have $|\lambda|<1$; consequently, $|1/\lambda^{4N}|$ becomes very large in the limit $\mathsf{R}\to\infty$. Thus, if there are solutions of Eq. (62) inside the unit disk, they must be located very close to the poles of either $R_L^+(\lambda)$ or $R_R^-(\lambda)$. In other words, these solutions are perturbations of the bound states generated by w_L (or w_R) alone, already discussed in the previous subsection. Poles on the unit circle exist if and only if both $|R_{L,R}^{\pm}(\lambda)|$ are equal to 1. In this case, the energy band degenerates into a discrete energy spectrum. Because the left-hand side of Eq. (62) is slowly varying compared to the right-hand side, one can get the qualitative picture by setting the left-hand side constant. If the amplitudes of $R_{L,R}^{\pm}(\lambda)$ are equal to 1 on the whole unit circle, then Eq. (62) has $4N$ solutions λ_k distributed on the unit circle; the spacing between two consecutive solutions is $2\pi/4N+O(1/N^2)$. This is the picture in the λ plane. In the E plane, the discrete energies are given by $E(\lambda_k)$. The spacing between two consecutive energies is $2\pi\partial_\lambda E(\lambda_k)/4N + O(1/N^2)$.

1. Metals

The effects of band quantization will be the strongest for metals, since the Fermi energy lies inside the band. We calculate the density change by integrating Eq. (61) along the contour of integration shown in Fig. 3(a) and map the integral into the complex λ plane. The asymptotic behavior of $\Delta n(x)$, for large R, is determined by the behavior of the integrand near λ_F and $1/\lambda_F$ [see Fig. 3(b)]. In the immediate vicinity of these points, we can replace the slowly varying functions in the right-hand side of Eq. (61) (the reflection coefficients and the Bloch functions) with their value at λ_F and $1/\lambda_F$. The integral then can be explicitly calculated and the result is

$$\begin{aligned}\Delta n(x,\mathsf{R}) \to{}& \frac{2}{\mathsf{R}}\,\mathrm{Im}\,\frac{\tanh^{-1}[\lambda_F^{2N}(R_L^+R_R^-)^{1/2}]}{(R_L^+R_R^-)^{1/2}}\\ &\times[R_L^+\psi_{\lambda_F}(x)^2 + R_R^-\psi_{1/\lambda_F}(x)^2]\\ &+\frac{2}{\mathsf{R}}\,\mathrm{Im}[\ln(1-\lambda_F^{4N}R_L^+R_R^-)]|\psi_{\lambda_F}(x)|^2,\end{aligned} \tag{63}$$

with the reflection coefficients evaluated at λ_F.

To prove NEM, we need to find an upper bound on the above expression. A simple analysis reveals that the largest density changes occur when $|R_{L,R}^{\pm}|=1$, i.e., when the band is quantized at the Fermi energy. In this case, we can rewrite Eq. (63) as

$$\begin{aligned}\Delta n(x,\mathsf{R}) \to{}& \frac{4}{\mathsf{R}}\,\mathrm{Im}[\tanh^{-1}[\lambda_F^{2N}(R_L^+R_R^-)^{1/2}]]\\ &\times\mathrm{Re}[(R_L^+R_R^{-*})^{1/2}\psi_{\lambda_F}(x)^2]\\ &+\frac{2}{\mathsf{R}}\,\mathrm{Im}[\ln(1-\lambda_F^{4N}R_L^+R_R^-)]|\psi_{\lambda_F}(x)|^2.\end{aligned} \tag{64}$$

The density change, as a function of R, has discontinuities every time when $\lambda_F^{4N}R_L^+R_R^-=1$, i.e., when a discrete energy crosses the Fermi level. These discontinuities are finite: since $|\mathrm{Im}[\tanh^{-1}(z)]|\leqslant\pi/4$ and $|\mathrm{Im}[\ln(1-z)]|\leqslant\pi/2$, for $|z|\leqslant 1$, the amplitude of the asymptotic term of $\Delta n(x,\mathsf{R})$ cannot exceed the upper bound

$$\overline{\Delta n}(x,\mathsf{R}) \to \frac{2\pi}{\mathsf{R}}|\psi_{\lambda_F}(x)|^2, \tag{65}$$

independent of $w_{L,R}$ potentials and of the position of the Fermi energy relative to the discretized energies. Similarly, the amplitude of $\Delta\epsilon(x,\mathsf{R})$ cannot exceed the upper bound

$$\overline{\Delta\epsilon}(x,\mathsf{R}) \to \epsilon_F\overline{\Delta n}(x,\mathsf{R}), \tag{66}$$

and this completes our discussion of NEM for metals.

These upper bounds are optimal, in the sense that there are $w_{L,R}$ potentials (the worst scenario) that generate a density and an energy density that are equal to these upper bounds. By comparing with the results of the previous section, one can see that interference has nontrivial effects: these upper bonds are not simply the superposition of the left and right upper bounds found in the previous section.

2. Insulators

We consider first the situation when there are no bound states in the insulating gap. We can take $\mathcal{C}_1'$ of Fig. 2 as the contour of integration in Eq. (16), which is mapped into $\mathcal{C}_1$ in the complex λ plane. For any λ on this curve, the denominator in the right side of Eq. (61) goes exponentially to 1 as $\mathsf{R}\to\infty$. Consequently, in this limit, the structure of the integrand becomes completely analogous with the one studied in the previous subsection. The asymptotic behavior can be extracted as previously and the result is

$$\begin{aligned}\Delta n(x,\mathsf{R}) \to{}& -2R_L^+(\lambda_n)\left(\frac{2\pi}{b\mathsf{R}}\right)^{1/2}u_n(x)^2e^{-2q_n\mathsf{R}} - 2R_R^-(\lambda_n)\\ &\times\left(\frac{2\pi}{b\mathsf{R}}\right)^{1/2}u_n'(x)^2e^{-2q_n\mathsf{R}} - 2R_L^+(\lambda_n)R_R^-(\lambda_n)\\ &\times\left(\frac{\pi}{b\mathsf{R}}\right)^{1/2}u_n(x)u_n'(x)e^{-4q_n\mathsf{R}}.\end{aligned} \tag{67}$$

Thus, in the limit of large R, $\Delta n(x,\mathsf{R})$ is just the sum of the independent density changes due to the left and right poten-

tials, plus an exponentially small correction. From the previous subsection, we can conclude that the amplitude of $\Delta n(x,\mathsf{R})$ cannot exceed, for large R, the upper bound

$$\overline{\Delta n}(x,\mathsf{R}) \to 2\left(\frac{2\pi}{b\mathsf{R}}\right)^{1/2}[u_n(x)^2+u_n'(x)^2]e^{-2q_n\mathsf{R}}. \quad (68)$$

Similarly,

$$\overline{\Delta\epsilon}(x,\mathsf{R}) \to \tilde{E}_n\overline{\Delta n}(x,\mathsf{R}). \quad (69)$$

For the case when there are bound or continuum states in the insulating gap, the conclusion is the same: for large R, the density change near the origin is the sum of the independent changes induced by the left and right potentials. Upper bounds on $\Delta n(x,\mathsf{R})$ can be trivially derived from the previous subsection.

V. THE NEARSIGHTEDNESS RANGE

The nearsightedness range $\mathsf{R}(x,\Delta n)$ was introduced as the range beyond which any perturbation, no matter how large, induces a density change at x less than the given Δn. The asymptotic $\mathsf{R}(x,\Delta n)$, in the limit $\Delta n\to 0$, can now be easily calculated from the upper bounds on $\Delta n(x)$, derived in this paper. Since the periodic systems are macroscopically homogeneous, the asymptotic $\mathsf{R}(x,\Delta n)$ will be independent of x.

When solving for R in $\overline{\Delta n}(x,\mathsf{R})=\Delta n$, we first average $\overline{\Delta n}(x,\mathsf{R})$ over one unit cell. For metals, Eq. (65) leads to the following asymptotic expression:

$$\mathsf{R}(\Delta n) \to 1/\Delta n. \quad (70)$$

Such universal behavior is characteristic only to one dimension; in higher dimensions, the nearsightedness range will depend on the average particle density.[2]

For insulators and w's that generate no bound states in the insulating gap, Eq. (68) leads to

$$\mathsf{R}(\Delta n) \to \frac{1}{2q_n}\ln\frac{\tilde{n}}{\Delta n}, \quad (71)$$

where

$$\tilde{n}=\frac{4\sqrt{2\pi q_n}}{b}\int_0^b [u_n(x)^2+u_n'(x)^2]dx. \quad (72)$$

In the small-gap and tight-binding limits, $\tilde{n}$ is completely determined by the exponential decay constant q_n, $\tilde{n}\to 4q_n\sqrt{2/\pi}$ and $\tilde{n}\to 4\sqrt{q_n/\pi b}$, respectively.

It is important to notice that the nearsightedness range does not depend on the details of the underlying potential $v(x)$, but on some simple parameters that can be defined also for nonperiodic potentials. For example, q_n can be identified with the exponential decay constant of the density matrix.

VI. DISCUSSION

The above analysis provides a quantitative analysis of the nearsightedness of electronic matter for noninteracting fermions, moving in one dimension under the action of periodic potentials. Although the simplest case possible, it allowed us to understand the different mechanisms behind NEM. The asymptotic behavior of $\Delta n(x)$ was found to be determined by the reflection coefficient. For specific cases, the amplitude of $\Delta n(x)$ cannot exceed an upper bound simply because the reflection coefficients are always smaller than 1. More general, and now including 2D and 3D, one will find that asymptotic behavior of $\Delta n(x)$ is determined by certain elements of the scattering matrix and, for specific cases, the unitarity of the scattering matrix imposes certain upper bounds. The situation is, however, more complicated when bound states appear in the insulating gap or when the bands become quantized.

We have introduced the concept of the nearsightedness range $\mathsf{R}(x,\Delta n)$, which is well defined only because there is this upper bound on $\Delta n(x)$. $\mathsf{R}(x,\Delta n)$ is a characteristic of the unperturbed system and gives a simple and effective measure of nearsightedness. For periodic metals, we found $\mathsf{R}(x,\Delta n)$ to have, in the asymptotic limit $\Delta n\to 0$, a universal expression, namely, $1/\Delta n$. For insulators, $\mathsf{R}(x,\Delta n)$ is strongly dependent on the band structure but has a weak, logarithmic dependence on Δn.

Although the estimates given in this paper can be applied only to 1D systems, we think we gain some knowledge that can be useful for more general situations. We are convinced that NEM exists in dimensions higher than one, where it can be quantified in a similar way. In particular, we believe that a complete theoretical analysis and optimization of the $O(N)$ divide and conquer algorithm is possible in all dimensions. Preliminary results in this direction have been already given in Ref. 2. The one-dimensional analysis proved to be extremely useful by providing a viable strategy and some understanding of the effects of the bound states in the insulating gap and of the band quantization on NEM.

ACKNOWLEDGMENTS

I would like to thank Professor Walter Kohn, who suggested and supervised this project. This work was completed when the author was visiting the Physics Department at the University of California at Santa Barbara and the Collaboratory for Advanced Computing and Simulation at the University of Southern California, and was supported by NSF Grants No. DMR03-13980 and No. DMR04-27188, and DOE Grant No. DE-FG02-04ER46130.

APPENDIX A

We estimate here the exponential decay rate q_n, related to the branch point by $|\lambda_n|=e^{-q_nb}$. According to Ref. 14,

$$q_n=\frac{1}{b}\ln[|\mu_n|+\sqrt{\mu_n^2-1}], \quad (A1)$$

where μ_n is the Kramers function evaluated at the branch point $\tilde{E}_n$, defined by $d\mu/dE|_{E=\tilde{E}_n}=0$.

For small gaps, the behavior of $\mu(E)$ inside the entire gap is well approximated by a quadratic function of E:

$$\mu(E) \simeq (-1)^n \left(1 - \frac{m_n^* b^2 (E - E_n^+)(E - E_{n+1}^-)}{2(E_{n+1}^- - E_n^+)}\right), \quad \text{(A2)}$$

where $E_n^\pm$ is the upper (lower) edge of the nth band, and m_n^* is the effective mass at the upper edge of the nth band. Since $\mu_n \gtrsim 1$, $q_n \simeq (1/b)\sqrt{2(|\mu_n|-1)}$, which, together with Eq. (A2), leads to

$$q_n = \frac{1}{2}\sqrt{m_n^*(E_{n+1}^- - E_n^+)}. \quad \text{(A3)}$$

We consider now a periodic potential $\Sigma_l V_a(x-lb)$, where $V_a(x)$ vanishes for $|x|>c$ and has atomic levels $E_n \equiv -k_n^2$, $n=1,\ldots$. In the limit $b\to\infty$, we show that

$$q_n \simeq \frac{1}{b} \ln \frac{8\sqrt{-E_n}}{ebW_n}, \quad \text{(A4)}$$

where W_n is the width of the nth energy band. For x in $[-b/2, b/2]$ and $|x|>c$, the solutions of the Schrödinger equation at an energy $E=-k^2$ are of the general form

$$\psi(x) = \begin{cases} a_-(k)e^{-kx} + b_-(k)e^{kx}, & x < -c, \\ a_+(k)e^{-kx} + b_+(k)e^{kx}, & x > c, \end{cases} \quad \text{(A5)}$$

with

$$\begin{pmatrix} a_+(k) \\ b_+(k) \end{pmatrix} = \hat{T}(k) \begin{pmatrix} a_-(k) \\ b_-(k) \end{pmatrix}, \quad \text{(A6)}$$

$\hat{T}(k)$ being the transfer matrix of the potential V_a. The energy levels of V_a correspond to the zeros of $T_{22}(k)$, already denoted by k_n. The Kramers function is given by

$$\mu(k) = \frac{1}{2}[T_{11}(k)e^{-kb} + T_{22}(k)e^{kb}]. \quad \text{(A7)}$$

We estimate first the bandwidths. We look for the solutions of $\mu(k)=\pm 1$, which give the band edges. For b large, the solutions of this equation must be located very close to the zeros of $T_{22}(k)$ since, otherwise, the second term in Eq. (A7) becomes very large. We can then linearize, $T_{22}(k) \simeq (k-k_n)T'_{22}(k_n)$, and neglect the exponentially small term in Eq. (A7), in which case the equation $\mu(k)=\pm 1$ can be trivially solved, leading to

$$W_n = \frac{8k_n e^{-k_n b}}{|T'_{22}(k_n)|}. \quad \text{(A8)}$$

We now calculate $\tilde{E}_n = -\tilde{k}_n^2$, defined by[14]

$$\left(\frac{d\mu}{dk}\right)_{k=\tilde{k}_n} = 0 \Leftrightarrow T'_{22}(\tilde{k}_n) \simeq -bT_{22}(\tilde{k}_n). \quad \text{(A9)}$$

For b large, the solutions of the above equation must also be close to the zeros of $T_{22}(k)$. Linearizing $T_{22}(k)$, we find $\tilde{k}_n = k_n - 1/b$ and the Kramers function evaluated at $\tilde{k}_n$ is

$$\mu_n = \frac{-T'_{22}(k_n)e^{k_n b}}{2eb} = \frac{4k_n}{ebW_n}. \quad \text{(A10)}$$

Since $\mu_n \gg 1$, $q_n \simeq (1/b)\ln[2\mu_n]$ and Eq. (A4) follows.

APPENDIX B

Let $w(x)$ be a perturbing potential of finite support and such that

$$\int w(x)\psi_+(x)^2 dx > 0 \quad \text{(B1)}$$

or

$$\int w(x)\psi_-(x)^2 dx < 0, \quad \text{(B2)}$$

where $\psi_\pm(x)$ is the Bloch function at the upper (lower) edge of an energy band. We show here that, even for infinitely small coupling constants, this potential will pull bound states out from the band.

Let H_0 denote the periodic Hamiltonian and $H \equiv H_0 + \gamma w$. It is known that H has a bound state at some energy E if and only if the operator[20]

$$\hat{K}_E = \gamma w^{1/2}(E - H_0)^{-1}|w|^{1/2} \quad \text{(B3)}$$

has an eigenvalue equal to 1.[20] Here, $w^{1/2} = w/|w|^{1/2}$. We show that, for any given energy E below (above) the band, $\hat{K}_E$ has an eigenvalue equal to 1 for some positive γ, which decreases to zero as E approaches the edges of the band, provided the condition Eq. (B1) [Eq. (B2)] is satisfied.

If n is odd, the lower edge of the band corresponds to $\lambda = 1$. We take an energy E below such band and let λ, which is real and less than 1, be its corresponding λ parameter. Equation (15) gives

$$\hat{K}_E(x,x') = -\frac{2\pi\gamma}{b} w(x)^{1/2} \frac{\psi_{1/\lambda}(x_<)\psi_\lambda(x_>)}{\lambda dE/d\lambda} |w(x')|^{1/2}, \quad \text{(B4)}$$

and we notice that $dE/d\lambda \propto \lambda - 1$, for $\lambda \to 1$, i.e., the kernel $\hat{K}_E(x,x')$ diverges at $\lambda = 1$. We can separate the diverging part by expanding

$$\psi_{\lambda^{-1}}(x_<)\psi_\lambda(x_>) = \psi_-(x)\psi_-(x') + (\lambda - 1)W_\lambda(x,x'). \quad \text{(B5)}$$

This provides the following decomposition:

$$\hat{K}_E = \gamma\alpha(\lambda)|\varphi_1\rangle\langle\varphi_2| + \gamma\hat{A}(\lambda), \quad \text{(B6)}$$

where

$$\alpha \equiv -\frac{2\pi}{b\lambda dE/d\lambda}, \quad \text{(B7)}$$

$$\hat{A}(\lambda) \equiv \frac{2\pi}{b}\frac{1-\lambda}{\lambda dE/d\lambda} w^{1/2} W_\lambda |w|^{1/2}, \quad \text{(B8)}$$

and

$$\varphi_1(x) \equiv w(x)^{1/2}\psi_-(x),$$

$$\varphi_2(x) \equiv |w(x)|^{1/2}\psi_-(x). \quad \text{(B9)}$$

The first term of Eq. (B6) diverges while the second one is analytic at $\lambda = 1$. Now let Ψ be given by

$$\Psi = [\gamma \hat{A}(\lambda) - 1]^{-1} \varphi_1, \tag{B10}$$

which is well defined for small γ. Then

$$\hat{K}_E \Psi = \Psi + \{1 + \gamma\alpha(\lambda)\langle\varphi_2|[\gamma\hat{A}(\lambda) - 1]^{-1}|\varphi_1\rangle\}\varphi_1. \tag{B11}$$

In other words, $\hat{K}_E$ has an eigenvalue at +1, if

$$1 + \gamma\alpha(\lambda)\langle\varphi_2|[\gamma\hat{A}(\lambda) - 1]^{-1}|\varphi_1\rangle = 0. \tag{B12}$$

We can rewrite this equation as

$$\gamma\langle\varphi_2|[\gamma\hat{A}(\lambda) - 1]^{-1}|\varphi_1\rangle - \frac{b\lambda}{2\pi}\frac{dE}{d\lambda} = 0. \tag{B13}$$

If we denote the left side by $F(\gamma,\lambda)$, then $F(0,1)=0$ and $\partial_\gamma F(0,1) = -\int w(x)\psi_-(x)^2 dx > 0$, i.e., the conditions of the analytic implicit function theorem are satisfied, which means that, for any λ near +1, there is always a solution $\gamma(\lambda)$ to Eq. (B12). Moreover,

$$\gamma(\lambda) = -\frac{b\lambda}{2\pi}\frac{dE}{d\lambda}\left(\int w(x)\psi_-(x)^2 dx\right)^{-1} + \cdots, \tag{B14}$$

where the ellipsis indicates terms of order $o((1-\lambda)^2)$. γ is real and positive, for E below the band, and goes to zero as E approaches the band edge.

The other possible cases, $\lambda=-1$ and n even, follow in the same way.

[1] W. Kohn, Phys. Rev. Lett. **76**, 3168 (1996).
[2] E. Prodan and W. Kohn, Proc. Natl. Acad. Sci. U.S.A. **102**, 11635 (2005).
[3] J. Friedel, Philos. Mag. **43**, 153 (1952).
[4] N. Bloemberg and T. J. Rowland, Acta Metall. **1**, 731 (1953).
[5] T. J. Rowland, Phys. Rev. **119**, 900 (1960).
[6] W. Kohn, Phys. Rev. **119**, 912 (1960).
[7] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).
[8] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
[9] G. Galli, Phys. Status Solidi B **217**, 231 (2000).
[10] W. Yang, Phys. Rev. Lett. **66**, 1438 (1991).
[11] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999).
[12] S. Y. Wu and C. S. Jayanthi, Phys. Rep. **358**, 1 (2002).
[13] F. Shimojo, R. K. Kalia, A. Nakano, and P. Vashishta, Comput. Phys. Commun. **167**, 151 (2005).
[14] W. Kohn, Phys. Rev. **115**, 809 (1959).
[15] E. Prodan, Phys. Rev. B **73**, 035128 (2006).
[16] R. E. Allen, Phys. Rev. B **20**, 1454 (1979).
[17] H. A. Kramers, Physica (Amsterdam) **2**, 483 (1935).
[18] R. de L. Kronig and W. G. Penney, Proc. R. Soc. London, Ser. A **130** (1931).
[19] F. W. J. Olver, *Asymptotic and Special Functions* (A. K. Peters, Wellesley, MA, 1997).
[20] B. Simon, *Trace Ideals and Their Applications* (Cambridge University Press, New York, 1978).