Wavefunction scaling in a quasi-periodic potential

D J Thouless and Q Niu

Department of Physics FM-15, University of Washington, Seattle, WA 98195, USA

Abstract. The nature of wavefunctions in a particular system with two incommensurate periods is discussed. It is shown that the wavefunction can be approximated by a product of functions corresponding to the different quasiperiods, and so these approximate periods serve as discrete length scales for the wavefunction. In this way it can be understood how localised or extended wavefunctions, and also two different types of wavefunctions corresponding to a singular continuous spectrum, can occur. Under certain critical conditions self-similar wavefunctions are found.

1. Introduction

Various studies have been made of one-dimensional wave equations in nearly periodic potentials. It was, for example, shown by Harper (1955) and Zil'berman (1956) that such an equation occurs for electrons in a two-dimensional periodic potential perturbed by a weak magnetic field if the number of flux quanta per unit cell is irrational, and the same equation occurs when the magnetic field is strong and the periodic potential is weak (Rauh *et al* 1974). Azbel (1964, 1979) has examined such equations and shown that the integrated density of states has the structure of a devil's staircase, and that the wavefunctions may be either localised or extended. Aubry and André (1980) have studied the equation of the form

$$Vc_{n-1} + 2V' \cos(2\pi n\varphi + \nu)c_n + Vc_{n+1} = Ec_n. \tag{1.1}$$

This has an important duality property, since the equation for its Fourier coefficients has exactly the same form, with V and V' interchanged, and with ν replaced by the Bloch wavenumber K. This allowed them to deduce that for $V' > V$ there is an exponential decay of solutions with a decay length equal to $1/\ln(V'/V)$ lattice spacings. For the generic case of φ irrational all eigenstates seem to be localised with this localisation length.

The self-dual case $V = V'$ is of particular interest, since it gives a sort of critical point at the boundary between localised and extended states. This corresponds to the problem of electrons in a magnetic field and a two-dimensional potential with square symmetry, and it was pointed out by Hofstadter (1976) that the spectrum in this case is concentrated on a non-denumerable set of measure zero. This spectrum cannot be absolutely continuous if the measure is zero, and it seems unlikely that it is a point spectrum, so the spectrum is probably singular continuous. Since it is known that a point spectrum corresponds to localised states (Kunz and Souillard 1980) and extended states are usually associated with an absolutely continuous spectrum, it is

of considerable interest to know what types of wavefunctions are associated with the intermediate case of a singular continuous spectrum.

Another case of a singular continuous spectrum has been demonstrated by Avron and Simon (1982). For $V' > V$ the states are localised (point spectrum) for almost all irrational values of φ, and the states are of course extended Bloch waves, with an absolutely continuous spectrum, for rational φ. In the case that φ is a Liouville number (which is in some sense an irrational number unusually well approximated by rationals) the spectrum is singular continuous.

It is the purpose of this paper to give a detailed description of the wavefunctions for these problems. In § 2 it is shown how successive rational approximants of φ lead to a scaling theory of the problem. In § 3 the scaling theory is applied to the various cases discussed in this introduction. In § 4 some comparison is made with numerical calculations of the wavefunctions.

2. Scaling theory of the wavefunctions

We follow Azbel (1964) in writing the parameter φ in its continued fraction form

$$\varphi = \frac{1}{N_1+} \frac{1}{N_2+} \frac{1}{N_3+} \dots \tag{2.1}$$

We take the N_i as positive or negative integers whose magnitude is 2 or more. The approximate periods of the system are then N_1, $|N_1N_2 + 1|$, $|N_1N_2N_3 + N_1 + N_3|$, etc. For simplicity in the discussion of the scaling theory we consider the case in which all the N_i are fairly large, but the numerical results given in § 4 suggest that this is not an essential restriction.

The scaling theory is developed by considering blocks whose size is equal to, or close to, the approximate periods of the system. Initially we consider blocks of size N_1 (with the occasional block of size $N_1 \pm 1$ for compatibility with the next period), and then we consider larger blocks made up of $|N_2|$ of the smaller blocks, so that there are $|N_1N_2 + 1|$ sites in these larger blocks, and so on. This is essentially the method used by Azbel (1964) and Sokoloff (1981b) in their discussion of the energy gap structure.

As the first step in this procedure we consider blocks of size N_1. We consider a particular energy E or a narrow energy range close to E, and for positive E we make the division between blocks where $n\varphi + \nu/2\pi$ is half an odd integer, while for negative E the division is made where $n\varphi + \nu/2\pi$ is an integer, so that the division between blocks is made as far as possible from the sites close in energy to E. We have to consider different values of V'/V and of E.

First we consider the case in which V'/V is rather large. If $|E|$ is greater than $2(V + V')$ there is of course no state with energy close to E. For $|E|$ less than $2(V + V')$ each block will have one or two sites close in energy to E. Since the Green function will be dominated by the eigenfunctions of each block lying close to E it is possible to eliminate all but the sites closest in energy to E. A convenient way of doing this, described in the appendix, is to use Feenberg (1948) perturbation theory, which results in the addition of a self-energy due to the eliminated sections, for each site which is kept, and the replacement of the coupling between sites by an effective coupling. Provided the eliminated sites differ in energy from E by an amount large compared

with V the self-energy corrections are small and the effective coupling is insensitive to energy; the condition for this is

$$V < V_1' = (2\pi/N_1)(4V'^2 - E^2)^{1/2}. \tag{2.2}$$

The effective coupling between two sites a distance N_1 apart is approximately

$$V_1 \approx V^{N_1} \prod_{m=n+1}^{n+N_1-1} [E - 2V\cos(2\pi n\varphi + \nu)]^{-1} \approx V(V/V')^{N_1-1}. \tag{2.3}$$

If the energy is close to

$$E_r = \pm 2V'\cos \pi r\varphi, \tag{2.4}$$

where r is some integer less than $1/2\varphi$ there will be some blocks that contain two sites with energy close to E_r, separated by r sites. The effective coupling between such sites within the block is of order $V^r V'^{1-r}$, while the coupling between sites in different blocks is at most of order $V^{N_1-r}V'^{1+r-N_1}$. As a result an energy gap whose width is of order $V^r V'^{1-r}$ opens up around each value of E_r.

Now we can consider the next level of the scaling theory, in which the new block size is $|N_1 N_2 + 1|$. The energies of the $|N_2|$ remaining sites within the block are spread more or less uniformly over a range of order V' as given by equation (2.2), while the coupling between them is of order

$$V_1 \approx V^{N_1}V'^{1-N_1}. \tag{2.5}$$

The condition for the scaling at this stage to proceed in the same way as at the previous stage is

$$V_1 < V_2' = (2\pi/|N_1 N_2 + 1|)(4V'^2 - E^2)^{1/2} \approx V_1'/N_2; \tag{2.6}$$

generally this will be true if (2.5) is satisfied unless N_2 is very large. For this new scale a new set of energy gaps will be introduced close to the energies given by equation (2.4) for larger values of r, because close to these energies there can be two sites with the same energies in the same blocks. The procedure is slightly different close to the energy gaps E_r for $r < N_1/2$, since there the existence of two sites close in energy leads to an increase in the density of states, and an increase in V_1, but it is still true that the ratio of V_1 to V_1' is much less than the ratio of V to V' in that region.

The rescaling can be carried out even when the condition (2.2) is not satisfied. Provided $V < V'$ there is a region around the boundary of each block for which $|E - 2V'\cos(2\pi n\varphi + \nu)| > 2V$, and in this region the solution of the difference equation (1.1) increases or decreases exponentially with distance. Within the block there are one or two regions in which $|E - 2V'\cos(2\pi n\varphi + \nu)| > 2V$, and there the solution oscillates. The quantity which is needed for the rescaling procedure is the end-to-end component of the Green function, and this is dominated by the eigenstate of the block lying closest to E. The most important contributions come from blocks which are close to resonance, and a single resonance approximation is adequate for the blocks that are far from resonance, so it is still possible to replace a block by a single site with the energy of this eigenvalue. The effective coupling of a block to its neighbour is given by V times the value of the eigenfunction at the boundary site. The eigenstates are fairly insensitive to the boundary conditions applied in the regions of exponential growth, and it is probably most convenient to define the eigenfunctions by isolating the block from its two neighbours. The eigenfunctions can be calculated within a block by an appropriate modification of the WKB method (Sokoloff 1981a).

The parameters of the rescaled theory can be estimated by considering the case $\varphi = 1/N_1$. In this case the original problem is periodic, and the rescaled problem involves a set of identical blocks. It can easily be shown that the characteristic equation for the original problem involves the phase ν and the Bloch wavenumber $N_1 K$ only in the E-independent combination

$$2 V^{N_1} \cos N_1 K + 2 V'^{N_1} \cos N_1 \nu. \tag{2.7}$$

The ratio of the ν-dependent term to the K-dependent term gives the ratio of V_1' to V_1, so we have

$$V_1'/V_1 = (V'/V)^{N_1}, \tag{2.8}$$

which is just the same as is given by equations (2.2) and (2.3). Equation (2.2) still gives an upper limit for V_1', but it is no longer a good approximation, because the band gaps may now occupy an appreciable proportion of the available range of energies.

Equation (2.8) shows that in the limit $V' = V$ the rescaled problem is similar to the original problem, so this case gives a critical fixed point of the scaling procedure.

3. Results of the scaling theory

3.1. Localised states

Equation (2.8) shows that if V'/V is greater than unity it scales to larger values. Unless N_i increases very rapidly with i there will be some value j for which

$$V_i < V_{i+1}', \qquad \text{all } i \geq j, \tag{3.1}$$

which is the condition analogous to (2.2) for the eigenstates to be localised on a single block of size $\Pi_{i=1}^{j} N_i$. Beyond this scale states are localised, and fall off exponentially with an exponent $\ln V/V'$. The state may spread over many sites within the block of size $\Pi_{i=1}^{j} N_i$, although the spread will be over a small fraction of the total number of sites in the block.

3.2. Liouville numbers

Avron and Simon (1982), by modifying a similar argument given by Gordon (1976) for the Schrödinger equation, have shown that if φ is a Liouville number then there are no localised eigenstates even for large V'/V. For the Liouville numbers the N_i increase so rapidly that the condition (3.1) is never satisfied. Since V_{i+1}' is of order V_i'/N_{i+1}, this will occur if

$$N_{i+1} \gg (V'/V)^{\Pi_{j=1}^{i} N_j}. \tag{3.2}$$

It is easy to show that eigenstates cannot be localised under these conditions. On each length scale the difference in energy between one block and its neighbour is so small that the very weak coupling between the blocks is sufficiently strong to cause the eigenstates to spread out as a standing wave over a number of blocks n_j, which is a small fraction of the total number of blocks N_j in an approximate period; except near the band edges n_j/N_j is of order V_{j-1}/V_{j-1}'. Thus the total number of sites over

which the main part of the wavefunction is spread is given by the unbounded product $\Pi_i n_i$.

The nature of the local density of states can be examined by a similar argument. We consider in succession systems with periods N_1, $|N_1 N_2 + 1|$, and calculate the local density of states for $n = 0$ at some fixed value of the phase ν. For the system with period N_1 the spectrum consists of N_1 narrow bands. Since V'/V is large there will be n_1 of these bands that contribute an appreciable weight to the density of states at $n = 0$, since these are bands whose wavefunctions oscillate near $n = 0$. The remainder of the bands have weights that decrease by successive factors of order $(V/V')^2$. For the system of period $|N_1 N_2 + 1|$ each band splits up into $|N_2|$ very narrow components. A small proportion $n_2/|N_2|$ of the components will have relatively large weights, and these have energies that lie within the range of the bands at the previous stage. The other components have weights that decrease by factors of order $(V/V')^{2N_1}$, and these have energies that are spread more or less uniformly through the gaps of the previous stage. This process by which energy bands break down into more and more components filling a larger proportion of the total range of energies leads to a singular continuous spectral density. It is not a point spectrum because the weight associated with any particular band is no more than $\Pi_i n_i^{-1}$, and it is not absolutely continuous, because at each stage in this procedure the maximum value of the spectral density, in any interval in which it is non-zero, increases by a factor of order $V_{j-1}/n_j V_j \approx V'_j/V_j$, which is a large number, so the spectral density is unbounded everywhere. The support of this spectrum has a measure $4|V' - V|$ (den Nijs, Nightingale and Thouless, unpublished).

3.3. The critical point

The point $V = V'$ is a critical point, as a lot of previous work has made clear. In particular Aubry and André (1980) showed that the localisation length becomes infinite everywhere in the spectrum as V approaches V' from below. It was also shown by Hofstadter (1976) that the support of the spectrum has zero measure for irrational φ in this limit, and so the spectrum is not absolutely continuous.

Equation (2.8) shows that V'_j/V_j is unity for all j, so the effective coupling constant is the same on all length scales. The form of the wavefunction for each length scale does depend also on the value of N_j and on the value of the energy. For example, the states whose energies are close to a sub-band edge on a certain length scale will have wavefunctions which oscillate only over a very few blocks, and which fall off exponentially over the rest of the blocks in an approximate period, while states whose energies are near the centre of the sub-band oscillate almost everywhere. The wavefunction is a product of factors for each length scale, where the factor depends on N_j and the position of E within the sub-band corresponding to the next shorter length scale. As a result, at each stage in the rescaling the strength corresponding to a particular energy band gets spread over a number of sites whose ratio to the number at the previous stage depends on N_j and the position of the band, but does not depend explicitly on j. The spectral density corresponding to each band therefore tends to zero, and so the spectrum is not a point spectrum. Since it is not absolutely continuous it should be singular continuous.

If all the N_j are identical, as they are for solutions of the equation

$$\pm\varphi^2 + N\varphi = 1,\tag{3.3}$$

and if the energy is in the same position relative to each set of sub-bands, as it is for $E = 0$ or for the highest or lowest state in the spectrum, then the wavefunction is a product of similar factors

$$c_n/c_0 \approx \prod_j f(n/p_j),$$ (3.4)

where

$$p_{j+1} = Np_j \pm p_{j-1}, \qquad p_1 = N, \qquad p_0 = 1,$$ (3.5)

and f is a Bloch function with period unity, with $f(0) = 1$. Hofstadter (1976) has illustrated a wavefunction with this type of structure, but he chose a case where the N_j were different.

An expression like (3.4) is neither exponentially localised nor an extended state of the usual sort. There is a sense in which the wavefunction falls off like a power of distance, since

$$\ln|c_n| \approx \sum_{j=1}^{\infty} \ln|f(n/p_j)|,$$ (3.6)

and f is close to unity for $n \ll p_j$. Therefore, the number of terms contributing significantly to this sum is of order $\ln n / \ln N$, and so we have

$$|c_n|/c_0 \approx |n|^{\ln \bar{f}/\ln N},$$ (3.7)

where $\bar{f}$ is the geometric mean of f. However, this is a very rough approximation, and the function is quasi-periodic, and returns close to its maximum value at some distant points.

3.4. Extended states

For $V > V'$ it is no longer possible to adopt the same scaling procedure, since the effective coupling becomes strong on large length scales, and it is no longer possible to argue that only a few eigenstates contribute, or that boundary conditions are unimportant. We know, however, that the density of states is the same as that for the dual problem with V and V' interchanged (Aubry and André 1980). We can therefore calculate eigenstates for this problem by taking the Fourier transform of eigenstates of the dual problem. From the solution of the dual problem

$$V'd_{m-1} + 2V \cos(2\pi m\varphi + K)d_m + V'd_{m+1} = Ed_m,$$ (3.8)

we can construct solutions of equation (1.1) of the form

$$c_n = (2\pi)^{-1} \sum_m d_m \exp(2\pi i\varphi mn + iKn + i\nu m).$$ (3.9)

Since the d_m are exponentially localised and there exist n for which $\mathrm{frac}(\varphi n)$ is arbitrarily small, this has the character of a running wave.

4. Numerical calculations

As a check on the arguments presented in this paper, a number of eigenfunctions were computed for rational values of φ, both for $V \neq V'$ and for the critical case $V = V'$. In general the expected qualitative features were observed, even when the

numbers N_i in the continued fraction were quite small. In order to do a more quantitative test we made a more detailed study of some cases with $V = V'$ where the self-similarity of the wavefunction was expected to be apparent.

One case that is expected to be self-similar is a wavefunction corresponding to an extremal eigenvalue for a value of φ for which all the N_i in the continued fraction expansion are identical. In figure 1 is shown the highest eigenvalue in the case

$$\varphi = \frac{12}{29} = \frac{1}{2+} \frac{1}{2+} \frac{1}{2+} \frac{1}{2+} \frac{1}{2} = \frac{1}{2+5/12},$$ (4.1)

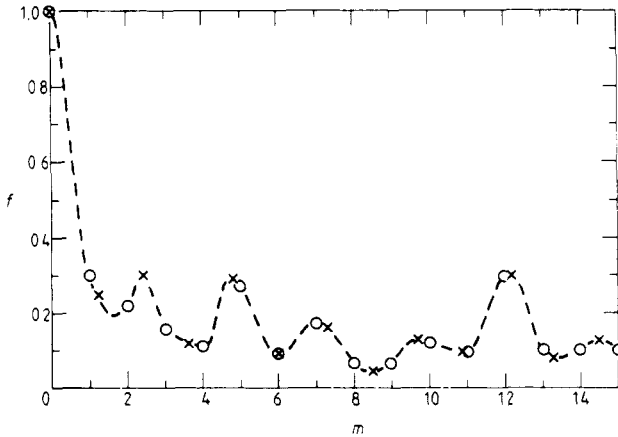which we have denoted by $f_{12/29}(m)$.

Figure 1. Plot of the lowest energy solution for $\varphi = 12/29$ compared with the product of the solutions for $\varphi = 5/12$ and $1/2$, with the scale altered so that the periods are the same. Solutions are normalised so that the maximum value is unity ($\times$, $f_{5/12}(12m/29) \times f_{1/2}(24m/29)$; $\bigcirc$, $f_{12/29}(m)$; $V = V' = 1.0$, $\varepsilon = \varepsilon_{\max}$).

To show that this result is compatible with the product form of the wavefunction given in equation (3.4), despite the small values of the N_i, we have also plotted the product $f_{5/12}(12m/29)f_{1/2}(24m/29)$. The two seem to follow the same somewhat irregular curve, and there are no substantial discrepancies visible.

Other examples of states which are expected to be self-similar can be found at the centre of the band. Here the situation is a little different because the central band generally splits up into a different number of components from the other bands. The sequence of fractions p_n/q_n generated by the recurrence relations

$$p_{n+1} = 2p_n + p_{n-1}, \qquad q_{n+1} = 3p_{n+1} + p_n,$$ (4.2)

with $p_0 = 0$, $p_1 = 1$, each of which has a continued fraction expansion with $N_1 = 3$ and all other N_i equal to 2, has the property that the central band splits into three components at each stage in the iteration. Figure 2 shows the modulus of the wavefunction for zero energy plotted over a half-period for the first five fractions in this sequence. Each wavefunction in this sequence looks similar to its predecessor, but has more structure of the expected sort than its predecessor.

We have made two quantitative tests of the scaling of these wavefunctions. With the amplitude at the site zero (the phase ν is chosen to make the energy of this site
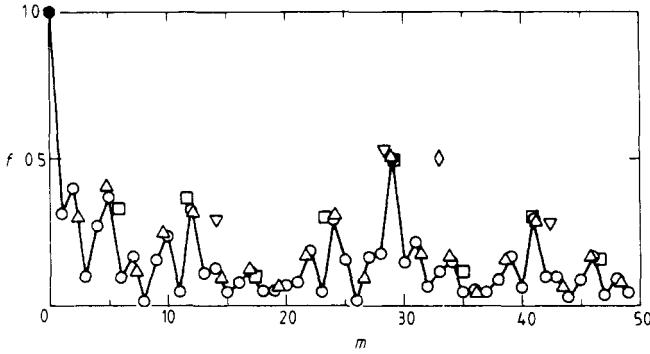
Figure 2. Plot of the modulus of $E = 0$ solutions for $\varphi = 29/99$, 12/41, 5/17, 2/7, and 1, 3, rescaled so that the periods are the same. Solutions are normalised so that the maximum value is unity ($\bigcirc$, $f_{29/99}(m)$; $\triangle$, $f_{12/41}(41m/99)$; $\square$, $f_{5/17}(17m/99)$; $\triangledown$, $f_{2/7}(7m/99)$; $\lozenge$, $f_{1/3}(m/33)$).

zero, so that the wavefunction has its maximum there) fixed as unity, both the normalisation per period and the current (of the complex Bloch wave) were calculated. The ratio of these two for the same value of φ gives the frequency with which the particle tunnels from one unit cell of the periodic system to another, which is proportional to the width of the energy band. Table 1 shows the results of this calculation, together with the ratios between successive members of the sequence. These ratios seem to tend rather rapidly to limits, as they should for this critical case. In the limit the normalisation tends to a value proportional to $p^{0.297}$ and the current to a value proportional to $p^{-1.280}$, so the width of the energy band should be proportional to $p^{-1.577}$, where p is the length of the period. These exponents are not expected to have any general significance.

Table 1. Normalisation N_i per unit cell, current J_i in units of $V/\hbar$, and ratios of these for a state at the centre of the band for the periodic systems given by equation (4.2). The maximum amplitude of the wavefunction is taken to be unity. The extrapolated limit is a guess.

φ_i	N_i	J_i	N_i/N_{i-1}	J_i/J_{i-1}
1/3	1.5000	0.5000	—	—
2/7	1.8622	1.6191×10^{-1}	1.2415	0.3238
5/17	2.4628	5.2958×10^{-2}	1.3225	0.3271
12/41	3.1792	1.7042×10^{-2}	1.2909	0.3218
29/99	4.1397	5.5254×10^{-3}	1.3021	0.3242
70/239	5.3744	1.7865×10^{-3}	1.2983	0.3233
169/577	6.9872	5.7846×10^{-4}	1.3001	0.3238
limit			1.2995	0.3235

Appendix

For the case of a linear chain with an equation of the form (1.1), Feenberg (1948) perturbation theory is well known to give a particularly simple prescription for the elimination of unwanted sites on blocks. This prescription can be derived from first principles in the following way. We take the two equations

$$-b_{i-1}c_{i-1} + a_i(E)c_i - b_i c_{i+1} = 0,$$
$$-b_{j-1}c_{j-1} + a_j(E)c_j - b_j c_{j+1} = 0,$$

(A1)

where the b_i will also depend on E if previous eliminations have been performed, and solve $j - i - 1$ intervening equations to get c_{i+1} and c_{j-1} in terms of c_i and c_j in the form

$$c_{i+1} = D_{i+1,j-1}^{-1}\left(c_i b_i D_{i+2,j-1} + c_j \prod_{i+1}^{j-1} b_k\right),$$
$$c_{j-1} = D_{i+1,j-1}^{-1}\left(c_i \prod_{i}^{j-2} b_k + c_j b_{j-1} D_{i+1,j-2}\right).$$

(A2)

Here D represents a determinant given by the recurrence relations

$$D_{k,j-1} = a_k D_{k+1,j-1} - b_k^2 D_{k+2,j-1} \qquad \text{for } j > k + 1,$$

(A3)

with $D_{j-1,j-1} = a_{j-1}, D_{j,j-1} = 1$. Since the sites chosen for elimination have a larger than b, the determinant can be approximated by

$$D_{i+1,j-1} \approx \prod_{i+1}^{j-1} a_k.$$

(A4)

Substitution of (A2) in (A1) gives the renormalised equations

$$-b_{i-1}c_{i-1} + (a_i(E) - b_i^2 D_{i+2,j-1}/D_{i+1,j-1})c_i - \left(\prod_{i}^{j-1} b_k/D_{i+1,j-1}\right)c_j = 0,$$
$$-\left(\prod_{i}^{j-1} b_k/D_{i+1,j-1}\right)c_i + (a_j(E) - b_{j-1}^2 D_{i+1,j-2}/D_{i+1,j-1})c_j - b_j c_{j+1} = 0.$$

(A5)

References

Aubry S and André G 1980 *Ann. Israel Phys. Soc.* **3** 133–64
Avron J and Simon B 1982 *Bull. Am. Math. Soc.* **6** 81–5
—— 1983 *Commun. Math. Phys.* to be published
Azbel M Ya 1964 *Sov. Phys.–JETP* **19** 634–45
—— 1979 *Phys. Rev. Lett.* **43** 1954–7
Feenberg E 1948 *Phys. Rev.* **74** 206–8
Gordon M Ya 1976 *Usp. Mat. Nauk* **31**, no. 4, 257–8
Harper P G 1955 *Proc. Phys. Soc.* A **68** 874–8
Hofstadter D R 1976 *Phys. Rev.* **14** 2239–49
Kunz H and Souillard B 1980 *Commun. Math. Phys.* **78** 201–46
Rauh A, Wannier G H and Obermair G 1974 *Phys. Status Solidi* b **63** 215–29
Sokoloff J B 1981a *Phys. Rev.* B **23** 2039–41
—— 1981b *Phys. Rev.* B **23** 6422–9
Zil'berman G E 1956 *Sov. Phys.–JETP* **3** 835–9