On the General Theory of Surface States and Scattering of Electrons in Solids

1963 Proc. Phys. Soc. 81 300

(http://iopscience.iop.org/0370-1328/81/2/311)

PROC. PHYS. SOC., 1963, VOL. 81

On the General Theory of Surface States and Scattering of Electrons in Solids

BY V. HEINE

Cavendish Laboratory, Cambridge

MS. received 31*st August* 1962

Abstract. In discussing the existence of surface states of electrons in solids or the scattering at a plane of discontinuity, it is necessary to match the wave function at the boundary plane, taken as the plane $x = 0$, using an infinite set of solutions of the Schrödinger equation in the solid. The solutions must all have the same energy E and the same k_y and k_z, but will have different values of k_x, which may be complex. The formal structure of the energy is therefore considered, when viewed as a complex function E of the complex variable k_x, and in particular the form of the lines along which E becomes real. This leads to some discussion about setting up approximate wave functions for surface states, differences between one, two and three dimensions, the energy of surface states relative to that of the Bloch bands, and the correction of a conclusion drawn by Goodwin in 1939.

§ 1. INTRODUCTION

IT is sometimes of interest to discuss the electron wave function in a solid near a surface of discontinuity, which may be a stacking fault, twin boundary or other crystal interface, or a free surface of the crystal. The problem of setting up the complete wave function Ψ can be viewed as a question of matching the allowed solutions on the two sides of the boundary surface, which we shall take as the yz plane. On one or both sides of the surface we have a perfect periodic crystal, or if the cells near the physical boundary are distorted we can choose the matching plane slightly inside the crystal. Then for the matching, we need to know all the solutions ψ_j of the Schrödinger equation in the crystal with a given energy E (Maue 1935, Howie 1960, Pippard 1962). Consider for instance the scattering of an electron at a twin boundary. We have the incident Bloch wave, a reflected and transmitted wave, and perhaps one or more diffracted waves; in total a small finite number of waves. These Bloch waves are not sufficient to make up a matched total wave function. Ψ has to be matched at every point of the boundary plane, or boundary line in two dimensions, and this cannot possibly be achieved by considering only a small number of Bloch waves. To match at every point of the boundary line or plane, we require Ψ to be expressed as an infinite series

$$\Psi = \sum a_j \psi_j \tag{1}$$

in terms of a complete set of solutions ψ_j on each side, all with the same energy E and the same k_y and k_z, the components of the reduced wave vector parallel to the boundary. As well as including a finite number of possible Bloch waves in (1), we need an infinite set of further solutions, and these are the Tamm evanescent or surface waves (Tamm 1932) that decay exponentially in the x direction normal to the boundary.

As is well known in one dimension (see, for example, Mott and Jones 1936, p. 58), the evanescent waves have the same form as Bloch waves

$$\psi = \exp(i\mathbf{k} \cdot \mathbf{r}) U_{\mathbf{k}}(\mathbf{r}) \tag{2}$$

where $U_{\mathbf{k}}(\mathbf{r})$ is periodic, but where $\mathbf{k}$ is now complex. Although we are only interested in real energies, a better picture is obtained by considering the band structure $E(\mathbf{k})$ as a complex function E of the complex vector $\mathbf{k}$. For the matching problem on a boundary plane, we then want to study the one-dimensional section of the complex $E(\mathbf{k})$ defined by keeping k_y and k_z real and constant, with k_x being a single complex variable. This one-dimensional section we shall denote by $E(k_x)$. Kohn (1959) has given an exhaustive treatment of the complex band structure for a one-dimensional periodic potential: his methods are based on the theory of ordinary (i.e. not partial) differential equations and cannot be extended to two or three dimensions, not even to a one-dimensional section $E(k_x)$ of a two- or three-dimensional band structure. Blount (1962) has given some general results about $E(\mathbf{k})$ with k_x, k_y, k_z all being considered complex. The main purpose of the present work is to give a more detailed discussion of the complex section $E(k_x)$, showing how it differs from a truly one-dimensional band structure, and showing where all the functions ψ_j in (1) come from. This is done for an arbitrary periodic potential without approximation, and therefore carries forward one step the general discussion which was laid down by Maue (1935) from equation (1). It is hoped that such a clarification of the systematic general theory will serve as a basis for future accurate calculations of surface states. It leads directly to some comments on various matters, including the differences between one, two and three dimensions (Kohn 1959), the setting up of suitable approximations for surface states (Antončík 1961) and the form of the energy spectrum of surface states (Grimley 1962, private communication, Holland 1963). We also find a result of Goodwin (1939), ostensibly derived from the nearly free electron approximation, to be incorrect.

§ 2. THE STRUCTURE OF $E(k_x)$ NEAR THE REAL AXIS

For purposes of finding the complete set of solutions ψ_j in (1), our interest in the one-dimensional section $E(k_x)$ of the complex band structure lies in determining for what ranges of k_x the energy is real and what the energy is along these ranges. We first review from this point of view the form of a truly one-dimensional band structure (figure 1), drawing on the results of Kohn (1959). The whole of $E(k)$ is periodic with

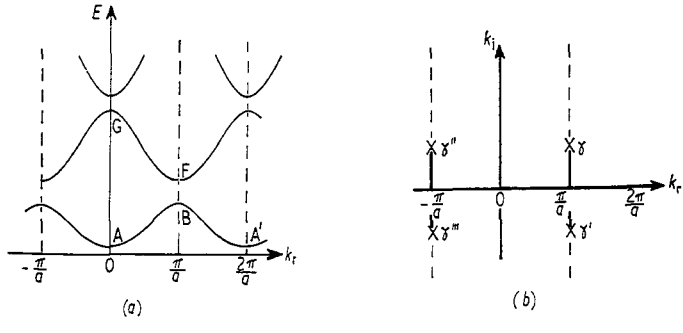

Figure 1. Band structure in one dimension: (*a*) energy for real k; (*b*) lines of real energy in the complex k plane.

period $2\pi/a$ in the direction of the real k axis, and has the following symmetries:

$$E(k) = E\left(k + \frac{2n\pi}{a}\right) = E(-k) = [E(k^*)]^*. \tag{3}$$

The real and imaginary parts of k we denote as follows:

$$k = k_r + ik_i. \tag{4}$$

E is of course real for k on the real axis, for example along the section AB of the band structure (figure 1(a)), while k goes from 0 to π/a along the real axis in the complex plane (figure 1(b)). Then the energy continues to be real when k goes off the real axis along the line $k_r = \pi/a$ (and corresponding lines $k_r = (2n+1)\pi/a$), E increasing monotonically until γ and γ' are reached. These are branch points of order 1 at which two consecutive sheets of $E(k)$ are degenerate, and around which the energy behaves like

$$E(k) \simeq E_\gamma + \text{const.}\,(k-k_\gamma)^{1/2}. \tag{5}$$

The energy is complex beyond γ and γ', but is real and increases monotonically again as k goes from γ or γ' back to π/a. The energy has then reached F (figure 1(a)), so that the two bands AB, FG are joined continuously in the complex plane across the band gap BF. All consecutive pairs of bands are joined in this way on the lines

$$k_r = 2n\pi/a \quad \text{or} \quad k_r = (2n+1)\pi/a.$$

We note that for any energy between A and B, there are two linearly independent Bloch waves ψ_1 and ψ_2 with real k and $k_1 = -k_2$; for E between B and F, there are two linearly independent evanescent waves with k_1, $k_2 = \pi/a \pm i\kappa$ lying in the range $\gamma\gamma'$. For energies below A (figure 1(a)), there are no further bands to join on to, and the energy is real and decreases monotonically to $-\infty$ as $k_i \to \pm\infty$ along the line $k_r = 0$. The periodic repetition of the band structure adds only trivially different solutions. All this follows immediately from the work of Kohn (1959). For any energy, therefore, we have two independent solutions as we expect from the theory of ordinary differential equations, and these two solutions are sufficient for any matching problem at some boundary point in the one-dimensional case.

We now turn to the two- and three-dimensional case. Following Blount (1962), we consider the complex band structure $E(\mathbf{k})$ and the wave function (2) to be defined by

$$H(\mathbf{k})U_{\mathbf{k}}(\mathbf{r}) = E(\mathbf{k})U_{\mathbf{k}}(\mathbf{r}) \tag{6}$$

where

$$H(\mathbf{k}) = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) - \frac{i\hbar^2}{m}\mathbf{k}\cdot\nabla + \frac{\hbar^2k^2}{2m} \tag{7}$$

and where $U_{\mathbf{k}}(\mathbf{r})$ satisfies periodic boundary conditions in the unit cell. It follows from (6) and (7) that $E(\mathbf{k})$ is periodic in reciprocal space

$$E(\mathbf{k}) = E(\mathbf{k}+\mathbf{g}_{lmn}) \tag{8}$$

where $\mathbf{g}_{lmn}$ is a reciprocal lattice vector and is purely real. The symmetry

$$E(\mathbf{k}^*) = [E(\mathbf{k})]^* \tag{9}$$

also follows from (6) and (7), or in the case of our one-dimensional section,

$$E(k_x^*) = [E(k_x)]^* \tag{10}$$

from which it follows that E must be real for real $\mathbf{k}$. $E(k_x)$ is everywhere an analytic

function of the complex variable k_x, except at branch points of the type (5) joining consecutive sheets (Blount 1962). It can be easily shown in the manner of Herring (1937) that the occurrence of these branch points is stable under small perturbation, but that for general values of k_y, k_z, to which we shall restrict ourselves in all that follows, the occurrence of higher order degeneracies is vanishingly probable. There are no other singularities in the finite plane (Blount 1962).

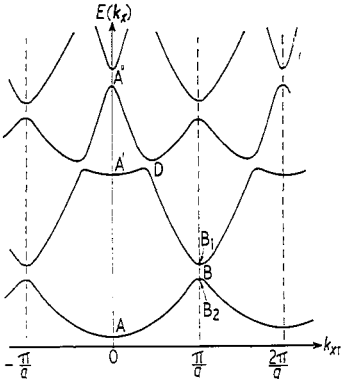

Figure 2. A section $E(k_x)$ for real k_x of a two-dimensional band structure in the nearly free electron approximation. The value of k_y is about $0{\cdot}4\pi/a$.

We consider next $E(k_x)$ in the region near a maximum or minimum in the band structure along the real axis, for example, near A, B or D in figure 2. At an extremum E_0 at k_0, we have $dE/dk_x = 0$, and E in the neighbourhood behaves like

$$E(k_x) \simeq E_0 + (\text{real constant})(k_x - k_0)^2. \tag{11}$$

This expression is real when $k_x - k_0$ is either real or pure imaginary. Thus we expect $E(k_x)$ to be real on a line which crossed the real axis at k_0 at right angles, as well as on the real axis itself. From (11) we also expect the energy along this other line to be a

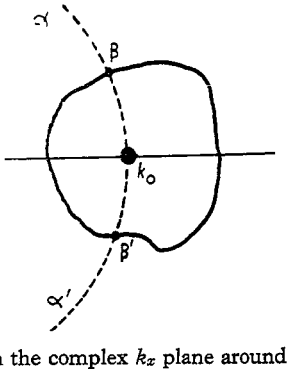

Figure 3. A contour in the complex k_x plane around a saddle point at k_0.

minimum at k_0 if it is a maximum along the real axis and conversely. In either case the energy when viewed in the complex plane has a real saddle point at k_0, and we shall sometimes refer to them as such rather than as extrema. We can prove this more rigorously as follows. Consider a small contour c (figure 3) surrounding k_0 and not

far enough away to include any branch points, and put $f(k_x) = E(k_x) - E_0$. By a well-known theorem

$$\int_c \frac{d}{dk_x} \ln f(k_x)\, dk_x = 2\pi i(Z-P) \tag{12}$$

where Z and P are the number of zeros and poles enclosed, counted with their multiplicity, i.e. $Z = 2$, $P = 0$. Thus (12) becomes

$$[\ln |f|]_c + i[\arg f]_c = 4\pi i \tag{13}$$

where $[\ ]_c$ denotes the change in a quantity in going round the contour. The first term on the left is zero because of single valuedness, and we conclude that $\arg f$ increases by 4π on going around the contour. Hence f is real four times on the contour, namely where $\arg f = 0,\ \pi,\ 2\pi,\ 3\pi$. If E_0 is a maximum along the real axis, the constant in (11) is negative, and $\arg f = \pi$ and 3π (i.e. $f < 0$) on crossing the real axis. The points $\arg f = 0$ and 2π ($f > 0$) must alternate with these, i.e. must lie one each on the upper and lower halves of the contour (β, β' in figure 3). This is true whatever the contour c, and we conclude from continuity that there is a line $\alpha\beta k_0\beta'\alpha'$ cutting the real axis at k_0, on which the energy $E(k_x)$ is real and has a minimum at k_0. From (10) the line is symmetrical about the real axis, and cuts it at right angles.

What happens to these lines, which we shall call 'real lines', as we follow them into the complex k_x plane? They cannot simply terminate. For suppose one did at k_1; then one can apply the argument of (12) and (13) to $f(k_x) = E(k_x) - E_1$, and conclude that the real line must cut the contour twice, contrary to hypothesis.—Similarly, real lines cannot branch, nor conversely can two or more coalesce into one. In principle, they can cross (the two lines being taken to refer to the same sheet of E), but this is vanishingly probable in the sense of Herring (1937) except where one line is the real axis itself. For suppose they do cross at some general complex point k_2. One can apply the argument of (12) and (13) backwards and hence show that $E(k_x) - E_2$ has a double zero at k_2, i.e. $dE/dk_x = 0$ there. Now it can easily be shown by the method of Herring (1937) that the slightest perturbation destroys $E(k_x)$ being real and dE/dk_x being zero simultaneously at the same point.—Similarly a real line can only cross the real axis where E has a maximum or minimum when plotted along the real axis.—As a further corollary we have that the energy varies monotonically along a real line except where it cuts the real axis, because if it did not we should have a point on the line where $dE/dk_x = 0$. At a saddle point on the real axis, the energy increases monotonically as one comes into the saddle point from either direction along one of the real lines (the real axis being counted as a real line) and continues to increase monotonically as one goes out in either direction along the other real line.—A real line can loop back to the real axis (figure 4), but only if it encloses at least one branch point γ. For suppose it did not enclose a branch point; then the contour $\alpha\beta\alpha'$ can be shrunk to a point without passing through a singularity, so that the whole contour lies on a single sheet of $E(k_x)$. Since $E(k_x)$ is analytic everywhere inside, its imaginary part ϕ, say, satisfies $\nabla^2\phi = 0$ everywhere inside, and is zero everywhere on the boundary. Hence from the theorem on the potential inside a hollow conductor, we conclude $\phi = 0$ everywhere inside the contour. Since $E(k_x)$ is analytic, this requires that $E(k_x) =$ constant, and is real everywhere inside, which is vanishingly probable.—A real line can pass through a branch point if the energy is real there, but this is vanishingly probable in the sense of Herring (1937) if the branch point is situated at some general point in the k_x plane. If, however, the crystal has a mirror plane parallel to the boundary plane which we have taken to

define the x axis, then

$$E(k_x) = E(-k_x) \tag{14}$$

and from (8) and (10)

$$\begin{aligned} E(k_x) &= E(-k_x+g_{100}) \\ &= [E(-k_x^*+g_{100})]^* \end{aligned} \tag{15}$$

where $\mathbf{g}_{100}$ is a reciprocal lattice vector in the x direction. If now k_x lies on the line

$$k_{xr} = \tfrac{1}{2}g_{100} \tag{16}$$

then $-k_x^*+g_{100} = k_x$, and from (15)

$$E(k_x) = [E(k_x)]^*. \tag{17}$$

The two sides of this equation refer to *some* energy, which may or may not belong to the same sheet. Thus on the line (16), either $E(k_x)$ is real, or the bands come in complex conjugate pairs. It can also be shown that it is no longer vanishingly probable to have a branch point with a real energy on the line (16) because of the extra symmetry there. If one does, then the energy is real and monotonically increases as one goes into the branch point from one side along the symmetry line, and increases monotonically as one goes out again on the same side. It is as if the real loop in figure 4 had been pulled tight around the branch points until it degenerated into a straight line, exactly as in figure 1(*b*). This would happen for instance around B and A″ in the band structure of figure 2.

To sum up, a real line leaves the real axis at a saddle point of the band structure, is always symmetric about the real axis, and turns back to the real axis after going round one or more branch points (figure 4). In this way the maxima of one band are joined

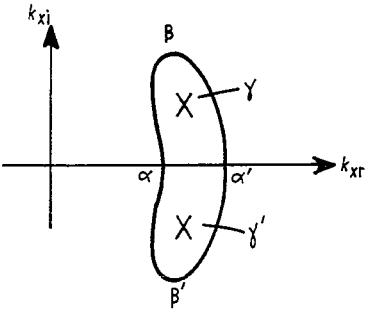

Figure 4. A loop of real line in the complex k_x plane linking real saddle points at α, α', and surrounding branch points at γ, γ'.

across the band gap by a real line to the minima of a higher band. The energy can be followed monotonically along two distinct paths: into the maximum from the left or right along the real axis, along the real line in the upper or the lower half plane to the minimum of the higher band, and out of the minimum to the left or the right into that band. A limiting case of this situation occurs with the saddle points lying on a symmetry line (16), when the real line runs up the symmetry line to a real branch point and back.

§ 3. REAL LINES RUNNING TO INFINITY

The only other possibility, apart from the above loops of real line discussed in § 2, is that a real line runs off to $k_{xi} = \pm\infty$. We can investigate the situation there by

taking k_{xi} large enough ($\hbar^2 k_{xi}^2/2m \gg |V(r)|$) and treating $V(r)$ in (7) as a perturbation. At this stage it becomes best to consider a specific example, which we shall choose as the simple square lattice of side a with a (1, 0) 'face' as the boundary line so that the x and y axes become identical with the crystal axes. The unperturbed free electron energy bands are given by

$$E_{lm}(k_x) = -k_{xi}^2 + (k_{xr} + lb)^2 + 2ik_{xi}(k_{xr} + lb) + (k_y + mb)^2 \tag{18}$$

in Hartree atomic units ($\hbar^2/2m = 1$), where l and m are integers and $b = 2\pi/a$ is the unit reciprocal lattice length. In the fundamental zone strip $-\frac{1}{2}b \leqslant k_{xr} \leqslant \frac{1}{2}b$ of k_x space, the energy (18) is real for given large k_{xi} only when $k_{xr} = 0$ and $l = 0$, but m is allowed to be any integer. We therefore have a single infinity of real lines running along the imaginary axis with energy

$$-k_{xi}^2 + (k_y + mb)^2. \tag{19}$$

We can label these lines by a symbol $(m, \pm)$ according to the energy (19) and whether they come in from large positive or negative k_{xi}. The corresponding wave functions are

$$\psi = \exp[-k_{xi}x + i(k_y + mb)y]. \tag{20}$$

From (17), by continuity, the energy along these lines remains real when the potential $V(r)$ is switched on, as can also be verified by perturbation theory. In the more general case when there is no mirror plane perpendicular to the x axis, the real line is displaced slightly sideways by the perturbation but not destroyed (figure 5). The bands (19) are not degenerate, and the wave functions (20) are also perturbed only to order V/k_{xi}^2.

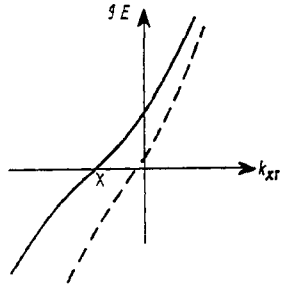

Figure 5. The imaginary part of E plotted along a line k_{xi}, which is large and constant. Broken curve, the unperturbed energy (18); full curve, after perturbation by $V(\mathbf{r})$. X shows the position of the real line.

A similar set of real lines on the axis $k_{xr} = -lb$ is obtained from (18) by taking l non-zero, but these are only a trivial repetition of the set at $k_{xr} = 0$ owing to the periodicity (8) of the band structure, and we will ignore them in the following discussion. The perturbation theory also shows that there are no more real lines on which the energy is large and negative. The real lines come in to join up with the band structure on the real axis at points such as A and A′, but not A″, in figure 2.

For E real and negative, well below the energy of the lowest band on the real axis, we can calculate from (19) approximately the intersection of this energy with all the real lines. For every value of m, we obtain two wave functions (k_{xi} positive and negative) approximating to (20), of which only the one with $k_{xi} > 0$ would be applicable on the side of the boundary $x > 0$. Taking the set with all integers m, we see from (20) that they form a complete set of functions on the boundary line $x = 0$ with reduced wave vector k_y. We have therefore found the complete set of solutions ψ_m to be inserted in

(1). In the case of one dimension, omitting the y term from (19), (20), we see that there are only two real lines going out to $k_1 = \pm\infty$, and only two solutions for any E, which are sufficient for matching at a point, as already remarked. For three dimensions we have $2\times\infty^2$ solutions

$$\psi_{mn} \simeq \exp[-k_{x1}x + i(k_y + mb)y + i(k_z + nb)z] \tag{21}$$

which form a complete set of matching functions in the form of a double Fourier series on the boundary plane $x = 0$, half (k_{x1} positive or negative) being used on one side and half on the other.

The discussion of the preceding paragraph applies only for large negative energies, and we consider now what happens for higher E. Each of the real lines $(m, \pm)$ can be followed with monotonically increasing energy until it joins with the band structure at the saddle points A, A′, etc., on the real axis (figure 2), then either to the left or the right along the real axis, and so on. We can follow the lines continuously through any saddle point by making some convention about which incoming line becomes which outgoing one, and can choose the paths in such a way that no piece of line is covered twice. For instance, one could choose the path of $(m = 0, +)$ in an arbitrary way, then the path $(m = 0, -)$ in such a way as not to overlap with the first one apart from touching at a saddle, etc. In this way we define a set of continuous paths $(m, \pm)$ with monotonically increasing energy along the real lines, winding their way through the band structure, taking in the sections on the real axis and the loops in the complex plane. Moreover, in this construction all sections of real line are automatically included. For suppose to the contrary that some piece of real line is omitted. Starting on it, we can follow it in the direction of *decreasing* energy. When we come to a saddle, if one of the outgoing (increasing E) lines has been omitted, then so has one of the incoming ones (E decreasing) because there are only four real lines at a saddle. Thus we can continue to follow the real line to lower E, and so on, till we reach a large negative E. The path can never join up with one of the set $(m, \pm)$ already counted because lines cannot coalesce. But at large negative E we must be on one of the lines $(m, \pm)$, so that we have a contradiction and hence all sections of real line are included in the set of paths as defined.

We conclude, therefore, that if we cut the band structure at any real energy E, we cut each of the paths once and obtain a complete set of solutions which are continuously related to the set of functions (20) or (21) along the paths. The solutions corresponding to small m (or small m and n in three dimensions) will come from the band structure on or near the real axis, but the solutions for large m will have $k_{xr} = 0$ and $|k_{xi}|$ large, as can be seen from (19). The latter are therefore evanescent waves with high Fourier components in the boundary line and a very rapid exponential decay in the x direction, and thus will resemble (20) or (21) closely.

The discussion proceeds in exactly the same way in the general case of arbitrary crystal structure and arbitrary boundary plane (which is, however, assumed to have rational indices). The geometry is as follows. We define the two shortest (non-parallel) translation vectors $\mathbf{a}_2$ and $\mathbf{a}_3$ in the boundary plane. It is then possible by Euclid's algorithm to find a third vector $\mathbf{a}_1$ such that the parallelopiped $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ forms one unit cell of the crystal and no more. We define the corresponding reciprocal vectors

$$\mathbf{g}_1 = 2\pi\frac{\mathbf{a}_2 \wedge \mathbf{a}_3}{\mathbf{a}_1 \,.\, \mathbf{a}_2 \wedge \mathbf{a}_3} \text{ etc.} \tag{22}$$

and form the fundamental unit cell ('reduced zone') of $\mathbf{k}$ space, as shown in figure 6. We note that $\mathbf{g}_1$ is perpendicular to the boundary plane, giving the required periodicity

of the band structure parallel to the real k_x axis. It also follows that a section k_y, k_z constant through the cell includes *all* points having a certain component of the reduced wave vector parallel to the boundary, which would not in general be true if the unit cell were chosen in other ways. The higher the indices of the boundary plane, the

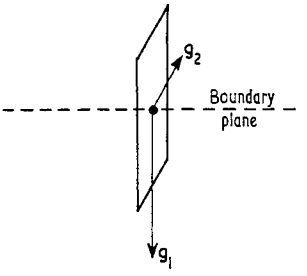

Figure 6. Construction of the reduced zone.

longer and narrower in the k_{xr} direction the reduced zone becomes, and a line k_y, k_z constant tends to intersect a constant energy surface in the Bloch band several times (figure 7). The ψ_j in (1) then include several Bloch waves, which in the case of a surface scattering problem may be described as diffracted waves (Pippard 1962). Incidentally,

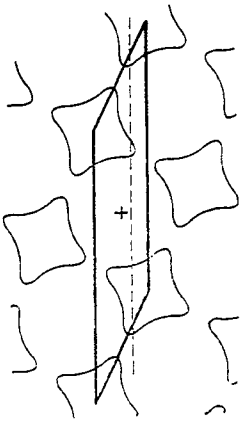

Figure 7. Contours of constant real E in the Bloch band for a two-dimensional square lattice, with the reduced zone relevant to a (1, 2) boundary surface.

in general a crystal would not have a symmetry plane parallel to the boundary, so that none of the real lines would follow the lines (16), and none of the evanescent waves would have $k_{xr} = 0$ or $\pm\frac{1}{2}g_1$ (except where a real line happens to cross (16)) in contradistinction to the one-dimensional case where they all do.

§ 4. APPLICATIONS

We wish to make four comments arising out of the theory above. Firstly, in the context of the second last paragraph of § 3, the waves with low m, n will not resemble (20), (21) closely although they are continuously related to them. These are of course the ones having the largest coefficients in (1) and which are of greatest importance for

the calculation of surface states, etc. Suppose E lies in a band gap. Since ψ changes continuously along a real line, the ψ_j with a complex k_j may be constructed to a first approximation from the $U_{\mathbf{k}}(\mathbf{r})$ (equation (2)) at the nearest saddle point where the loop of real line meets the real axis. A better approximation is obtained from (6) and (7) by treating $H(\mathbf{k}+\delta\mathbf{k})-H(\mathbf{k})$ as a perturbation in the manner of Herring (1937) and Blount (1962). If $\mathbf{k}_j$ lies on the loop of real line linking the band gap, for instance at B (figure 2), the effect of the perturbation is to mix strongly the $U_{\mathbf{k}}(\mathbf{r})$ at B_1 and B_2. This therefore justifies the intuitive assumption of Antončík (1961), and also gives us a perturbation method of estimating the contributions of the $U_{\mathbf{k}}(\mathbf{r})$ from higher bands. A discussion of which states are linked by loops of real line to which other states is therefore important to developing rational approximate wave functions in surface state calculations. Incidentally, when the real line is of the form of figure 1(*b*) associated with a symmetry line (16), the wave function varies continuously along the real line even where it passes through the branch point, as shown by Blount (1962). Moreover the symmetry element $x \to -x$ now links k_x with $-k_x$, which is equivalent to k_x^* and not to k_x, so that $\psi_{\mathbf{k}}$ (or $U_{\mathbf{k}}$) does not have to be symmetric or antisymmetric in x, and indeed changes continuously from one to the other, for example between B_1 and B_2 in figure 2. When k_y and k_z have special values, the symmetry considerations become of course more complicated.

Secondly we note that there is no essential difference in § 2 between the loop of real line associated with a region of the type D (figure 2) in the band structure, and a region of type A″ or B, the latter being just a limiting case of the former. From the theory of § 3, it is equally possible to have surface states associated with either type of band gap. This contradicts the statement by Goodwin (1939), that only the A″ or B type band gaps which are associated with (l, 0, 0) reciprocal lattice planes in the nearly free electron approximation can give rise to surface states.† In fact there is no difficulty in calculating the equation of the real line at D in the nearly free electron approximation.

The third comment concerns the situation in one dimension. Here we have only two real lines coming in along the imaginary axis from $k_i = \pm\infty$ and joining at $k = 0$ on to the band structure on the real axis. The maxima (and minima) of the bands can only occur at $k = 0$ or the zone 'face' $k = \pm\pi/a$, because if one occurred somewhere in between, at k_0 say, then there would also be one at $-k_0$, and for an energy slightly below the maximum (or above a minimum) there would be four linearly independent solutions which is impossible for an ordinary differential equation. (For a much more complicated proof of this result see Shockley (1950).) When the real line leaves the real axis at such a maximum, it has to remain on the symmetry line $k_r = 0$ or $\pm\pi/a$. This follows either from equations (14) to (17) if $V(x)$ has a centre of symmetry, or otherwise by noting that complex conjugation (time reversal) applied to (6) and (7) gives

$$E(k_r+ik_i) = [E(-k_r+ik_i)]^* = [E(-k^*)]^* \tag{23}$$

whence (15), (16), (17) again follow. Thus all that the real line can do is to reach a branch point on the symmetry line and return to a minimum on the real axis. The bands therefore come alternatively concave and convex upwards in the reduced zone $-\pi/a \leqslant k \leqslant \pi/a$ (figure 1(*a*)). We have now rederived all the results of Kohn (1959),

† Since it took the author quite some time to discover how the contradiction could arise, it may help others to point out that equations (14*a*, *b*) of Goodwin define two solutions W_1, q_1 and W_2, q_2, according to whether one takes the plus or minus sign, and either solution if real suffices to satisfy the original equations. The statement below equation (14) is the additional condition for $q_1 = q_2$, but this is by no means necessary.

except for the facts that (i) the branch points come monotonically closer to the real axis with increasing E, and (ii) there are no other branch points in $E(k)$ apart from one pair associated with each band gap. These appear to be specifically one-dimensional results without analogy in two or three dimensions and hence are not derivable by the general methods used here.

Fourthly, we discuss a result of Maue (1935) on the form of the energy spectrum and the existence of localized surface states. We use the term 'surface state' to refer to a properly matched wave function Ψ localized around a free surface of a crystal, and not just any evanescent solution of the Schrödinger equation inside the crystal not satisfying any particular boundary conditions. For a state to be properly localized near the surface, all the ψ_j in (1) on the crystal side of the boundary have to be evanescent waves. If the energy lies inside a 'Bloch sub-band', i.e. a range of energy spanned by Bloch waves in the section $E(k_x)$ of the band structure, the series (1) in general includes Bloch waves and the state is not a localized one. We have, therefore, that the energy of a surface state with given k_y and k_z must lie outside the corresponding Bloch sub-bands, as pointed out by Maue (1935). We wish to point out that the result includes and goes beyond what has recently been proved by Grimley (1962, private communication) and Holland (1963†), namely that if the energy of a state lies outside the Bloch sub-bands then the state is a localized one, because the latter result leaves open the possibility of localized states with energy inside the Bloch sub-bands. Incidentally, for high index surface planes, a section k_y, k_z constant samples a large part of the band structure from both near the top and the bottom of a band (figure 7), and the Bloch sub-bands tend to overlap more. Thus in the case of a metal where there is no complete band gap as in an insulator we expect high index surfaces not to give rise to surface states.

ACKNOWLEDGMENTS

I am indebted to Professor A. B. Pippard, Dr. E. Antončík, Dr. M. Tomašek, Dr. T. B. Grimley and Dr. A. Howie for stimulating my interest in this subject and for helpful discussion.

REFERENCES

Antončík, E., 1961, *J. Phys. Chem. Solids*, **21**, 137.
Blount, E. I., 1962, *Solid State Physics*, Vol. 13, Eds F. Seitz and D. Turnbull. (New York: Academic Press), Appendix C.
Goodwin, E. T., 1939, *Proc. Camb. Phil. Soc.*, **35**, 205.
Herring, C., 1937, *Phys. Rev.*, **52**, 365.
Holland, B. W., 1963, *Phil. Mag.*, in the press.
Howie, A., 1960, *Phil. Mag.*, **5**, 251.
Kohn, W., 1959, *Phys. Rev.*, **115**, 809.
Maue, A. W., 1935, *Z. Phys.*, **94**, 717.
Mott, N. F., and Jones, H., 1936, *The Theory of the Properties of Metals and Alloys* (Oxford: Clarendon Press).
Pippard, A. B., 1962, *Low Temperature Physics*, Eds C. DeWitt, B. Dreyfus and P. G. de Gennes (New York: Gordon and Breach).
Shockley, W., 1950, *Electrons and Holes in Semiconductors* (New York: Van Nostrand), p. 406.
Tamm, I., 1932, *Phys. Z. Sowjet*, **1**, 732.

† *Note added in proof.* These authors have also proved now the result of the present paper by their methods (private communication).