

TOPICAL REVIEW

On the equivalence of different techniques for evaluating the Green function for a semi-infinite system using a localized basis

Julian Velev and William Butler

Metals and Ceramics Division, Oak Ridge National Laboratory, Oak Ridge,
TN 37831-6114, USA

and

Center for Materials for Information Technology, University of Alabama, Tuscaloosa,
AL 35487, USA

Received 1 March 2004

Published 14 May 2004

Online at stacks.iop.org/JPhysCM/16/R637

DOI: 10.1088/0953-8984/16/21/R01

Abstract

We discuss several different techniques for evaluating the Green function for a semi-infinite system using a localized basis. We demonstrate that the different techniques are different ways of calculating the self-energy associated with the surface. They give equivalent results but have different convergence properties.

(Some figures in this article are in colour only in the electronic version)

Contents

1. Introduction	638
2. Localized basis and separation of transverse modes	639
3. Self-energy of surfaces and interfaces	641
3.1. Connecting two systems	641
3.2. Green function of a system with interfaces	643
3.3. Bond cutting and atom removal	643
4. Recursion relation for the self-energy	645
5. Accelerated iteration	648
6. Eigenvalue method	649
7. Convergence	651
8. Off-diagonal matrix elements	654
9. Conclusions	655
Acknowledgments	656
Appendix. Complex band structure	656
References	657

1. Introduction

The single-particle Green function (GF) is a useful tool for studying the properties of materials because it can be used to express all of the observable properties of the system of interest [1] and because it has other advantages including the following:

- (1) It allows one to efficiently treat complicated systems starting from more idealized ones by treating the complexity as perturbation. For example, in order to obtain the electronic structure of a periodic system with a localized impurity or defect one starts from the GF of the perfect periodic system and treats the impurity as a perturbation. Similarly, the presence of a surface can be considered to be a perturbation to the GF of an infinite medium.
- (2) The Green function is very useful for calculating the response of a system to external fields, e.g. transport properties in linear response [2, 3].
- (3) The Green function is convenient for treating disordered systems because its average over configurations (instances of a system) can be used to obtain configuration properties, whereas the configurational average of a wavefunction is typically meaningless [4].

For these reasons, the ability to calculate the GF of a system with an arbitrary geometry is very important. For a finite system, the GF may be evaluated using localized basis functions quite simply as the inverse of the Hamiltonian operator in real space. Similarly, the GF for an infinite periodic system may be obtained as the inverse of the Hamiltonian in reciprocal space. However, in the case of semi-infinite systems, e.g. an infinite periodic system with a surface, we encounter a problem because the Hamiltonian cannot be directly inverted in real space due to the fact that the Hamiltonian matrix is infinite, nor can the problem be trivially reformulated in reciprocal or momentum space because translational symmetry in the direction perpendicular to the surface is lacking.

We shall review several approaches that have been used to obtain the GF of semi-infinite systems, sometimes referred to as the surface Green function (SGF). The same methods can be used to obtain the GF of infinite periodic systems or bulk GF (BGF). Without attempting a detailed classification, we have grouped the methods into three groups—(1) those based on the recursion relation for the SGF, (2) accelerated iteration based on amplitude transfer matrices and (3) an eigenvalue approach based on direct calculation of the wavefunction coefficients and the relative phase of wavefunctions on adjacent layers. In addition, we review a ‘bond-cutting’ approach that can be used to calculate the GF for a semi-infinite system by cleaving the infinite system into two semi-infinite pieces.

We shall show that these different approaches yield equivalent results for the Green function, but that their convergence properties differ. These methods and their variations were developed at different times and by different groups. Moreover, they were each formulated and implemented in the framework of the particular problems of immediate concern to their developers. This obscures the relationships between the methods and makes it difficult to extend a method to a more general problem. For these reasons, it is not straightforward for a person venturing into this field to see the whole picture and to choose the optimal technique for the problem at hand. In this paper, we attempt to present a unified treatment of the various methods of calculating the GF for semi-infinite systems or systems with a surface or interface.

We discuss systems that are either finite in the plane of the surface or have translational symmetry in the plane of the surface or interface. However, the physical requirement is a separation of the transverse normal modes of the system so that each of these modes (sometimes called channels) can be treated independently. In the case of systems finite in the plane of the interface no separation is necessary, the whole system can be treated at once (at least in

principle). In the case of systems that are infinite but translationally invariant in the plane of the interface such a separation can be achieved by means of a Fourier transform. Systems that are infinite and disordered can be treated within the coherent potential approximation [5] which essentially recovers the translational invariance.

We consider only ideal interfaces; topics such as surface reconstruction are beyond the scope of this paper. However, the GF of the ideal surface is usually the starting point for studies of more general systems. All of the methods discussed can treat a *finite* number of inequivalent layers near the surface or interface. The final results are expressions for the on-site GF of an arbitrary layer in bulk or the surface layer in the case of a semi-infinite geometry. The knowledge of the on-site GF, however, is enough to obtain the GF matrix elements between arbitrary layers. Moreover, the GF of a system with multiple interfaces can be obtained from the GFs of the parts by mapping the influence of one part of the system on the other via a self-energy term. A more formal approach to obtaining the GF of a system with interfaces is the method of surface Green function matching [6].

Although we are mostly concerned with the development of the GF methods, it is worth mentioning a few of the applications. This area of research is vast and in this paper we do not attempt to give a comprehensive review of it. Our discussion is based on our experience in electron transport in solids, which is treated using an electronic Hamiltonian on a discrete set of localized basis functions on a lattice. In this form the GF formalism has been used to study transport through mesoscopic devices [7], exchange coupling, giant and tunnelling magnetoresistance [8, 9] and surface and interface states [6]. Other excitations, e.g. phonons and magnons, can be treated using an equivalent formalism [5]. GF methods have been applied, for example, to study phonons and lattice dynamics [10]. The zero-frequency limit of this approach, i.e. static elastic Green functions, have been used to treat strains associated with planar defects in crystals [11, 12]. Our discussion is restricted to Green functions that can be represented by matrix inverses; however, the GF formalism is equally well developed for the case of continuous media [2, 6]. Moreover, by exploiting the analogy between electronic and classical waves, the GF method has been formulated to study elastic waves [13].

We will show that the different methods for calculating the surface GF amount to different ways of calculating the self-energy associated with the surface. The work is organized as follows. In section 2 we introduce a general localized basis in which we write the secular and the GF equation. For systems which are infinite and periodic in the plane of the interface systems we use a Fourier transform in the plane of the interface to reduce the secular equation to a one dimensional form. In section 3 we introduce the self-energy associated with the surface; in sections 4, 5, and 6 we discuss the recursion, iterative and eigenvalue methods respectively; and the convergence of the different methods is discussed in section 7; finally, in section 8 we discuss how to obtain off-diagonal GF matrix elements.

2. Localized basis and separation of transverse modes

The GF formalism that we employ is based on the assumption that the wavefunction can be expanded in terms of orbitals localized around each atomic site (LCAO). However, the formalism is more general. For example, similar techniques for the surface GF have been employed in the context of multiple-scattering theory (MS) [14] using a non-localized plane wave basis set. Also, analogous expressions for the GF have been obtained in the basis of linear combinations of muffin-tin orbitals (LMTO) [15]. We employ LCAO because the physical picture is clearer and the formulation simpler; however, it is not an inherent limitation of the methods.

We assume the system to be translationally invariant in the plane of the surface or interface. However, the translational invariance is broken in the direction perpendicular to the surface, say z . The *principal layer* (PL) is the unit cell in the z direction which can consist of one or more atomic layers. Interaction beyond neighbouring PLs is neglected. The planar orbital is defined as a state formed out of the atomic states explicitly satisfying the Bloch theorem in the plane of the surface

$$|\bar{\mathbf{k}}, z, \alpha, i\rangle = \frac{1}{\sqrt{N_x N_y}} \sum_{\bar{\mathbf{R}}} e^{i\bar{\mathbf{k}} \cdot \bar{\mathbf{R}}} |\bar{\mathbf{R}}, z, \alpha, i\rangle \quad (1)$$

where $\bar{\mathbf{R}}$ is the projection of the lattice vector in the plane, z labels the principal layer, i labels the atomic layer within the principal layer and α is a combined spin and orbital index.

The wavefunction of the system can be expanded in terms of the planar orbitals,

$$|\psi(\bar{\mathbf{k}}k_z)\rangle = \sum_{z,\alpha i} C_{z,\alpha i} |\bar{\mathbf{k}}, z, \alpha, i\rangle, \quad (2)$$

where

$$C_{z,\alpha i} = e^{ik_z z} C_{\alpha i}, \quad (3)$$

and all normalization coefficients are absorbed into the expansion coefficient.

In the basis of planar orbitals, the Schrödinger equation $H|\psi\rangle = E|\psi\rangle$ becomes

$$\sum_m \bar{H}_{nm} C_m = 0, \quad (4)$$

where matrix multiplication over the indices within the PL is understood and \bar{H} is defined

$$\bar{H}_{nm|\alpha i \beta j} = \langle \bar{\mathbf{k}}, n, \alpha i | H - E | \bar{\mathbf{k}}, m, \beta j \rangle = H_{nm|\alpha i \beta j} - E S_{nm|\alpha i \beta j}, \quad (5)$$

where H is the Hamiltonian matrix in the planar orbital basis and S is the overlap matrix of the basis functions. In many applications the planar orbitals are assumed orthogonal $S_{nm|\alpha i \beta j} = \delta_{nm} \delta_{\alpha \beta} \delta_{ij}$. We have not made this assumption but to simplify the notation we have assimilated the basis overlap S in the definition of the Hamiltonian matrix elements \bar{H} .

The equation for the Green function in operator form $(E - H)G = 1$ can be written in the same representation as

$$\sum_m \bar{H}_{n,m} G_{m,n'} = -\delta_{nn'}, \quad (6)$$

where $G_{m,n'}$ stands for the matrix elements of the GF in the planar orbital basis.

Due to our use of the concept of principal layers, the sum over z in the equations above contains only three terms

$$\bar{H}_{n,n+1} C_{n+1} + \bar{H}_{n,n} C_n + \bar{H}_{n,n-1} C_{n-1} = 0, \quad (7)$$

$$\bar{H}_{n,n+1} G_{n+1,n'} + \bar{H}_{n,n} G_{n,n'} + \bar{H}_{n,n-1} G_{n-1,n'} = -\delta_{nn'}. \quad (8)$$

Further simplification can be accomplished using the translational invariance of the bulk $\bar{H}_{n,m} = \bar{H}_{m-n}$. Choosing $n' = 0$ as a reference layer we can write equations (7) and (8) as

$$\bar{H}_1 C_1 + \bar{H}_0 C_0 + \bar{H}_{\bar{1}} C_{\bar{1}} = 0, \quad (9)$$

$$\bar{H}_1 G_{n+1,0} + \bar{H}_0 G_{n,0} + \bar{H}_{\bar{1}} G_{n-1,0} = -\delta_{n0}. \quad (10)$$

For a finite system, we expand the wavefunction in the atomic orbitals, $|\bar{\mathbf{R}}, z, \alpha, i\rangle$, instead of the planar orbitals. All equations derived remain exactly the same except that there is no $\bar{\mathbf{k}}$ index and all quantities are matrices with respect to $\bar{\mathbf{R}}$ which labels the atomic positions in the plane.

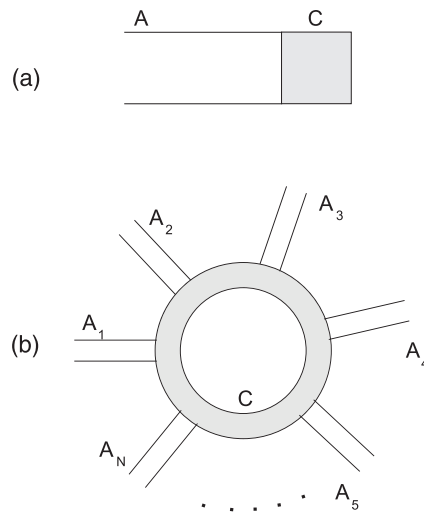


Figure 1. (a) Surface of a semi-infinite system showing the surface region (C) and semi-infinite lead (A) in which all principal layers are identical; (b) multiprobe geometry showing sample region (C) and semi-infinite leads (A_i).

3. Self-energy of surfaces and interfaces

There are many reasons that one may need to calculate the GF for a semi-infinite system. Most obviously, if one is interested in the properties of a surface, figure 1(a), the surface projection of the GF for a semi-infinite system is needed. Another important use for the GF for a semi-infinite system is the calculation of transport properties, figure 1(b). In the most general geometry involving electron transport through a sample, there are a number of long leads $A_i, i = 1, \dots, N$, attached to a finite central region C. The leads are assumed much longer than the region of interest and are usually considered to be semi-infinite. In this application, the GF for each of the semi-infinite leads is needed. In general, the leads will have a finite cross-section. In the simplest cases, however, both the leads and the sample may be approximated as translationally invariant in the plane of the interface.

3.1. Connecting two systems

Let us first consider two systems A and B before they are brought into contact as shown in figure 2(a). We denote the Green functions of the two disconnected systems as G_{AA}^0 and G_{BB}^0 , respectively. We then bring the systems into contact by introducing some overlap V_{ab} and V_{ba} which is localized around the interface. Then the GF of the combined system can be found by treating V as a local perturbation

$$G = (E - H_0 - V)^{-1} = ((G^0)^{-1} - V)^{-1}. \quad (11)$$

Explicitly,

$$G = \begin{pmatrix} (G_{AA}^0)^{-1} & -V_{AB} \\ -V_{BA} & (G_{BB}^0)^{-1} \end{pmatrix}^{-1}, \quad (12)$$

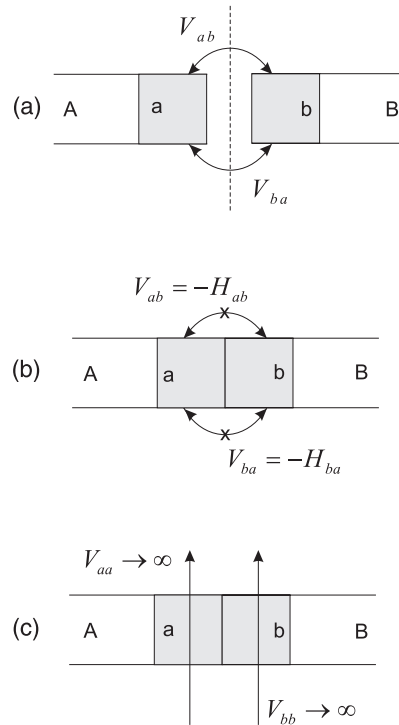


Figure 2. (a) Creating one material from two isolated pieces (A) and (B); (b) splitting the material into two isolated pieces (A) and (B) by bond cutting and (c) by atom removal.

where the matrices V_{AB} and V_{BA} are zero except for the small part around the interface, V_{ab} and V_{ba} respectively. After inverting we obtain

$$\begin{aligned}
 G_{AA} &= ((G_{AA}^0)^{-1} - V_{ab}G_{bb}^0V_{ba})^{-1}, \\
 G_{BB} &= ((G_{BB}^0)^{-1} - V_{ba}G_{aa}^0V_{ab})^{-1}, \\
 G_{BA} &= G_{Bb}^0V_{ba}G_{aA}, \\
 G_{AB} &= G_{Aa}^0V_{ab}G_{bB}.
 \end{aligned} \tag{13}$$

Here the lower case indices on the potentials, V , and on the GFs, G , indicate that the perturbation is confined to the interface between the two systems.

This simple result is quite significant. We can take into account the influence of a large part of the system on the other part via a small matrix due to the limited range of the interaction between the two parts. The term has the form of a self-energy $\Sigma = G_0^{-1} - G^{-1}$. The self-energy added to the GF of part B due to the presence of part A is $\Sigma = V_{ba}G_{aa}^0V_{ab}$. The self-energy added to the GF of part A due to the presence of part B is $\Sigma = V_{ab}G_{bb}^0V_{ba}$. Thus, for the system with a surface in figure 1(a), we can include the effects of the substrate A on the GF of the surface C through a small self-energy term that depends only on the substrate.

In a similar way, we can attach many regions A to region B . The self-energies due to the influence of different adjacent regions ($A_i, i = 1, \dots, N$) on a region B are additive if the adjacent regions overlap only with region B and not with each other

$$G_B = \left((G_B^0)^{-1} - \sum_{i=1}^N \Sigma_{A_i} \right)^{-1}, \tag{14}$$

where $\Sigma_{A_i} = V_{B_i} G_{i_i}^0 V_{i_B}$ and the superscript 0 stands for the GFs of the isolated parts. Thus, for the electron transport set-up in figure 1(b), we can map the influence of all semi-infinite leads A_i onto the GF of the probe C via small size additive terms. These self-energies can be viewed as the escape rates of the carriers in the leads.

3.2. Green function of a system with interfaces

From a different perspective equation (13) can be used to obtain all of the GF matrix elements of a system with an interface given the GFs of the separate parts. G_{AA} and G_{BB} represent the matrix elements for the total system Green function that connect sites within the A and B parts of the system respectively, and G_{AB} and G_{BA} represent matrix elements of the total GF that connect sites across the interface. Thus we can use the results of the previous subsection to derive the formulas of the surface Green function matching method [6]. To do this, we rewrite equation (11) in the form

$$\begin{aligned} G &= G^0 + G^0 T G^0, \\ T &= V(1 - G^0 V)^{-1}, \end{aligned} \quad (15)$$

where G^0 represents the GF of the separate parts and V is the localized potential at the surface. The reason we do this is that the matrix T has elements only in the surface region i.e. $T = V_{ss}(1 - G_{ss}^0 V_{ss})^{-1}$ where $s = a \cup b$ is the surface region.

Then projecting equation (15) onto the interface region and solving for T we obtain

$$T = \mathcal{G}^{0-1}(\mathcal{G} - \mathcal{G}^0)\mathcal{G}^{0-1}, \quad (16)$$

where $\mathcal{G}^0 = G_{ss}^0$ is the surface projection of the GFs of the parts and $\mathcal{G} = G_{ss}$ is the surface projection of the matched GF. Then we can obtain the matched GF of the whole system from equation (15) to be

$$\begin{aligned} G_{AA} &= G_{AA}^0 + G_{Aa}^0 \mathcal{G}_{aa}^{0-1} (\mathcal{G}_{aa} - \mathcal{G}_{aa}^0) \mathcal{G}_{aa}^{0-1} G_{aA}^0, \\ G_{AB} &= G_{Aa}^0 \mathcal{G}_{aa}^{0-1} \mathcal{G}_{ab} \mathcal{G}_{bb}^{0-1} G_{bB}^0, \end{aligned} \quad (17)$$

because the GFs of the parts do not have matrix elements across the interface. Thus, we have managed to obtain all of the GF matrix elements of the system with interfaces via the surface projection of the matched GF \mathcal{G} .

The surface projection of the matched GF can be obtained directly from equation (12) when applied to the interface region

$$\mathcal{G} = (\mathcal{G}_{aa}^0 + \mathcal{G}_{bb}^0 - \Delta\mathcal{H})^{-1}, \quad (18)$$

where $\Delta\mathcal{H}$ accounts for surface perturbation. The surface perturbation in the simplest case is just the cross coupling through the interface V . In principle it can contain more involved physics e.g. a Schottky barrier, defect or other surface effects.

3.3. Bond cutting and atom removal

In section 3.1 we showed that the BGF could be obtained by joining two SGFs. Not surprisingly, a similar technique can be used to solve the opposite problem, that is, to obtain the GF of the parts from the GF of the whole. The GF of a semi-infinite medium can be obtained by splitting the GF of the bulk into two isolated parts by either *bond cutting* or *atom removal* [16]. In order to use this approach, one removes enough atoms or cuts enough bonds at the surface that the parts on either side no longer interact with each other. The orbital cutting technique is illustrated in figure 2(b). One switches off the interaction between atoms by introducing an additional interaction of equal magnitude and opposite sign. If G^0 is the bulk GF, $V_{ab} = -H_1$,

and $V_{ba} = -H_1^\dagger$, where $H_1 = H_1^\dagger$ is the bulk Hamiltonian matrix between principal layers, equation (11) becomes

$$G = - \left[\begin{pmatrix} \ddots & & & \\ & \bar{H}_0 & \bar{H}_1 & \\ & \bar{H}_1 & \bar{H}_0 & \\ & & & \ddots \end{pmatrix} + \begin{pmatrix} \ddots & & & \\ & 0 & -\bar{H}_1 & \\ & -\bar{H}_1 & 0 & \\ & & & \ddots \end{pmatrix} \right]^{-1} = \begin{pmatrix} G_{AA} & 0 \\ 0 & G_{BB} \end{pmatrix}, \quad (19)$$

where G_{AA} and G_{BB} are by definition the surface GFs. For a practical calculation, the most convenient starting point is equation (15) where G^0 is this time the bulk GF. The GF matrix elements are

$$G_{SS} = G_{SS}^0 + G_{Ss}^0 T_{ss} G_{sS}^0, \quad (20)$$

where $S = A \cup B$. The GF matrix element between layers on the opposite side of the surface must be zero which is a useful check for ensuring that the bonds have been cut correctly.

The atom removal technique, figure 2(c), produces the same result by setting the atomic energy levels of the atoms at the interface to infinity. If G^0 is again the bulk GF and $V_{ab} = V_{ba} = 0$, $V_{aa} = V_{bb} = u \rightarrow \infty$ then equation (11) becomes

$$G = - \left[\begin{pmatrix} \ddots & & & \\ & \bar{H}_0 & \bar{H}_1 & \\ & \bar{H}_1 & \bar{H}_0 & \\ & & & \ddots \end{pmatrix} + \begin{pmatrix} \ddots & & & \\ & u & 0 & \\ & 0 & u & \\ & & & \ddots \end{pmatrix} \right]^{-1} \rightarrow \begin{pmatrix} G_{AA} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & G_{BB} \end{pmatrix}, \quad (21)$$

where we have explicitly inverted the block matrix and set $u \rightarrow \infty$. The matrices G_{AA} and G_{BB} in the final result are exactly the SGFs and thus the equivalence between the bond cutting and atom removal techniques is clear. For practical calculations one starts again from equation (15). The GF matrix elements between the surface region s and the remainder of the system $\bar{S} = (A \cup B)/s$ are

$$\begin{aligned} G_{ss} &= G_{ss}^0 + G_{s\bar{s}}^0 T_{\bar{s}s} G_{s\bar{s}}^0, \\ G_{\bar{s}s} &= G_{\bar{s}s}^0 + G_{\bar{s}s}^0 T_{\bar{s}s} G_{s\bar{s}}^0, \\ G_{s\bar{s}} &= G_{s\bar{s}}^0 + G_{s\bar{s}}^0 T_{\bar{s}s} G_{s\bar{s}}^0, \\ G_{\bar{s}\bar{s}} &= G_{\bar{s}\bar{s}}^0 + G_{\bar{s}s}^0 T_{\bar{s}s} G_{s\bar{s}}^0. \end{aligned}$$

The next step is to observe (from equation (15)) that $T \rightarrow -(G_{ss}^0)^{-1}$ when $u \rightarrow \infty$. Therefore, the matrix elements between the surface region and the remainder of the system vanish and all others are determined from the BGF matrix elements

$$\begin{aligned} G_{ss} &= G_{\bar{s}s} = G_{s\bar{s}} = 0 \\ G_{\bar{s}\bar{s}} &= G_{\bar{s}\bar{s}}^0 - G_{\bar{s}s}^0 (G_{ss}^0)^{-1} G_{s\bar{s}}^0. \end{aligned} \quad (23)$$

In general, the atom removal method involves surface matrices a factor of two smaller compared to the bond cutting method and is less prone to errors.

This technique is powerful if one already has the BGF. The BGF is normally obtained from the spectral representation of the GF [16]. However, we will show in section 6 that if we solve for the eigenfunctions and eigenvalues of the Hamiltonian we can construct the SGF directly without any need to first calculate the BGF and then cleave the medium.

4. Recursion relation for the self-energy

We can now concentrate on obtaining the GF of a single semi-infinite medium. The *recursion* or *continued fraction method* [17]¹ relates the GF of a given layer $G_{n,n}$ to the GFs of the neighbouring layers ($G_{n+1,n+1}$ and/or $G_{n-1,n-1}$). If we start with the surface layer $n = 0$ and execute this process to infinite depth we will account for the influence of all layers on the surface layer. However, in practice one must end the recursion at some finite depth, which will amount to taking into account the influence of a finite number of neighbouring layers on the surface layer. Equivalently, the recursion can be easily reformulated as *iteration* [19], where the estimation of the GF of a given layer on a given step $G^{(n)}$ is related to the estimation of the GF on the previous step $G^{(n-1)}$. However, a finite number of iterations is equivalent to executing the recursion to a certain depth. Thus, both formulations amount to truncations of the semi-infinite medium at a finite number of layers from the surface. However, the recursion or iterative methods have a significant advantage over simple truncation because successive iterates can be used to estimate the error in the truncation as we discuss in section 7.

These techniques effectively truncate the medium to a finite number of layers at the surface. This leads to a finite GF with a finite set of poles and a density of states (DoS) that is everywhere discontinuous. As more and more layers are added, however, the distribution of the poles becomes more and more dense. If a small imaginary part is added to the energy, the analytic properties of the GF calculated for the truncated system can be made to approach arbitrarily close to that of the exact GF. This is demonstrated and discussed more carefully in section 7. The three approaches of simple truncation, recursion and simple iteration are implemented differently but they share the same physical idea and their complexity increases linearly with the depth to which the calculation is carried. Due to their common convergence properties we will refer to these techniques collectively as the recursion method. Variations of these methods have been so popular that one cannot hope to list all their applications, but refer the reader to [18].

One way to find an expression for the SGF is to map the semi-infinite problem onto the problem discussed in the previous section. According to the prescription of that section, we split the semi-infinite medium into two parts. The first part is the PL at the surface and the second part is the remainder of the system which is itself semi-infinite. In the notation of the previous section, C is the isolated PL at the surface, labelled 0 hereafter, and A are the PLs numbered 1 to ∞ . Thus, an expression for the G_{CC} part of the GF can be obtained immediately from equation (14) accounting for region A through the surface self-energy term

$$G_{00}^R = \left((G_{00}^0)^{-1} - H_{01}G_{11}H_{10} \right)^{-1} = \left((G_{00}^0)^{-1} - \Sigma_{00}^R \right)^{-1}, \quad (24)$$

where $G_{ii}^0 = (ES_{ii} - H_{ii})^{-1}$ is the GF of an isolated PL.

We can continue this process of splitting the semi-infinite part into a surface PL and a semi-infinite remainder. At the next step this will amount to applying the formula recursively for G_{11} to obtain

$$\Sigma_{00}^R = H_{01} \left((G_{11}^0)^{-1} - \Sigma_{11}^R \right)^{-1} H_{10}. \quad (25)$$

In general, the self-energy in PL i due to the next PL is given by

$$\Sigma_{i,i}^R = H_{i,i+1}G_{i+1,i+1}H_{i+1,i} = H_{i,i+1} \left((G_{i+1,i+1}^0)^{-1} - \Sigma_{i+1,i+1}^R \right)^{-1} H_{i+1,i}, \quad (26)$$

¹ It should be noted that the technique described here applied to a general three-dimensional medium is more general than the version for a layered medium we utilize. The application here is a special case of their procedure.

for the case of a left surface ($i > 0$) and

$$\Sigma_{i,i}^L = H_{i,i-1} G_{i-1,i-1} H_{i-1,i} = H_{i,i-1} \left((G_{i-1,i-1}^0)^{-1} - \Sigma_{i-1,i-1}^L \right)^{-1} H_{i-1,i}, \quad (27)$$

in the case of a right surface ($i < 0$).

The same procedure can be used to determine the GF for an infinite solid. This time we split the bulk into three parts—central, left and right—and then using equation (14) the GF for the central region can be written as

$$G_{00} = \left((G_{00}^0)^{-1} - H_{01} G_{11} H_{10} - H_{0\bar{1}} G_{\bar{1}\bar{1}} H_{\bar{1}0} \right)^{-1} = \left((G_{00}^0)^{-1} - \Sigma_{00}^L - \Sigma_{00}^R \right)^{-1}, \quad (28)$$

where Σ_{00}^L and Σ_{00}^R are given by equations (27) and (26) respectively.

In the particular case in which the hopping integrals are scalar, the expression for the surface GF acquires the familiar continued fraction form

$$G_{00}^R = \frac{1}{E - H_{00} - \frac{H_{01} H_{10}}{E - H_{11} - \frac{H_{12} H_{21}}{E - H_{22} - \dots}}}, \quad (29)$$

where there are only two types of term entering this expression: $a_i = (E - H_{i,i})^{-1}$ and $b_i = H_{i,i+1} a_{i+1} H_{i+1,i}$ (in the case of a homogenous system these are all the same). A similar expression can be written for the bulk GF.

Notice that to this point we have not made an assumption that the medium is homogenous i.e. consists of repeating identical layers. In this respect, the recursive method is more general than simple iteration, accelerated iteration (section 5) or eigenvalue (section 6) methods and it is the only method to use for infinite, non-periodic solids.

If the solid consists of identical layers, then $H_{n,n+1} = H_1$, $H_{n,n-1} = H_{\bar{1}}$, $G_{nn}^0 = G_{00}^0$ and the recursive expression equation (27) becomes

$$\Sigma = H_1 \left((G_{00}^0)^{-1} - \Sigma \right)^{-1} H_{\bar{1}}. \quad (30)$$

An identical expression holds for the self-energy on the other side.

Since all quantities are in general matrices this is a non-linear system of equations which could be solved exactly using a non-linear solver such as Newton–Raphson [20]. If H_0 and H_1 are scalars, the solution is trivial. For example, for a linear chain with on-site energies ε_0 and nearest neighbour hopping matrix elements w , the recursion for the self-energy of a semi-infinite chain is

$$\Sigma = w^2 [\varepsilon - \varepsilon_0 - \Sigma]^{-1}. \quad (31)$$

Defining $\sigma = \Sigma/w$ and $x = (\varepsilon - \varepsilon_0)/w$, equation (31) can be written as $\sigma = (x - \sigma)^{-1}$ which has solutions $\sigma = \frac{x}{2} \pm \sqrt{\frac{x^2}{4} - 1}$. The choice of solution is determined by the requirement that $\text{Im } \Sigma < 0$ for a causal Green function.

A practical way of solving the non-linear equation (30) is through iteration

$$\Sigma^{(n+1)} = H_1 \left((G_{00}^0)^{-1} - \Sigma^{(n)} \right)^{-1} H_{\bar{1}}, \quad (32)$$

where $\Sigma^{(0)} = 0$. This is in essence the simple iteration technique. Another way to reformulate the recursive method into an iterative method is to use objects called *amplitude transfer matrices* (ATMs). The ATMs are defined by

$$\begin{aligned} G_{n+1,m} &= T G_{n,m} & n \geq m, \\ G_{n-1,m} &= \bar{T} G_{n,m} & n \leq m, \end{aligned} \quad (33)$$

so that they transfer the amplitude of a unit source to the right/left respectively. This is the method we will develop further since it leads naturally into the accelerated iteration technique described in the next section.

The expression for the site-diagonal part of the BGF is obtained from equation (10) when $n = 0$

$$\bar{H}_1 G_{10} + \bar{H}_0 G_{00} + \bar{H}_{\bar{1}} G_{\bar{1}0} = -1, \quad (34)$$

from which using the definition of T and \bar{T} we obtain

$$G_{00} = -(\bar{H}_0 + \bar{H}_1 T + \bar{H}_{\bar{1}} \bar{T})^{-1}. \quad (35)$$

In the case of left or right surface the corresponding Hamiltonian matrix elements are set to zero and one obtains the expressions

$$\begin{aligned} G_{00}^R &= -(\bar{H}_0 + \bar{H}_1 T)^{-1}, \\ G_{00}^L &= -(\bar{H}_0 + \bar{H}_{\bar{1}} \bar{T})^{-1}, \end{aligned} \quad (36)$$

for the right and left SGF, respectively. By comparison with equations (28) and (24), it becomes clear that the amplitude transfer matrix is proportional to the self-energy or to the SGF

$$\begin{aligned} \Sigma^R &= H_1 T & T &= G_{00}^R H_{\bar{1}}, \\ \Sigma^L &= H_{\bar{1}} \bar{T} & \bar{T} &= G_{00}^L H_1. \end{aligned} \quad (37)$$

An expression for the ATMs is obtained from equation (8) for $n' = 0$ and $n \leq 0$

$$G_{n,0} = -\bar{H}_0^{-1} (\bar{H}_{-1} G_{n-1,0} + \bar{H}_1 G_{n+1,0}). \quad (38)$$

For $n > 0$, using the definition of T , equation (33), we can obtain

$$T = -(1 + \bar{H}_0^{-1} \bar{H}_1 T)^{-1} \bar{H}_0^{-1} \bar{H}_{-1}. \quad (39)$$

Equivalently, for $n < 0$ using the definition of \bar{T} we can obtain

$$\bar{T} = -(1 + \bar{H}_0^{-1} \bar{H}_{-1} \bar{T})^{-1} \bar{H}_0^{-1} \bar{H}_1. \quad (40)$$

We can label $u_0 = -\bar{H}_0^{-1} \bar{H}_{-1}$ and $v_0 = -\bar{H}_0^{-1} \bar{H}_1$ and the expressions would be

$$\begin{aligned} T &= (1 - v_0 T)^{-1} u_0, \\ \bar{T} &= (1 - u_0 \bar{T})^{-1} v_0. \end{aligned} \quad (41)$$

Note that if $\bar{H}_{-1} = \bar{H}_1$ the equations for T and \bar{T} become equivalent.

The equations (41) are solved iteratively

$$\begin{aligned} T^{(n+1)} &= (1 - v_0 T^{(n)})^{-1} u_0, \\ \bar{T}^{(n+1)} &= (1 - u_0 \bar{T}^{(n)})^{-1} v_0, \end{aligned} \quad (42)$$

where $T^{(0)} = 0$ and $\bar{T}^{(0)} = 0$.

This method for calculating the self-energy accounts for the influence of one more layer at every order of the approximation. In other words, every iterative step adds the next order of diagrams to the GF expansion. In order to visualize this process, we assign to H_1 an arc going forward, $H_{\bar{1}}$, an arc going backwards and a vertex to \bar{H}_0 . Then the lowest order approximation to the self-energy of the surface is a loop shown in figure 3(a). In this case $T^{(0)} = u_0$, $\Sigma^{(0)} = -\bar{H}_1 \bar{H}_0^{-1} \bar{H}_{-1}$ and the diagram represents one excursion to the next PL. The first order approximation is obtained by replacing $T^{(0)}$ in equation (33). The diagrams obtained from a geometric progression with multiple $v_0 u_0$ are shown in figure 3(b). These diagrams represent all possible excursions to the neighbouring layer. The second order approximation to T has the same form where every vertex is replaced with $T^{(1)}$. As a result, in the expression for the self-energy will appear diagrams which are all possible products of two loops. A typical diagram is

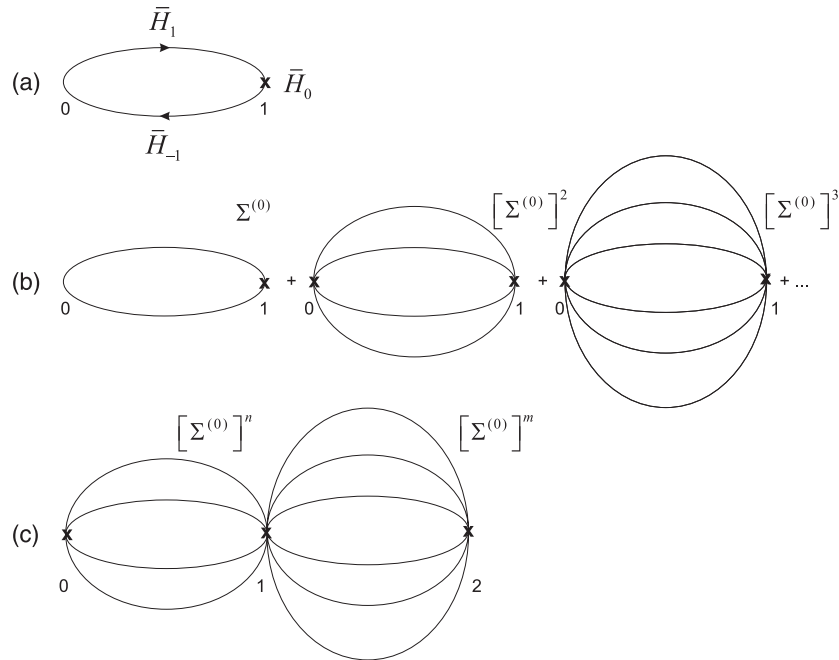


Figure 3. Diagrammatic expansion of the surface self-energy: (a) lowest order contribution, (b) one-layer contribution and (c) typical two-layer diagram.

shown in figure 3(c); it represents several excursions to layer 1 followed by excursions to layer 2. In this manner, every successive approximation will add an extra layer to the approximate expression for the self-energy. This perturbation theory relies on the smallness of the inter-layer hopping elements for convergence. However, one can imagine that the convergence would not be that good if the system has strong although limited range interaction. Summing to a given order N corresponds to truncation to N layers at the surface.

5. Accelerated iteration

In this section, we show that it is possible to accelerate the iteration scheme developed in the last section. The idea is that of layer doubling, i.e. on every step two slabs of the same number of layers are connected as opposed to connecting the slab to the neighbouring layer only. This technique can be implemented in terms of GFs [21] or ATMs [22]. We will show in section 7 that the convergence properties of the accelerated iteration scheme dramatically differ from the recursion method.

Following the derivation in [22] we apply equation (38) recursively for $G_{n-1,0}$ and $G_{n+1,0}$. We obtain an equation that relates $G_{n,0}$ to $G_{n\pm 2,0}$

$$G_{n,0} = u_1 G_{n-2,0} + v_1 G_{n+2,0} \quad (43)$$

where we labelled $u_1 = (1 - u_0 v_0 - v_0 u_0)^{-1} u_0^2$ and $v_1 = (1 - u_0 v_0 - v_0 u_0)^{-1} v_0^2$. After i iterations we obtain the general equation

$$G_{n,0} = u_i G_{n-2^i,0} + v_i G_{n+2^i,0} \quad (44)$$

where

$$\begin{aligned} u_i &= (1 - u_{i-1}v_{i-1} - v_{i-1}u_{i-1})^{-1}u_{i-1}^2, \\ v_i &= (1 - u_{i-1}v_{i-1} - v_{i-1}u_{i-1})^{-1}v_{i-1}^2. \end{aligned} \quad (45)$$

It is clear that after i iterations we have taken into account the interactions with 2^i adjacent layers.

Let us now take $n = 2^i$ and write down the first few equations

$$\begin{aligned} G_{1,0} &= u_0G_{0,0} + v_0G_{2,0} & i = 0 \\ G_{2,0} &= u_1G_{0,0} + v_1G_{4,0} & i = 1 \\ \dots & \\ G_{2^i,0} &= u_iG_{0,0} + v_iG_{2^{i+1},0} & i \end{aligned} \quad (46)$$

then plugging them in in reverse order and taking into account that we can iterate until v_i is arbitrarily small we obtain the formula for $G_{1,0} = TG_{0,0}$ where

$$T = u_0 + v_0u_1 + v_0v_1u_2 + \dots + v_0 \dots v_{i-1}u_i + \dots \quad (47)$$

Alternatively, for $n < 0$, we can show that $G_{-1,0} = \tilde{T}G_{0,0}$ where

$$\tilde{T} = v_0 + u_0v_1 + u_0u_1v_2 + \dots + u_0 \dots u_{i-1}v_i + \dots \quad (48)$$

We can explicitly show the equivalence of T 's and the self-energy by plugging the definitions of u and v into the expression for the transfer matrices. The successive approximations to the ATMs double the number of layers taken into account. Looking at equation (45) we can see that while u_0 represents a single loop, u_1 already contains all possible diagrams between two single layers. Similarly, u_i contains all possible diagrams between two super-layers each consisting of n individual layers. The role of $v_0 \dots v_{i-1}$ is to gradually build all diagrams of a super-layer of order $i - 1$ starting from a single layer; finally, u_i produces the diagrams between two such layers.

It should be noted that both this accelerated iteration technique and the iteration technique discussed in the previous section effectively calculate the self-energy for a finite system. For the simple iterative procedure, the self-energy produced after N iterations corresponds to a system with N layers while for the accelerated technique it corresponds to a system of 2^N layers. It is easy to see, for example, by examining the analytic structure of the self-energy of the simple linear chain, that the analytic structure at any level of approximation using the iterative procedures differs from that of the truly infinite system. The convergence of the Green function and self-energy will be studied in detail in section 7. In the next section we shall develop a procedure for calculating the self-energy that does not rely on an iterative procedure.

6. Eigenvalue method

The set of propagating and evanescent states of the Hamiltonian is referred to as the *complex band structure* (CBS) of the system. Knowledge of the CBS is necessary in order to construct the GF from its spectral representation [16, 23]. However, this approach has been hampered by the problem of obtaining the CBS by finding complex roots of the characteristic polynomial of the secular equation. It turns out, however, that the CBS can be obtained by diagonalizing the companion matrix of the Hamiltonian [24, 25].

For an infinite periodic system, the wavefunction on adjacent layers is related by a phase factor equation (3). Determining the phase factor from equation (9) is difficult because it involves finding roots of a complex polynomial. However, it is possible, following the

derivation in [25], to write equation (9) as an eigenvalue equation so that the phase factor is just the eigenvalue of the companion matrix of the Hamiltonian

$$\begin{pmatrix} -(\bar{H}_1)^{-1}\bar{H}_0 & -(\bar{H}_1)^{-1}\bar{H}_1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \end{pmatrix} = e^{ik_z \Delta z} \begin{pmatrix} C_0 \\ C_1 \end{pmatrix}, \quad (49)$$

where the second line is just the definition $C_0 = e^{ik_z \Delta z} C_1$. The CBS consists of both the real and complex values of k_z which satisfy equation (49). A brief summary of the method employed to solve for the CBS is given in the appendix. The eigenstates can be separated into propagating (decaying) to the right $k_z^>$ and propagating (decaying) to the left $k_z^<$ with eigenvectors $C^>$ and $C^<$ respectively.

In a similar way, following [24], we can write equation (10) for the Green function in the form

$$\begin{pmatrix} -(\bar{H}_1)^{-1}\bar{H}_0 & -(\bar{H}_1)^{-1}\bar{H}_1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} G_{n,0} \\ G_{n-1,0} \end{pmatrix} = \begin{pmatrix} G_{n+1,0} \\ G_{n,0} \end{pmatrix} + \begin{pmatrix} (\bar{H}_1)^{-1}\delta_{n,0} \\ 0 \end{pmatrix}. \quad (50)$$

Here the second line is the tautology $G_{n,0} = G_{n,0}$. Equation (50) defines a transfer matrix \mathcal{T} because it relates the GF of PLs $n-1$ to the GF at PL $n+1$ (n). The inverse of the transfer matrix \mathcal{T}^{-1} will do the opposite, i.e. it will relate the GF of PLs $n+1$ (n) to the GF at PLs n ($n-1$). Incidentally, the transfer matrix is just the companion matrix of the Hamiltonian and equation (49) can be reinterpreted as the eigenvalue equation of the transfer matrix.

We can use equation (50) to obtain an expression for the GF. If the GF satisfies equation (50) for $n > 0$ and it is normalizable at infinity

$$\mathcal{T} \begin{pmatrix} G_{n,0} \\ G_{n-1,0} \end{pmatrix} = \begin{pmatrix} G_{n+1,0} \\ G_{n,0} \end{pmatrix} \quad (51)$$

then the GF must expand only into right going states as follows:

$$\begin{pmatrix} G_{n+1,0} \\ G_{n,0} \end{pmatrix} = \begin{pmatrix} C_{n+1}^> g_1 \\ C_n^> g_1 \end{pmatrix} \quad (52)$$

where g_1 is a coefficient to be found. Vice versa, if the GF satisfies equation (50) for $n < 0$ and it is normalizable at minus infinity

$$\begin{pmatrix} G_{n,0} \\ G_{n-1,0} \end{pmatrix} = \mathcal{T}^{-1} \begin{pmatrix} G_{n+1,0} \\ G_{n,0} \end{pmatrix} \quad (53)$$

then the GF must expand only into left going states

$$\begin{pmatrix} G_{n,0} \\ G_{n-1,0} \end{pmatrix} = \begin{pmatrix} C_n^< g_2 \\ C_{n-1}^< g_2 \end{pmatrix} \quad (54)$$

where g_2 is another coefficient to be determined. Notice that the same wavefunction coefficients C are used in both expressions because \mathcal{T} and \mathcal{T}^{-1} must have the same set of eigenvectors since every matrix commutes with its inverse. For $n = 0$ we have

$$\begin{pmatrix} -(\bar{H}_1)^{-1}\bar{H}_0 & -(\bar{H}_1)^{-1}\bar{H}_1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} C_0^< g_2 \\ C_1^< g_2 \end{pmatrix} = \begin{pmatrix} C_1^> g_1 \\ C_0^> g_1 \end{pmatrix} + \begin{pmatrix} (\bar{H}_1)^{-1} \\ 0 \end{pmatrix}. \quad (55)$$

The second line of this equation yields $g_1 = (C_0^>)^{-1} C_0^< g_2$ and the first line

$$\bar{H}_0 C_0^< g_2 + H_1 C_1^< g_2 + H_1 C_1^> g_1 = 1 \quad (56)$$

which can be solved for g_2 and for the GF, respectively,

$$G_{00} = -C_0^< g_2 = -(\bar{H}_0 + H_1 C_1^< (C_0^<)^{-1} + H_1 C_1^> (C_0^>)^{-1})^{-1}. \quad (57)$$

Again, by comparison with equation (35) one can obtain the expressions for the amplitude transfer matrices to be

$$\begin{aligned} T &= C_1^> (C_0^>)^{-1} \\ \bar{T} &= C_1^< (C_0^<)^{-1}. \end{aligned} \quad (58)$$

In this notation the ATMs have very clear physical meaning; they destroy amplitude at the reference layer and create the corresponding amplitude one layer to the right/left. One can check that these are the same amplitude transfer matrices we had before, i.e. they satisfy the formal definition for the ATMs. If we label Λ^{\gtrless} to be the diagonal matrix of the eigenvalues of T then

$$\begin{aligned} T &= C_0^> \Lambda^> (C_0^>)^{-1} \\ \bar{T} &= C_0^< (\Lambda^<)^{-1} (C_0^<)^{-1} \end{aligned} \quad (59)$$

using equation (3) again. Applying the ATMs to the wavefunction yields

$$\begin{aligned} T C_0^> &= C_0^> \Lambda^> = C_1^> \\ \bar{T} C_0^< &= C_0^< (\Lambda^<)^{-1} = C_1^< \end{aligned} \quad (60)$$

in other words the action of the transfer matrices is like multiplying the states (columns of C) by their respective phases.

Similarly, we can find that the action on the GFs

$$\begin{aligned} T G_{n,0} &= T C_n^> g_1 = C_{n+1}^> g_1 = G_{n+1,0} & n \geq 0 \\ \bar{T} G_{n,0} &= C_n^< g_2 = C_{n-1}^< g_2 = G_{n-1,0} & n \leq 0 \end{aligned} \quad (61)$$

relates the GFs of adjacent layers as expected. Thus the eigenvalue approach yields the ATMs which can be used to construct the surface self-energies and Green function as demonstrated in section 4.

7. Convergence

In the simple truncation approach, the GF is calculated using a finite number of layers close to the surface. In the recursion (simple iteration) and the accelerated iteration methods, performing a given finite number of steps also amounts to truncating the semi-infinite medium at a given number of layers from the interface. However, unlike simple truncation, the error of the truncation can be estimated. In this respect these methods represent controlled approximations. The ATM is approximated by a series $T = \sum_i a_i$ and $T^n = \sum_{i=1}^n a_i$ is the n th partial sum of the series. The Cauchy convergence criterion states that $\forall \epsilon > 0$ and $\forall p > 0 \exists n > 0 : |T^{n+p} - T^n| < \epsilon$. In particular ($p = 1$), this means that $\forall \epsilon > 0 \exists n > 0 : |a_n| < \epsilon$. Thus the condition that the term in the series $a_n \rightarrow 0$ is a necessary but not a sufficient condition for convergence. However, checking the Cauchy criterion would require the ability to calculate T to arbitrary order; therefore, in practical calculations the weaker criterion is used.

The density of states of a finite number of layers consists of discrete poles which approximate a branch cut only if we introduce broadening via a small complex part added to the energy. The small complex part is also necessary for numerical stability. As an illustration, we consider a one-dimensional (1D) semi-infinite chain of atoms with only one band. The analytical results for such a system are well known [26]. The parameters of the model are $H_0 = t$ and $H_{\bar{1}} = H_1 = w$. In figures 4(a) and (b), the imaginary and real parts of the surface GF of this chain are shown. The converged result is compared to the result obtain using the simple iteration method (recursion and simple truncation have the same convergence properties) for small number of iterations ($n = 30$). In figures 4(c) and (d), the converged

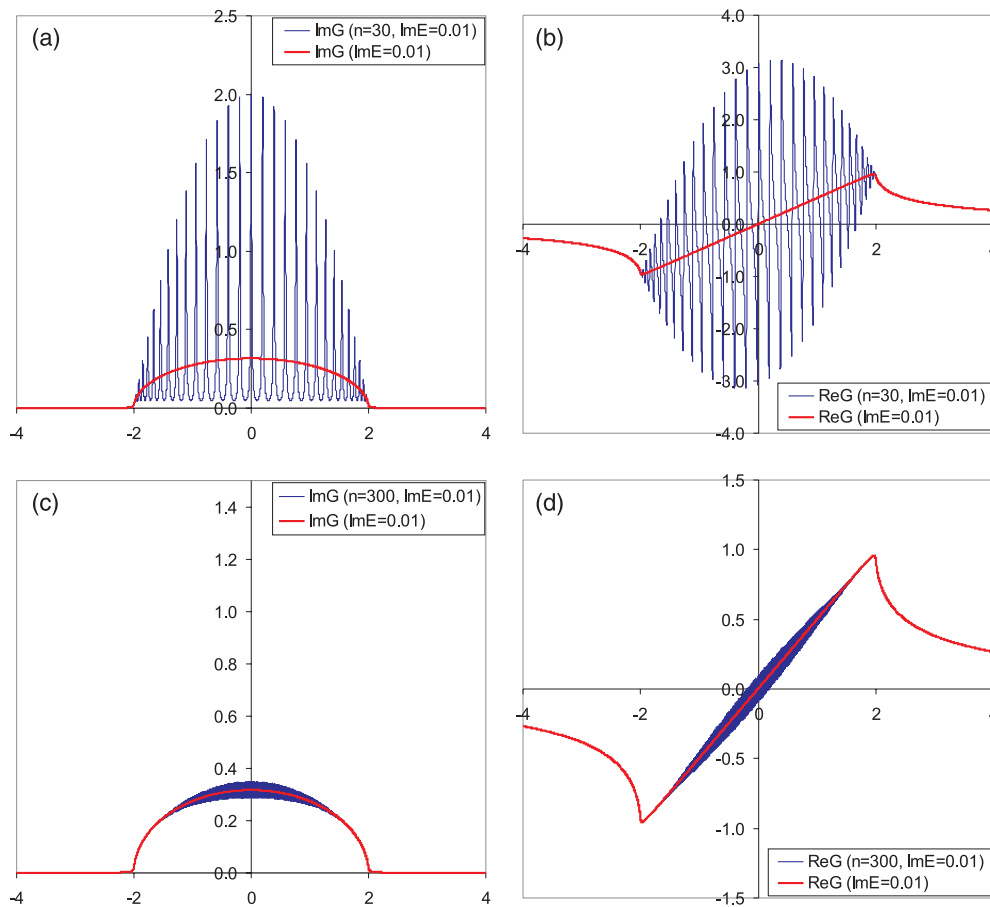


Figure 4. Surface Green function versus energy for a 1D semi-infinite chain: (a) $-\text{Im} G (n = 30)$, (b) $\text{Re} G (n = 30)$, (c) $-\text{Im} G (n = 300)$, (d) $\text{Re} G (n = 300)$ after n iterations of the recursive approximation; an imaginary part $\delta = 10^{-2}$ is added to the energy. The result is compared to the result converged to precision $\epsilon = 10^{-6}$.

GF is compared to the GF obtained after a much larger number of iterations ($n = 300$). The imaginary part of the energy is $\delta = 0.01$ and the convergence criterion is $\epsilon = 10^{-6}$. The graphs are shown in dimensionless units $g = wG$ versus $x = (E - t)/w$ where G and E are the GF and the energy respectively.

In figures 5(a) and (b), the imaginary and real parts of the bulk GF are shown. The converged result is compared with the result obtained using this accelerated iteration method for small number of iterations ($n = 5$). In figures 5(c) and (d), the same functions are plotted after a few more iterations of the iterative method. Compared to the recursive method (figure 4), fewer iterations are needed to obtain convergence.

The result of the eigenvalue method is indistinguishable from the converged result in figures 4 and 5. The eigenvalue expression for the ATMs represents a continued fraction that is not truncated.

The analytic result is approached as we let the imaginary part of the energy $\delta \rightarrow 0$. Use of large δ converges the result much faster but may lead to inaccurate results. This is especially true for the region around singularities which is illustrated in figures 6(a) and (b) for

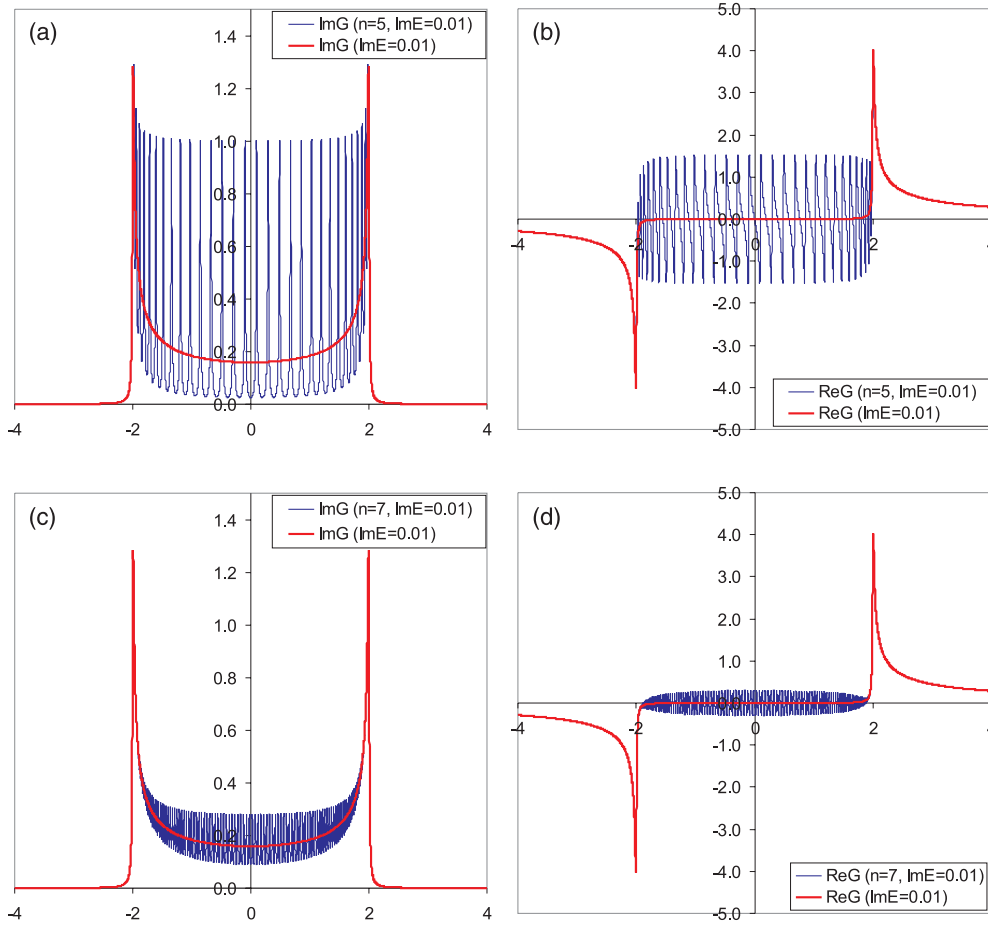


Figure 5. Bulk Green function versus energy of an infinite 1D chain: (a) $-\text{Im} G (n = 5)$, (b) $\text{Re} G (n = 5)$, (c) $-\text{Im} G (n = 7)$, (d) $\text{Re} G (n = 7)$ after n iterations of the iterative method; an imaginary part $\delta = 10^{-2}$ is added to the energy. The result is compared to the result converged to precision $\epsilon = 10^{-6}$.

the imaginary and real parts of the surface GF of the 1D chain calculated using the recursion method. The same behaviour is observed for the imaginary and real parts of the bulk GF calculated using the iteration method, as illustrated in figures 6(c) and (d). In these figures two different imaginary parts of the energy are compared. The deviation of the result obtained with $\delta = 10^{-1}$ is much less than the result obtained with $\delta = 10^{-2}$ for the same number of iterations (figures 4 and 5), therefore closer to convergence. However, the result is further away from the exact result. Thus, higher precision requires a lower imaginary part to the energy which, however, makes the convergence much slower.

The imaginary part of the energy determines the broadening and the height of the peaks around the poles. Roughly speaking, convergence is reached when the poles overlap to cover the energy interval separating the poles $\delta \simeq W/N_{\text{poles}}$ where N_{poles} is the number of poles and W is a measure of the band width. More precisely, δ should be greater than $1/D^0(E_0)$ where $D(E_0)$ is the minimum density of states. If we decrease δ exponentially holding the precision constant we need an exponentially growing number of poles to maintain a given level

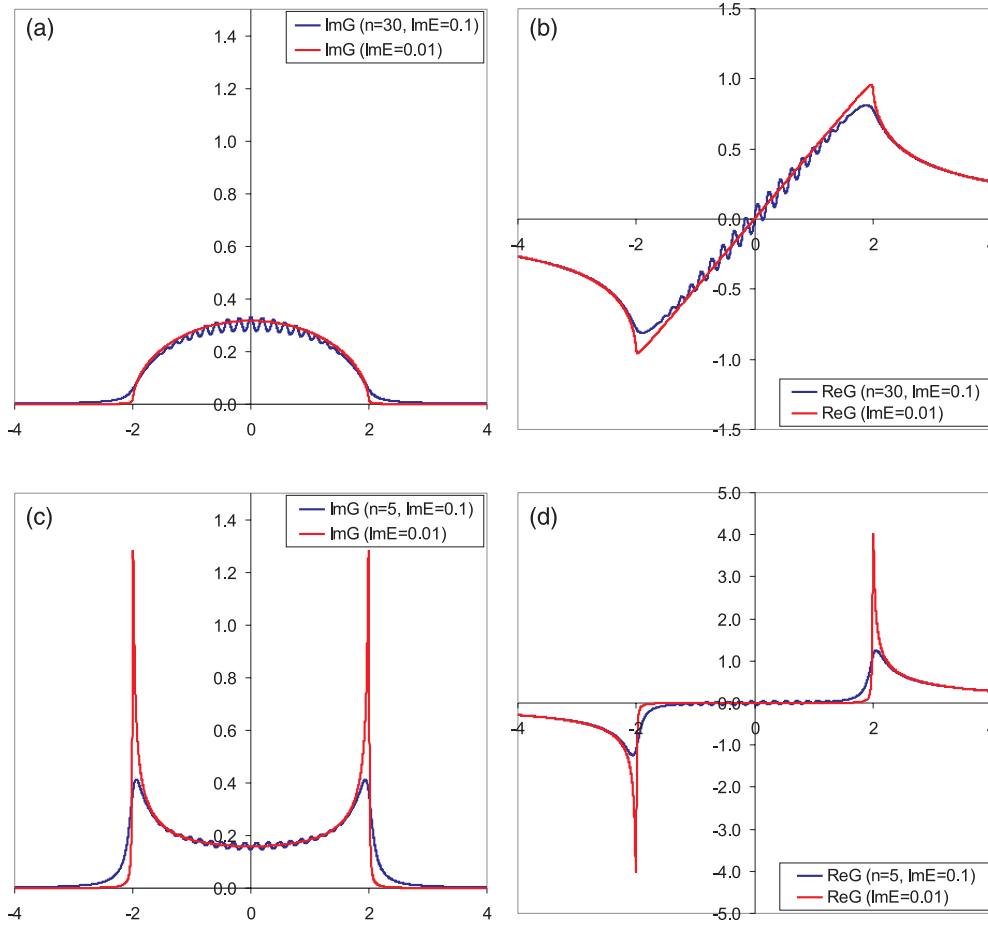


Figure 6. Green functions versus energy for a 1D chain: (a) $-\text{Im} G$ for surface ($n = 30$), (b) $\text{Re} G$ for surface ($n = 30$) after n iterations of the recursive method, (c) $-\text{Im} G$ ($n = 5$), (d) $\text{Re} G$ ($n = 5$) after n iterations of the iterative method; an imaginary part $\delta = 10^{-1}$ is added to the energy. The results are compared to the result converged to precision $\epsilon = 10^{-6}$ and $\delta = 10^{-2}$.

of convergence. The convergence of the recursive and iterative methods at one energy point $x = 0$ is shown in table 1 for different values of δ . As we see, the exponential increase in precision leads to exponential slowing down of the recursive method because the number of poles scales linearly with the number of iterations (N for the surface, $2N$ for the bulk GF). The iterative method suffers only linear decrease in performance because the number of poles increases exponentially (2^N for the surface, 2^{N+1} for the bulk GF). The eigenvalue method remains of constant complexity with increasing precision because the direct methods used for solving for the CBS depend only on the rank of the Hamiltonian matrix.

8. Off-diagonal matrix elements

From the on-site matrix element of the GF one can obtain all necessary GF matrix elements [22]. In bulk, due to the translational invariance ($G_{nm} = G_{00}$), we do not need more than G_{00} and T and \bar{T} to obtain all GF matrix elements using the definition of the ATMs, equation (33). In the

Table 1. Number of iterations needed for convergence versus imaginary part of the energy at a fixed precision $\epsilon = 10^{-6}$.

δ	0.1	0.01	0.001	0.0001	0.00001
Recursive	151	1519	15 201	152 015	1520 158
Iterative	10	13	16	20	23
Eigenvalue	1	1	1	1	1

case of a surface, the G_{nn} deeper in the material are not the same as those closer to the surface. We can devise recurrence relations that will allow us to obtain all diagonal matrix elements and from them all off-diagonal matrix elements. Starting from equation (8) for $n = n' > 0$

$$\bar{H}_0 G_{n,n} = -1 - \bar{H}_1 G_{n+1,n} - \bar{H}_{-1} G_{n-1,n} \tag{62}$$

we can write this using the definition of T as

$$(\bar{H}_0 + \bar{H}_1 T) G_{n,n} = -1 - \bar{H}_{-1} G_{n-1,n}, \tag{63}$$

from which, using the definition of the SGF equation (24) and the relation of the ATM to the SGF equation (37), we obtain

$$G_{n,n} = \begin{cases} G_{00} + T G_{n-1,n-1} S & n \geq 1 \\ G_{00} + \bar{T} G_{n+1,n+1} \bar{S} & n \leq -1 \end{cases} \tag{64}$$

where the second relation was obtained in a similar way. The matrices S and \bar{S} are the analogues of T and \bar{T} used to raise or lower the second index of the GF. They are defined as follows:

$$\begin{aligned} G_{n,m+1} &= G_{n,m} S & n \leq m, \\ G_{n,m-1} &= G_{n,m} \bar{S} & n \geq m. \end{aligned} \tag{65}$$

An expression for the BGF and the SGF can be derived in terms of S and \bar{S} using a similar argument to that in the derivation of equations (35) and (36). The bulk GF becomes

$$G_{00} = -(\bar{H}_0 + S \bar{H}_1 + \bar{S} \bar{H}_{-1})^{-1}, \tag{66}$$

and the surface GFs to the left and right are

$$\begin{aligned} G_{00}^R &= -(\bar{H}_0 + S \bar{H}_1)^{-1}, \\ G_{00}^L &= -(\bar{H}_0 + \bar{S} \bar{H}_{-1})^{-1}. \end{aligned} \tag{67}$$

Comparing equation (35) with (66) one can find a simple relationship between the two kinds of ATM

$$S = \bar{H}_1 T (\bar{H}_1)^{-1} \quad \bar{S} = \bar{H}_{-1} \bar{T} (\bar{H}_{-1})^{-1}. \tag{68}$$

Finally, we can use $T(\bar{T})$ and $S(\bar{S})$ as raising and lowering operators on $G_{n,n}$ to obtain all GF matrix elements.

9. Conclusions

We consider various methods of calculating the Green function of surfaces. In our unified treatment it becomes easy to see that the methods essentially amount to different ways of calculating the self-energy associated with the surface. The recursive method being the oldest is also the slowest but it has an advantage that it could be applied to a non-homogeneous system virtually constructing the GF of the system layer by layer. The iterative method provides the possibility for an accelerated scheme which is very practical and convenient if the problem is formulated completely in the GF formalism. It bypasses the calculation of wavefunctions and the complex band structure. If the wavefunctions are needed the eigenvalue method presents the better alternative because it allows the calculation of the wavefunctions and the GF in one step. It can also be the fastest method for higher precision applications.

Acknowledgments

This research was supported in part by the Oak Ridge Institute for Science and Education and the Oak Ridge National Laboratory. Work at the University of Alabama was supported in part by National Science Foundation MRSEC grant number DMR0213985 and by the Office of Naval Research N00014-03-100358.

Appendix. Complex band structure

According to Bloch's theorem, the wavefunctions on adjacent principal layers are related by $\psi(\mathbf{r} + \Delta z) = e^{ik_z \Delta z} \psi(\mathbf{r})$ where k_z is the component of the crystal momentum in the z direction. This relation implicitly assumes that the principal layers are chosen so that there is translational invariance in the z direction with the period of the principal layer. This is usually the case. If this assumption is violated an additional phase factor involving $\bar{\mathbf{k}}$ must be included as discussed below.

In an infinite solid only Bloch states with real crystal momentum can exist. States with complex crystal momentum would diverge at infinity. In a finite solid, however, both propagating and evanescent states (with complex crystal momentum) can exist. In this appendix we shall review techniques for finding the complex band structure, i.e. all values of crystal momentum, both real and complex. The complex crystal momenta are important for determining the GF near a surface or interface.

The problem of determining the CBS has been revisited in the literature many times; here we follow [25]. We have reformulated this approach for principal layers. The objective is to find all (both real and complex) eigenvalues of the Schrödinger equation (4) in the basis of planar orbitals (1). We construct the companion matrix of the matrix \bar{H} by the following manipulations:

$$-(\bar{H}_1)^{-1} \bar{H}_0 C_0 - (\bar{H}_1)^{-1} \bar{H}_1 C_1 = C_1 = e^{ik_z} C_0, \quad (\text{A.1})$$

where the last equality follows from equation (3). We can rewrite these equations in the equivalent form,

$$\begin{pmatrix} -(\bar{H}_1)^{-1} \bar{H}_0 & -(\bar{H}_1)^{-1} \bar{H}_1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \end{pmatrix} = e^{ik_z} \begin{pmatrix} C_0 \\ C_1 \end{pmatrix}, \quad (\text{A.2})$$

which defines the eigenvalue problem for the companion matrix of the Hamiltonian which is also identified as the transfer matrix \mathcal{T} in section 6. The complex band structure is derived from the eigenvalues $\{\lambda_n\}$ of the companion matrix and $k_z^n = -i \log \lambda_n$.

This approach has the problem that sometimes the matrix \bar{H}_1 may not have an inverse. In that case we can solve the generalized eigenvalue problem

$$\begin{pmatrix} \bar{H}_0 & \bar{H}_1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \end{pmatrix} = e^{ik_z} \begin{pmatrix} -\bar{H}_1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} C_0 \\ C_1 \end{pmatrix}. \quad (\text{A.3})$$

The principal layer is defined as a unit cell in the direction perpendicular to the layers. If the number of topologically different layers is larger than the range of the interaction there will be many zero blocks in the Hamiltonian matrix elements. In these cases it is convenient to keep the PL just large enough so there is no interaction beyond neighbouring PLs although now the neighbouring PLs would be displaced by some vector $\Delta \bar{\mathbf{d}}$ with respect to each other. In this case the wavefunction will acquire an additional phase factor going from layer to layer $C_1 = e^{i(k_z + \bar{\mathbf{k}} \cdot \Delta \bar{\mathbf{d}})} C_0$, where $\bar{\mathbf{k}}$ is the wavevector in the plane of the layers. In this case, equation (A.2) still applies, only the eigenvalues will be modified $k_z^n = -i \log \lambda_n - \bar{\mathbf{k}} \cdot \Delta \bar{\mathbf{d}}$.

The eigenvalues come in pairs: propagating or decaying to the right $k^>$ with $\text{Im}(k_z) > 0$ if $k_z \in \mathbb{C}$ or $V_g > 0$ if $k_z \in \mathbb{R}$ and propagating or decaying to the left ($k^<$) with $\text{Im}(k_z) < 0$ or $V_g < 0$. In practical calculations, one has to add a small imaginary part to the energy in order to make the matrix factorization stable. This will displace the propagating states off the real axis. Fortunately, one can predict the direction of the change. If we add $\delta > 0$ the eigenstate will change as follows:

$$|\exp(ik_z(\epsilon + i\delta))| = \left| \exp\left(i\left\{k_z(\epsilon) + \frac{\partial k_z(\epsilon)}{\partial \epsilon}(i\delta) + \dots\right\}\right) \right| = \left| \exp\left(-\frac{\delta}{v_g}\right) \right|; \quad (\text{A.4})$$

therefore, the states with $v_g > 0$ will acquire a positive imaginary part and the states with $v_g < 0$ a negative imaginary part.

References

- [1] Negele J W and Orland H 1998 *Quantum Many-Particle Systems* (New York: Perseus Book Group)
- [2] Stone A D and Szafer A 1988 *IBM J. Res. Dev.* **32** 384
- [3] Butler W H, Zhang X-G and Nicholson D M C 1995 *Phys. Rev. B* **52** 13399
- [4] Butler W H 1985 *Phys. Rev. B* **31** 3260
- [5] Elliott R, Krumhansl J A and Leath P L 1974 *Rev. Mod. Phys.* **46** 465
- [6] García-Moliner F and Velasco V R 1992 *Theory of Single and Multiple Interfaces* (Singapore: World Scientific)
- [7] Datta S 1997 *Electronic Transport in Mesoscopic Systems* (Cambridge: Cambridge University Press)
- [8] Tsymbal E Y and Pettifor D G 2001 *Solid State Phys.* **56** 113
- [9] Mathon J and Umerski A 2001 *Phys. Rev. B* **63** 220403(R)
- [10] Maradudin A A 1971 Theory of lattice dynamics in the harmonic approximation *Solid State Physics Suppl.* 3 (New York: Academic)
- [11] Tewary V K and Thomson R 1992 *J. Mater. Res.* **7** 1018
- [12] Ting T C T 1996 *Anisotropic Elasticity* (Oxford: Oxford University Press)
- [13] Ferreira M S, Bauer G E W and Wapenaar C P A 2002 *Ultrasonics* **40** 355
- [14] Gonis A, Zhang X-G, MacLaren J M and Crampin S 1990 *Phys. Rev. B* **42** 3798
- [15] Wenzien B, Kudrnovský J, Drchal V and Šob M 1989 *J. Phys.: Condens. Matter* **1** 9893
- [16] Pollmann J and Pantelides S 1978 *Phys. Rev. B* **18** 5524
- [17] Haydock R, Heine V and Kelly M 1972 *J. Phys. C: Solid State Phys.* **5** 2845
Haydock R, Heine V and Kelly M 1975 *J. Phys. C: Solid State Phys.* **8** 2591
- [18] Pettifor D G and Weaine D L (ed) 1985 *The Recursion Method and its Applications* (Berlin: Springer)
- [19] Falicov L and Yndurain F 1975 *J. Phys. C: Solid State Phys.* **8** 147
- [20] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1997 *Numerical Recipes in Fortran 77* (Cambridge: Cambridge University Press)
- [21] Lannoo M and Friedel P 1991 *Atomic and Electronic Structure of Surfaces* (Berlin: Springer)
- [22] López-Sancho M, López-Sancho J and Rubio J 1984 *J. Phys. F: Met. Phys.* **14** 1205
López-Sancho M, López-Sancho J and Rubio J 1985 *J. Phys. F: Met. Phys.* **15** 851
- [23] Allen R 1979 *Phys. Rev. B* **20** 1454
- [24] Lee D H and Joannopoulos J D 1981 *Phys. Rev. B* **23** 4988
Lee D H and Joannopoulos J D 1981 *Phys. Rev. B* **23** 4997
- [25] Chang Y-C and Schulman J 1982 *Phys. Rev. B* **25** 3975
- [26] Gonis A 1992 *Green Functions for Ordered and Disordered Systems* (Amsterdam: North-Holland)