Tight-binding surface states in finite crystals

2000 J. Phys.: Condens. Matter 12 5565

(http://iopscience.iop.org/0953-8984/12/26/304)

J. Phys.: Condens. Matter **12** (2000) 5565–5573. Printed in the UK PII: S0953-8984(00)11193-2

Tight-binding surface states in finite crystals

J Heinrichs

Institut de Physique B5, Université de Liège, Sart-Tilman, B-4000 Liège, Belgium

Received 20 January 2000

Abstract. The electronic states of a finite crystal are studied using Goodwin's model of a tight-binding linear chain of N one-level atoms with nearest-neighbour overlap. Using a transfer matrix approach we obtain the explicit form of the secular equation which correctly yields N eigenvalues in the interval $(0, \pi)$ of wavenumber q, unlike Goodwin's equation which involves spurious solutions at $q = 0$ and $q = \pi$. We present a new general analysis of bulk- and surface-state eigenvalues as a function of the parameter ε_0/γ describing the difference (ε_0) of Coulomb integrals for surface and bulk atoms relative to the overlap integral γ. We identify four distinct domains of values of $|\varepsilon_0/\gamma|$ in three of which one or two surface states of different origins exist, which we determine explicitly. Our discussion is valid for both signs of ε_0/γ and differs considerably in detail from Goodwin's analysis. In particular, it does not require distinct analyses for chains with even and odd numbers of sites.

1. Introduction

The general properties of electronic surface states in crystals derived for simple models in the 1930s [1–6] are the basis for the more realistic studies of such states which have appeared since: their energy levels lie within the forbidden gaps and they have an imaginary component of the wavevector perpendicular to the boundary, which determines the spatial decay of the wavefunction inside the crystal. A surface-state wavefunction also decays exponentially in free space, in contrast to conventional Bloch states whose wavefunctions extend throughout the crystal but vanish essentially outside. Surface states owe their existence to the deviation of the one-electron potential near the surface from its form deeper inside the crystal which, of course, also perturbs the Bloch-like extended states. The general properties of surface states have been established in two basic approximations for crystal electronic structure, namely the nearly free-electron [3, 4] and the tight-binding [5, 6] approximations. Physically, the binding of an electron at a surface may be visualized as arising from a repeated process whereby an electron wave scattered by a hard-wall potential supposed to mimic the surface is totally reflected by the periodic one-electron potential deeper inside the crystal.

This paper concerns Goodwin's pioneering work on the tight-binding approach to electronic surface states in crystals [5, 6]. In reference [5], Goodwin studied a model of a finite linear chain consisting of N identical equally spaced one-level atoms. In reference [6] he extended his work on finite linear chains to the study of surface states in a cubic crystal slab consisting of N infinite atomic layers and, furthermore, he considered semi-infinite linear chains in the framework of non-degenerate as well as degenerate two-band electronic models. Goodwin's calculations for crystals in two-band models have been further generalized by Artmann [7] and the criteria for the existence of surface states in various lattices, using his approach, have been elaborated on by Baldock [8]. The continuing relevance of Goodwin's

0953-8984/00/265565+09$30.00

tight-binding studies of surface states is demonstrated by their fundamental role in two important reviews [9, 10] as well as e.g. by a more recent realistic study of surface and bulk electronic states at high-symmetry points of the two-dimensional Brillouin zone of semi-infinite crystals [11].

By reanalysing Goodwin's tight-binding model for the electronic eigenstates in a finite linear chain using a transfer matrix approach [12], we have recently discovered an inconsistency in his work, namely the occurrence of eigenvalues corresponding to the exact band edges of an infinite chain. Our treatment resolves this inconsistency. Furthermore, a new analysis of the secular equation for the eigenvalues of the chain has led us to a more detailed and satisfactory picture of the properties of surface states in Goodwin's model. Unlike Goodwin's discussion, the present one is valid for both signs of the difference between Coulomb integrals for surface and bulk atoms and does not require distinct analyses for even and odd N. In view of the importance of Goodwin's work [5, 6] in the surface physics literature and, in particular, its pedagogical value, it seems useful to present our complementary results, which is the object of the following three sections.

2. Electronic states in a linear tight-binding chain

In Goodwin's model the electronic states for a chain of N one-level atoms at $a, 2a, \ldots, Na$ (with lattice parameter $a = 1$) are determined by the set of difference equations

$$a_{n+1} + a_{n-1} + \frac{\varepsilon}{\gamma} a_n = 0 \qquad n = 2, 3, \ldots, N-1 \tag{1}$$

$$a_2 + \frac{(\varepsilon - \varepsilon_0)}{\gamma} a_1 = 0 \tag{2}$$

$$a_{N-1} + \frac{(\varepsilon - \varepsilon_0)}{\gamma} a_N = 0 \tag{3}$$

where

$$\varepsilon = E - E_0 + \alpha \tag{4}$$

$$\varepsilon_0 = \alpha - \alpha'. \tag{5}$$

The eigenfunction corresponding to the energy E is expanded in terms of atomic s states of energy E_0 centred at sites n, with amplitudes $\{\ldots, a_n, \ldots\}$. As is necessary for the existence of surface states (see section 1), one assumes different values of the Coulomb integrals for atoms situated inside (α) and at the ends of the chain (α'). Finally γ denotes the unperturbed nearest-neighbour overlap integral.

As mentioned in section 1, we obtain the energy eigenvalues E and the corresponding amplitudes $a_n(E)$ in (1)–(3) using a transfer matrix, namely the 2×2 unimodular matrix

$$\widehat{T} = \begin{pmatrix} -\varepsilon/\gamma & -1 \\ 1 & 0 \end{pmatrix} \tag{6}$$

which allows us to rewrite (1) as

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \widehat{T} \begin{pmatrix} a_n \\ a_{n-1} \end{pmatrix} \qquad n = 2, 3, \ldots, N-1. \tag{7}$$

Iteration of (7) in terms of fixed values a_1 and a_2 for the amplitudes at sites 1 and 2 yields

$$\begin{pmatrix} a_{n+1} \\ a_n \end{pmatrix} = \widehat{T}^{n-1} \begin{pmatrix} a_2 \\ a_1 \end{pmatrix} \qquad n = 2, 3, \ldots, N-1. \tag{8}$$

The secular equation giving the allowed energy eigenvalues then follows by combining (8) for $n = N - 1$ with (2) and (3) which yields (with $\widehat{T}_{ij}$ the ij-element of $\widehat{T}$)

$$(\widehat{T}^{N-2})_{21} - \frac{\gamma}{\varepsilon - \varepsilon_0}(\widehat{T}^{N-2})_{22} = -\frac{(\varepsilon - \varepsilon_0)}{\gamma}(\widehat{T}^{N-2})_{11} + (\widehat{T}^{N-2})_{12}. \tag{9}$$

On the other hand, the amplitude at site n is given by

$$\frac{a_n}{a_1} = -\frac{(\varepsilon - \varepsilon_0)}{\gamma}(\widehat{T}^{n-2})_{11} + (\widehat{T}^{n-2})_{12} \qquad n = 2, 3, \ldots, N. \tag{10}$$

The explicit calculation of the elements of $\widehat{T}^m$, $m = n - 2$ or $N - 2$, in (9) and (10) proceeds as follows. We first find the eigenvalues λ_1 and λ_2 of $\widehat{T}$:

$$\lambda_{1,2} = -\frac{1}{2}\left(\frac{\varepsilon}{\gamma} \pm \sqrt{\left(\frac{\varepsilon}{\gamma}\right)^2 - 4}\right). \tag{11}$$

This expression simplifies to

$$\lambda_{1,2} = \mathrm{e}^{\pm \mathrm{i}q} \tag{12}$$

on putting

$$\varepsilon = -2\gamma \cos q \tag{13}$$

where q is either real or a complex parameter of the form

$$q = k\pi + \mathrm{i}\kappa \qquad k = 0, 1, 2, \ldots \tag{14}$$

since ε defined in (4) is real. The matrix $\widehat{U}$ which diagonalizes the transfer matrix $\widehat{T}$ is then

$$\widehat{U} = \begin{pmatrix} \mathrm{e}^{\mathrm{i}q} & \mathrm{e}^{-\mathrm{i}q} \\ 1 & 1 \end{pmatrix} \qquad \text{with } \widehat{U}^{-1} = \frac{1}{2\mathrm{i}\sin q}\begin{pmatrix} 1 & -\mathrm{e}^{-\mathrm{i}q} \\ -1 & \mathrm{e}^{\mathrm{i}q} \end{pmatrix}. \tag{15}$$

Since $\widehat{U}^{-1}\widehat{T}\widehat{U}$ is diagonal, $(\widehat{U}^{-1}\widehat{T}\widehat{U})_{ij} = \lambda_i \delta_{ij}$, $i = 1, 2$, we obtain from (13) and (15)

$$\widehat{T}^m = \widehat{U}(\widehat{U}^{-1}\widehat{T}\widehat{U})^m\widehat{U}^{-1} = \frac{1}{\sin q}\begin{pmatrix} \sin(m+1)q & -\sin mq \\ \sin mq & -\sin(m-1)q \end{pmatrix}. \tag{16}$$

Finally, on inserting the elements of (16) for $n = N - 2$ in (9) we get

$$\frac{1}{\sin q}\left[\left(\frac{\varepsilon - \varepsilon_0}{\gamma}\right)^2 \sin(N-1)q + 2\left(\frac{\varepsilon - \varepsilon_0}{\gamma}\right)\sin(N-2)q + \sin(N-3)q\right] = 0 \tag{17}$$

which is the secular equation for the allowed values of q giving the quantized energy levels of the chain,

$$E = E_0 - \alpha - 2\gamma \cos q. \tag{18}$$

Similarly we obtain from (10)

$$a_n = \frac{1}{\sin q}\left[\sin nq + \frac{\varepsilon_0}{\gamma}\sin(n-1)q\right]a_1 \qquad n = 2, 3, \ldots, N. \tag{19}$$

By simple algebra one verifies that (17) coincides with the secular equation obtained by Goodwin (equation (13) of reference [5]) up to the overall factor $1/\sin q$, noting that Goodwin uses the definition $\varepsilon = 2\gamma \cos q$ instead of (13). Similarly, the expression (19) for the amplitude a_n when re-expressed in terms of imaginary exponentials, i.e.

$$a_n = \frac{1}{2\mathrm{i}\sin q}\left[\left(1 + \frac{\varepsilon_0}{\gamma}\mathrm{e}^{-\mathrm{i}q}\right)\mathrm{e}^{\mathrm{i}nq} - \left(1 + \frac{\varepsilon_0}{\gamma}\mathrm{e}^{\mathrm{i}q}\right)\mathrm{e}^{-\mathrm{i}nq}\right] \qquad n = 2, 3, \ldots, N \tag{20}$$

differs from the corresponding expressions (10) and (12) of [5] in terms of constants A and B by the presence of the additional factor $1/\sin q$. The effects of the factor $1/\sin q$ in (17)

and (19) are the following. On the one hand, it ensures that the secular equation (17) is free of the spurious solutions $q = 0$ and $q = \pi$ corresponding to the band-edge eigenvalues for an infinite chain obtained by Goodwin [5]. On the other hand, it implies that the amplitudes (19) and (20) are finite for $q \rightharpoonup 0$ and $q \to \pi$, while Goodwin's amplitudes vanish in these limits. We recall that Goodwin used this property of his amplitudes a_n as a physical argument for excluding his spurious solution $q = 0$ (and $q = \pi$) of the secular equation.

3. Detailed analysis of eigenvalues

As indicated above, we wish to present a more general and detailed analysis of the roots of the secular equation (17) than that given by Goodwin [5]. In particular, we shall derive various threshold values of ε_0/γ for the appearance of surface states and discuss the number and the properties of these states between the successive thresholds.

We find that a very useful way of visualizing the bulk- and surface-state eigenvalues of the chain results from transforming (17) into the equivalent form (with $t = \varepsilon_0/\gamma$)

$$\frac{[(2\cos q + t)^2 + 1]\cos q - 2(2\cos q + t)}{1 - (2\cos q + t)^2} = \sin q \cot(N-2)q \tag{21}$$

using (13).

In the interval $(0, \pi)$ the function $g(q) \equiv \sin q \cot(N-2)q$ on the r.h.s. of (21) has $N-2$ branches defined in the $N-2$ intervals $(0, \pi/(N-2))$ ('0-branch'), $(\pi/(N-2), 2\pi/(N-2)), \ldots, (n\pi/(N-2), (n+1)\pi/(N-2)), \ldots, ((N-3)\pi/(N-2), \pi)$ ('π-branch'), respectively. In the intervals $(n\pi/(N-2), (n+1)\pi/(N-2)), n \neq 0, N-3$, it varies from ∞ at $n\pi/(N-2) + 0^+$ (here 0^+ denotes a positive infinitesimal) to $-\infty$ at $(n+1)\pi/(N-2) - 0^+$, while the '0-branch' decreases from a maximum value of $(N-2)^{-1}$ at $q = 0$ ($(\mathrm{d}g/\mathrm{d}q)|_0 = 0$, $(\mathrm{d}^2g/\mathrm{d}q^2)|_0 < 0$) towards $-\infty$ at $q = \pi/(N-2) - 0^+$ and the 'π-branch' increases from a minimum ($(\mathrm{d}^2g/\mathrm{d}q^2)|_\pi > 0$) of value $-(N-2)^{-1}$ at $q = \pi$ to ∞ at $q = (N-3)\pi/(N-2) + 0^+$.

On the other hand, in the interval $(0, \pi)$, the function

$$f(q) \equiv \{[(2\cos q + t)^2 + 1]\cos q - 2(2\cos q + t)\}[1 - (2\cos q + t)^2]^{-1}$$

on the l.h.s. of (21) has two real poles given by $2\cos q \equiv 2\cos q_\pm = -t \pm 1$ for $-1 < t < 1$, a single real pole (q_-) for $-3 < t < -1$ and a single real pole (q_+) for $1 < t < 3$, and, finally, no poles for $t < -3$ and for $t > 3$. Thus, in the case where two real poles exist, $f(q)$ has three branches. The branch to the left of the lowest pole increases from the value

$$f(0) = -\frac{t+1}{t+3} \tag{22}$$

at $q = 0$ to ∞ at the pole; the branch to the right of the upper pole increases from $-\infty$ to the value

$$f(\pi) = \frac{t-1}{t-3} \tag{23}$$

at $q = \pi$; the branch between the two poles increases from $-\infty$ at the lower pole to ∞ at the upper one. The cases where $f(q)$ has only one pole differ from the previous case in that the middle branch is absent.

The solutions of (21) for real q correspond to the intersection points of the functions $f(q)$ and $g(q)$. Clearly, the function $f(q)$ intersects each of the $N-4$ interior branches (i.e. those other than the 0- and π-branches) of $g(q)$ at least once, except for accidental values of ε_0/γ for which a pole of $f(q)$ coincides with a pole of $g(q)$. For the sake of brevity, we shall

omit discussing the case of such special values of (ε_0/γ) in the following. Consider now the conditions under which the 0- and π-branches of $g(q)$ are intersected by a branch of $f(q)$ situated to the left of the lower pole (or of the single pole in the case of a one-pole $f(q)$) and to the right of the upper pole, respectively, or by the continuous function $f(q)$ in the zero-pole case. From the properties of $g(q)$, it follows that for $f(q)$ to intersect the 0-branch of $g(q)$ it is sufficient that

$$f(0) < g(0) = \frac{1}{N-2}$$

since $f(q)$ varies much more slowly than $g(q)$ over the interval $(0, \pi/(N-2))$ for $N \gg 1$. It follows that a real root for q exists in the interval $0 < q < \pi/(N-2)$ for all positive $t \equiv \varepsilon_0/\gamma$ (including $\varepsilon_0 = 0$) and for negative values such that

$$\frac{\varepsilon_0}{\gamma} > -\frac{N+1}{N-1} \qquad \text{or} \qquad \frac{\varepsilon_0}{\gamma} < -3.$$

Similarly we find that existence of a real root of (21) in the interval $((N-3)\pi/(N-2), \pi)$ requires

$$f(\pi) > g(\pi) = -\frac{1}{N-2}$$

which, by using (23), is seen to be satisfied for all negative ε_0/γ and for positive values such that

$$\frac{\varepsilon_0}{\gamma} < \frac{N+1}{N-1} \qquad \text{or} \qquad \frac{\varepsilon_0}{\gamma} > 3.$$

Using the above properties of the functions $f(q)$ and $g(q)$ in (21), we are now able to discuss the solutions of this equation in various domains of the parameter $-\infty < \varepsilon_0/\gamma < \infty$.

First we find that in the domain

$$-1 < \frac{\varepsilon_0}{\gamma} < 1 \tag{24}$$

and, in particular, for $\varepsilon_0 = 0$, equation (21) has N real-q roots. Indeed, since $f(q)$ has two poles in the domain (24) it intersects each of the $N-2$ branches of $g(q)$ once except the two particular branches whose domains, $(n\pi/(N-2), (n+1)\pi/(N-2))$, are crossing a pole of $f(q)$, which are seen to be intersected twice by a branch of $f(q)$. The N real roots in the domain (24) correspond to Bloch-like extended (bulk) states which are only weakly or moderately perturbed by the effect of the difference of Coulomb integrals for surface and bulk atoms.

Consider now the case where $f(q)$ has a single pole. This case arises firstly for negative $t = -1 - 0^+$ for which $-t + 1 = 2 + 0^+$, so the pole $q_+ \to 0$ ceases to exist. With the disappearance of the q_+-pole, the branch of $f(q)$ to the left of q_+ also disappears and hence so does the root corresponding to its intersection with the 0-branch of $g(q)$. Clearly, the disappearance of this real root of (21) near $q = 0$ is associated with the occurrence of a corresponding imaginary root $q = \mathrm{i}\kappa$ (since $\cos q \to \cosh\kappa \geqslant 1$) which will be discussed further, below. Secondly, when t approaches unity the q_--pole of $f(q)$ near $q = \pi$ disappears and with it the real root of (21) associated with the intersection of the branch of $f(q)$ to the right of q_- with the π-branch of $g(q)$. Again, for $t = 1 + 0^+$ this real root near $q = \pi$ has been replaced by a complex root of the form $q = \pi + \mathrm{i}\kappa$ ($\cos q \to -\cosh\kappa \geqslant -1$).

Within the domain $-3 < \varepsilon_0/\gamma < -1$ and $1 < \varepsilon_0/\gamma < 3$ where $f(q)$ has only one pole, we now distinguish the two subdomains:

$$-\frac{N+1}{N-1} < \frac{\varepsilon_0}{\gamma} < -1 \qquad \text{and} \qquad 1 < \frac{\varepsilon_0}{\gamma} < \frac{N+1}{N-1} \tag{25}$$

and

$$-3 < \frac{\varepsilon_0}{\gamma} < -\frac{N+1}{N-1} \qquad \text{and} \qquad \frac{N+1}{N-1} < \frac{\varepsilon_0}{\gamma} < 3 \tag{26}$$

respectively.

In the domain (25), all $N-2$ branches of $g(q)$ are intersected once by a branch of $f(q)$ except the particular branch whose domain crosses the pole of $f(q)$ which is intersected twice. This yields $N-1$ real-q roots which together with the complex root which has appeared at the thresholds $\varepsilon_0/\gamma = \pm 1$ correspond to $N-1$ bulk eigenstates and a single surface eigenstate for the chain in the domain (25).

On the other hand, in the domain (26), two real roots of (21) which existed in the domain (24) have disappeared, namely one root (near $q = 0$ for $\varepsilon_0/\gamma < 0$ and near $q = \pi$ for $\varepsilon_0/\gamma > 0$) corresponding to the intersection with a branch of $g(q)$ whose domain was crossing the pole of $f(q)$ having disappeared, and another real root (near $q = 0$ or near $q = \pi$) due to the intersection of $f(q)$ with the 0- or with the π-branch of $g(q)$ which existed for values $\varepsilon_0/\gamma > -(N+1)/(N-1)$ and $\varepsilon_0/\gamma < (N+1)/(N-1)$, respectively. The disappearance of a $q \rightharpoonup 0$ real root for $\varepsilon_0/\gamma = -(N+1)/(N-1)$ and of a $q \rightarrow \pi$ root for $\varepsilon_0/\gamma = (N+1)/(N-1)$ is associated with the appearance of complex roots $q = \mathrm{i}\kappa$ and $q = \pi + \mathrm{i}\kappa$ at $\varepsilon_0/\gamma = -(N+1)/(N-1) - 0^+$ and at $\varepsilon_0/\gamma = (N+1)/(N-1) + 0^+$, respectively. The N roots of (21) in the domain (26) thus include $N-2$ bulk-state roots and two surface-state roots of distinct origin both emanating from $q \rightarrow 0$ roots or both emanating from $q \rightharpoonup \pi$ roots of (21).

Finally, in the domain

$$\frac{\varepsilon_0}{\gamma} < -3 \qquad \text{or} \qquad \frac{\varepsilon_0}{\gamma} > 3 \tag{27}$$

the two real poles of $f(q)$ which existed in the domain (24) have disappeared. Indeed, the poles q_- and q_+ which also existed in the domain (26) for $\varepsilon_0/\gamma < 0$ and $\varepsilon_0/\gamma > 0$, respectively, disappeared when crossing the boundaries $\varepsilon_0/\gamma = \pm 3$. In this case, the N roots of (21) are composed of the real roots arising from the intersection of the $N-2$ individual branches of $g(q)$ with the continuous function $f(q)$ and of two complex surface-state roots in each one of the regions (27), which replace the two additional real-q roots associated with the poles of $f(q)$ in the domain (24), as discussed above.

We now turn to the properties of the surface eigenstates corresponding to the complex-q roots obtained in the domains (25)–(27) where $|\varepsilon_0/\gamma| > 1$. We have shown above that the two surface states for the parameter ranges (26) and (27) correspond to

$$q = \mathrm{i}\kappa \qquad E = E_0 - \alpha - 2\gamma \cosh\kappa \tag{28}$$

for $\varepsilon_0/\gamma < 0$ and to

$$q = \pi + \mathrm{i}\kappa \qquad E = E_0 - \alpha + 2\gamma \cosh\kappa \tag{29}$$

for $\varepsilon_0/\gamma > 0$. Their energies lie outside the real-q band (18), either beyond the $q = 0$ edge (for $\varepsilon_0/\gamma < 0$) or beyond the $q = \pi$ edge (for $\varepsilon_0/\gamma > 0$). The equations determining the complex wavenumbers, (28) and (29), are obtained by inserting these expressions (with $\varepsilon = -2\gamma \cos q$) in equation (17):

$$\frac{1}{\sinh|\kappa|}\big[(2\cosh|\kappa| \pm t)^2 \sinh(N-1)|\kappa| - 2(2\cosh|\kappa| \pm t)\sinh(N-2)|\kappa| + \sinh(N-3)|\kappa|\big] = 0 \tag{30}$$

where the upper sign corresponds to $q = \mathrm{i}\kappa$ and the lower sign to $q = \pi + \mathrm{i}\kappa$. For analysing the explicit solutions of these equations it is convenient to transform them into the exact forms

$$\left(\frac{\mathrm{e}^{|\kappa|} \pm t}{\mathrm{e}^{-|\kappa|} \pm t}\right)^2 = \mathrm{e}^{-2(N-1)|\kappa|} \tag{31}$$

where the equation with the upper sign corresponds to $q = \mathrm{i}\kappa$ and $\varepsilon_0/\gamma < 0$ and the equation with the lower sign to $q = \pi + \mathrm{i}\kappa$ and $\varepsilon_0/\gamma > 0$. For $(N-1)|\kappa| \gg 1$, equation (31) yields two surface-state solutions for the domains (26) and (27) given by

$$\mathrm{e}^{|\kappa|} = -\frac{\varepsilon_0}{\gamma} \qquad \text{for } \frac{\varepsilon_0}{\gamma} < 0 \tag{32}$$

and

$$\mathrm{e}^{|\kappa|} = \frac{\varepsilon_0}{\gamma} \qquad \text{for } \frac{\varepsilon_0}{\gamma} > 0. \tag{33}$$

The two states have approximately the same energy given, for both signs of ε_0/γ, by

$$E = E_0 - \alpha + \gamma\left(\frac{\varepsilon_0}{\gamma} + \frac{\gamma}{\varepsilon_0}\right) \tag{34}$$

where we have used (28), (32) and (29), (33), respectively.

For completeness of our analytical discussion of surface states, it remains to verify that in the domain (25) where we obtain only one surface-state root the equations (31) do indeed yield one, rather than two, real solutions for each sign of ε_0/γ. First we note that the asymptotic solutions (32) and (33) are not valid in the intervals (25): since in these intervals $|\varepsilon_0/\gamma| = 1 + \mathrm{O}(1/N)$, equations (32) and (33) would correspond to values $|\kappa| = \mathrm{O}(1/N)$ which are incompatible with the asymptotic condition $(N-1)|\kappa| \gg 1$ for these solutions. The existence of a single solution of (31) in the domain (25) may be demonstrated by solving (31) as a second-order algebraic equation for $t = \varepsilon_0/\gamma$. This yields

$$\frac{\varepsilon_0}{\gamma} = \mp\frac{1}{\sinh(N-1)|\kappa|}\left[\sinh N|\kappa| \pm \sqrt{\frac{\cosh 2|\kappa| - 1}{2}}\right] \tag{35}$$

where the upper sign corresponds to values $\varepsilon_0/\gamma < 0$ and the lower sign to values $\varepsilon_0/\gamma > 0$. For $|\kappa| \rightharpoonup 0$ this expression reduces to the values $\varepsilon_0/\gamma = \mp 1$ and $\varepsilon_0/\gamma = \mp(N+1)/(N-1)$, as expected since they are the threshold values at which surface states of the two different origins first appear near $q = 0$ (for $\varepsilon_0/\gamma < 0$) and near $q = \pi$ (for $\varepsilon_0/\gamma > 0$), as discussed above. By expanding (35) to first order in $|\kappa|$, one then easily verifies that for values of ε_0/γ near the boundaries of the intervals (25), one of the two solutions is rejected in all cases because it corresponds to a negative value of $|\kappa|$.

Finally we make explicit the form of the amplitudes at sites n for the surface-state wavefunctions corresponding to the complex wavenumbers $q = \mathrm{i}\kappa$ and $q = \pi + \mathrm{i}\kappa$ given by (32) and (33), respectively. On putting $q = \mathrm{i}\kappa$ in (19) and using (32), the amplitudes for the surface state near the $q = 0$ edge reduce exactly to

$$a_n = \mathrm{e}^{-(n-1)|\kappa|}a_1 \qquad n = 2, 3, \ldots, N \tag{36}$$

i.e. they are exponentially damped away from the site $n = 1$ where the iteration of (1) and (2), (3) was started. Similarly, for $q = \pi + \mathrm{i}\kappa$, with κ defined by (33), equation (19) reduces to

$$a_n = (-1)^{n-1}\mathrm{e}^{-(n-1)|\kappa|}a_1 \qquad n = 2, 3, \ldots, N. \tag{37}$$

Alternative equivalent forms of (36), (37) are obtained by replacing the factor $\exp[-(n-1)|\kappa|]$ by $(-\varepsilon_0/\gamma)^{-(n-1)}$ and $(\varepsilon_0/\gamma)^{-(n-1)}$, respectively.

4. Discussion and concluding remarks

We conclude by comparing our results with those obtained by Goodwin [5] and by commenting further on his approach.

In section 3 we have identified four distinct domains (24)–(27) of (positive and negative) values of the parameter ε_0/γ for which the number and/or the origins of surface eigenstates of the chain are different. In the domain $-1 < \varepsilon_0/\gamma < 1$ the secular equation yields N real eigenstate roots corresponding to Bloch-like bulk states, in agreement with Goodwin [5], provided that one excludes his spurious roots $q = 0$ and $q = \pi$. The energy eigenvalues in this domain are bounded by the band edges $E = E_0 - \alpha \pm 2\gamma$ in (4) and (13), which, however, do not correspond to roots of (21). These bounds of the energy spectrum of a finite chain have been discussed earlier (for $\varepsilon_0 = 0$) [13], using a continued-fraction analysis of Brillouin–Wigner perturbation theory. In the domain (25), there exists a single complex root corresponding to a surface eigenstate, while in the domains (26) and (27), we find two surface-state solutions of different origins. In contrast, by approximating $(N + 1)/(N - 1)$ by unity, Goodwin [5] obtains two surface-state solutions for $\varepsilon_0/\gamma < -1$ (we recall that the Goodwin parameter ε_0/γ corresponds to our $-\varepsilon_0/\gamma$) but does not identify their distinct origins demonstrated in section 3. Also, his analysis is restricted to negative ε_0/γ and requires distinguishing between even and odd numbers of atomic sites N, unlike the discussion above.

Next we compare our amplitudes (36) and (37) of the surface-state wavefunctions with Goodwin's results [5]. While the expressions (36), (37) decay exponentially at sites away from our reference site $n = 1$, Goodwin's surface-state wavefunctions have the form of a linear combination of two states, one which decays exponentially away from site 1 and another one which decays exponentially from site N. This seems rather artificial since for $N \gg 1$ the two ends of the chain are decoupled, so a surface state localized at $n = 1$ should be independent of a state localized at $n = N$.

Goodwin's solution of (1) and (2), (3) consists in writing a_n as a linear combination of two linearly independent plane waves (solutions of (1)):

$$a_n = (-1)^n (A\mathrm{e}^{\mathrm{i}nq} + B\mathrm{e}^{-\mathrm{i}nq}) \qquad n = 1, 2, \ldots, N \tag{38}$$

satisfying the boundary conditions (2), (3). While this procedure yields, of course, the correct secular equation, Goodwin's claim that the amplitudes (38) vanish for the roots $q = 0$ (and $q = \pi$) of this equation is actually incorrect. To show this, we observe that the general solution of the system (1) and (2), (3) is given in terms of an arbitrary value of one of the a_n, say a_1, rather than in terms of an arbitrary value of e.g. B in (38) as implicitly assumed by Goodwin. Therefore B has to be re-expressed in terms of a_1, which, on inserting the ratio $A/B = (\gamma - \varepsilon_0 \mathrm{e}^{-\mathrm{i}q})(\varepsilon_0 \mathrm{e}^{\mathrm{i}q} - \gamma)^{-1}$ obtained from (2) [5] in equation (38) for $n = 1$, yields

$$B = \frac{\mathrm{i}}{2\gamma \sin q}(\varepsilon_0 \mathrm{e}^{\mathrm{i}q} - \gamma)a_1. \tag{39}$$

This expression displays the overall factor $1/\sin q$ arising naturally in our transfer matrix solution (19). As discussed above, this factor causes the amplitudes a_n to be finite for the roots $q = 0$ and $q = \pi$ of the secular equation (17), contrary to Goodwin's claim [5].

Finally, by means of a crude estimate of his parameter ε_0/γ for a copper lattice, using Slater orbitals, Goodwin obtained a negative value slightly exceeding -1. This appears to be one of the reasons why he restricted consideration to negative ε_0/γ in his study of the roots of the secular equation for the eigenvalues of the chain. However, with sophisticated modern computing techniques it should be possible to obtain more accurate values for the Coulomb and overlap integrals. Also, the existence of systems where ε_0/γ would be positive cannot be ruled out *a priori*.

References

[1] Tamm I 1932 *Phys. Z. Sowjet.* **1** 733
[2] Shockley W 1939 *Phys. Rev.* **56** 317
[3] Maue A W 1935 *Z. Phys.* **94** 717
[4] Goodwin E T 1939 *Proc. Camb. Phil. Soc.* **35** 205
[5] Goodwin E T 1939 *Proc. Camb. Phil. Soc.* **35** 221
[6] Goodwin E T 1939 *Proc. Camb. Phil. Soc.* **35** 232
[7] Artmann K 1952 *Z. Phys.* **131** 244
[8] Baldock G R 1952 *Proc. Camb. Phil. Soc.* **48** 457
[9] Davison S G and Levine J P 1970 *Solid State Physics* vol 25 (New York: Academic) p 1
[10] Koutecky J 1965 *Adv. Chem. Phys.* **9** 65
[11] Bortolani V, Celli V and Marvin A 1978 *Nuovo Cimento* B **46** 121
[12] Pendry J B 1994 *Adv. Phys.* **43** 461
[13] Heinrichs J and Jones R O 1972 *J. Phys. C: Solid State Phys.* **5** 2149