Hierarchical clustering in the spectra of incommensurate systems

1989 J. Phys. A: Math. Gen. 22 717

(http://iopscience.iop.org/0305-4470/22/6/022)

J. Phys. A: Math. Gen. **22** (1989) 717-729. Printed in the UK

Hierarchical clustering in the spectra of incommensurate systems

S C Bell† and R B Stinchcombe

Department of Theoretical Physics, University of Oxford, 1 Keble Road, Oxford OX1 3NP, UK

Received 24 May 1988

Abstract. Using degenerate perturbation theory to all orders we derive the Hofstadter clustering rules for the spectrum of Harper's equation. We also extend our method to calculate the fractal dimension, D_f, of the spectrum and speculate that $D_f = \frac{1}{2}$ for typical values of the incommensurability parameter ϕ. The clustering rules are also shown to be valid for a one-dimensional model of quasicrystals.

1. Introduction

Incommensurability occurs when there are two or more competing length scales in a problem. A simple example is that of the mismatch of lattice constant between a substrate and an absorbed layer (Ying 1981). The effect of the competition between the forces favouring different atomic separations is often to stabilise modulated states whose wavelength is an irrational ratio of the underlying lattice constant. Another less obvious example, which motivates the main equation to be investigated in this paper, is that of Bloch electrons in a magnetic field. Here the two length scales are the de Broglie wavelength and the cyclotron radius. The competition is between the crystal field, which on its own produces a band spectrum, and the magnetic field which on its own produces a discrete spectrum. As will be seen, the effect of these two forces is to produce a spectrum which is neither discrete nor continuous but, in fact, Cantor-set-like. It is the analysis of this kind of spectrum which is the focus of attention in this paper.

The first objective of this work is to investigate a model equation (Harper 1955) for this problem using perturbative methods. Following Stinchcombe and Bell (1987) we show that a beautiful hierarchical description of the spectrum due to Hofstadter (1976) can be derived analytically. From this basic analytic result further properties of the spectrum can be calculated such as the fractal dimension, D_f. We then show that the derivation of the rules is sufficiently general that it can be applied to similar equations which have occurred in the theory of quasicrystals (e.g. Luck and Petritis 1985). We begin here with a more detailed introduction to Harper's equation.

2. Harper's equation

Harper (1955) showed that using some ideas due to Peierls (1933) the Schrödinger

† Present address: Department of Physics, University of Manitoba, Winnipeg, Manitoba, Canada R3T 2N2.

0305-4470/89/060717+13$02.50

equation for a Bloch electron in a magnetic field could be approximated by

$$c_{n+1} + c_{n-1} + 2\lambda \cos(2\pi\phi n + \nu)c_n = Ec_n \tag{1}$$

where c_n is related to the wavefunction. An isotropic lattice corresponds to the case $\lambda = 1$. ϕ is the basic incommensurability parameter and is equal to the number of flux quanta per unit cell. It has been shown (Bell 1986) that (1) is also the equation for phonons in a modulated atomic chain. In the case of Bloch electrons the physical spectrum is obtained by finding the union of the spectra for all ν. Equation (1) is a linear eigenvalue problem; if $\phi = p/q$, for p, q integer, then it reduces to a $q \times q$ matrix equation. Here we are interested in the incommensurate limit ($q \to \infty$).

One of the fundamental properties of (1) is that it admits a dual transformation

$$c_n = \sum_m d_m \exp(2\pi i\phi mn + i\nu m + ikn) \tag{2}$$

such that

$$\lambda(d_{m+1} + d_{m-1}) + 2\cos(2\pi\phi m + K)d_m = Ed_m. \tag{3}$$

Thus $\lambda = 1$ is a self-dual point, which suggests that it locates some kind of critical point. Because (2) is a special kind of Fourier transformation it interchanges extended states with localised states and vice versa, so it seems likely that there is a localisation transition at $\lambda = 1$. This has been verified by numerical calculation. Aubry and André (1980) have shown that if a Lyapanov exponent (inverse localisation length) is associated with the wavefunction then $L = 1/\xi = \ln(\lambda)$ for $\lambda > 1$. ξ diverges at the transition in a similar way to the thermal correlation length in statistical mechanics. In addition Thouless (1983) has shown that the bandwidth (sum of the widths of the individual sub-bands) for the union of the spectra for all ν is $4|\lambda - 1|$, which vanishes at the self-dual point. There is, then, the possibility of a fractal or Cantor-set-like spectrum at this point. Hofstadter made an extensive numerical investigation of the spectrum for this case and its highly intricate patterns in the $E - \phi$ plane have inspired people to christen it the 'butterfly' diagram (figure 1).

There are many ways of attempting to analyse Harper's equation. In the original paper he performed a WKB calculation but it did not reveal the full complexity of the problem. However, this work was followed up by Azbel (1964) who extended the WKB analysis and put forward the basic suggestion that the clustering of the bands is related to the continued fraction representation (CFR) of ϕ. In particular he conjectured that if

$$\phi = \cfrac{1}{n_1 + \cfrac{1}{n_2 + \cfrac{1}{n_3 + \cfrac{1}{n_4 + \ldots}}}} \tag{4}$$

(which shall be written as $\phi = [n_1, n_2, n_3, n_4, \ldots]$) then the spectrum could be described as consisting of n_1 bands which split up into n_2 smaller bands and so on. This is not quite correct as the total number of bands produced by this scheme is

$$n_1 n_2 n_3 \ldots \tag{5}$$

while the actual number is

$$n_1 n_2 n_3 \ldots + n_2 n_3 \ldots \tag{6}$$

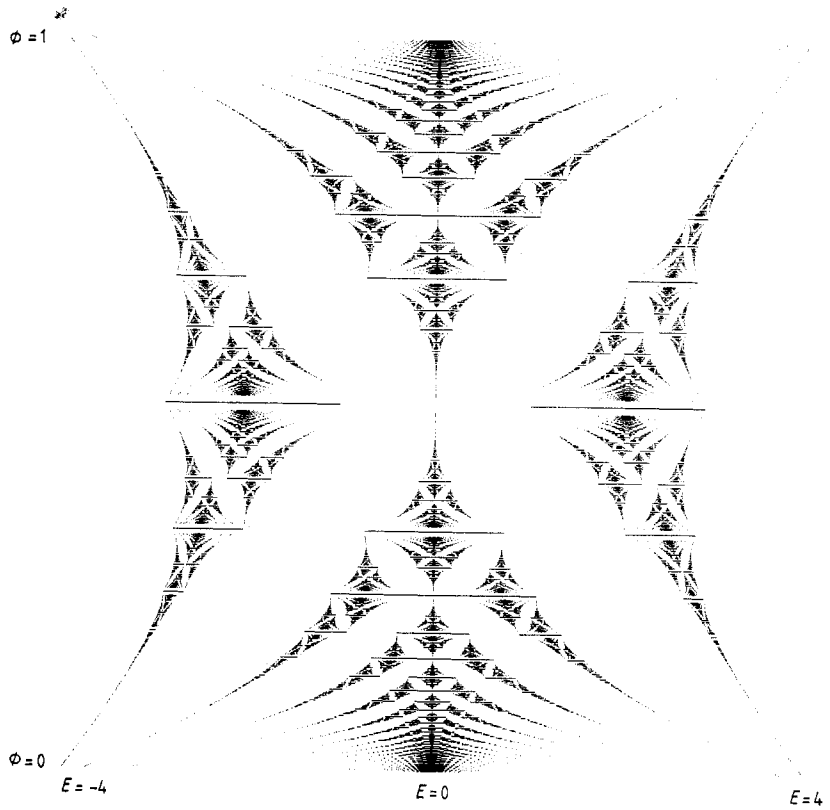

Figure 1. The spectrum of Harper's equation at the critical point. Energy is the horizontal variable ranging between −4 and 4 while ϕ is the vertical variable ranging between zero and one.

i.e. the denominator q of the fraction ϕ. This follows trivially from the fact that q is the dimension of the matrix to be diagonalised. In a numerical investigation of the spectrum Hofstadter discovered a set of rules which relate the band clustering to the CFR of ϕ in a more complicated way than that of Azbel but without its inconsistencies. Unfortunately this work was numerical and the origin of the rules was unexplained. In the next section we show that these rules are directly related to a perturbation theory of the equation.

3. Proof of the Hofstadter clustering rules

Hofstadter's clustering rules are a recursive way of analysing the spectrum of Harper's equation for a given value of ϕ. They were first deduced for the special case $\lambda = 1$ but here they will be derived for all λ. Inspecting figure 1 shows that the spectrum is split into two side clusters and a central cluster separated by two principal gaps. The idea is to consider these clusters as rescaled versions of the full spectrum but with new effective values of ϕ. Hofstadter deduced that for the side clusters this rescaled value of ϕ is

$$\begin{aligned} R(\phi) &= \{1/\phi\} && 0 < \phi < \tfrac{1}{2} \\ &= \{1/(1-\phi)\} && \tfrac{1}{2} < \phi < 1 \end{aligned} \tag{6}$$

while for the central cluster

$$S(\phi) = \{\phi/(1-2\phi)\} \qquad 0 < \phi < \tfrac{1}{2}$$
$$= \{1-\phi/(2\phi-1)\} \qquad \tfrac{1}{2} < \phi < 1 \tag{7}$$

where $\{x\}$ denotes the fractional part of x. These rules can be applied recursively, at each stage locating the principal gap of each cluster. The iteration ends when the rescaled values of ϕ are equal to zero, in which case the sub-clusters consist of only one band.

To derive these rules we shall use perturbation theory in powers of λ. Writing the Hamiltonian as

$$H = \delta_{n,m+1} + \delta_{n,m-1} + 2\lambda \cos(2\pi\phi n + \nu)\delta_{n,m} \tag{8}$$

we shall take the cosine term as the perturbation so that plane wave states are the unperturbed wavefunctions. For simplicity we shall firstly apply Rayleigh-Schrödinger perturbation theory which gives

$$E(k) = E_0(k) + \sum_{k'} \left(\frac{|V_{kk'}|^2}{E_0(k) - E_0(k')} \right) \tag{9}$$

for the second-order shift. $V_{kk'}$ is the matrix element between unperturbed states and is non-zero only for $k' = k \pm 2\pi\phi$. $E_0(k)$ is the unperturbed energy. The denominator vanishes (and hence perturbation theory breaks down) when $E_0(k) = E_0(k \pm 2\pi\phi)$ which happens at $k = \pm\pi\phi$ and at $k = \pm\pi(1-\phi)$. These are the wavevectors at which the first gaps open up. In mth order the perturbation will open up gaps at the wavectors $k = \pm\pi m\phi$ and $k = \pm\pi(1-m\phi)$ which were not present at lower order. Reducing to the zone $(-\pi, \pi)$ this becomes

$$k = \pm\pi\{m\phi\} \qquad \text{or} \qquad \pm\pi(1-\{m\phi\}). \tag{10}$$

The above equation for the location of the gaps which first appear in mth order is the first crucial statement in the proof of the rules. It is believed to be generally true for quasiperiodic Schrödinger operators (Ostlund and Pandit 1984). The other assumption which we shall use is that the gap formed by the mth order of perturbation will be of order λ^m and so m labels the gaps in order of decreasing size (Thouless *et al* 1982). It may seem that for $\lambda > 1$ the perturbation theory should break down but by duality the $\lambda > 1$ equation can be transformed into a $\lambda < 1$ equation (see (1), (2) and (3)) so that our arguments will hold in this case too. There may still appear to be a problem at $\lambda = 1$ but this can be resolved by a simple argument. For small λ the principal gap width is just 2λ while for large λ it must tend towards 2 by using duality. By extrapolation between these two limiting cases it is clear that for $\lambda = 1$ the gap width must be less than 2λ, so there is an additional convergence factor. In general we expect the gap at mth order to be narrower than λ^m and so our perturbation theory should work for all λ.

The problem of vanishing denominators in the Rayleigh-Schrödinger perturbation theory can be eliminated by using degenerate perturbation theory. Taking matrix elements of the Hamiltonian between states $|k\rangle$ and $|k+2\pi\phi\rangle$ and diagonalising the resulting sub-Hamiltonian gives

$$E(k) = 2\cos(k+2\pi\phi)\cos(\pi\phi) \pm [4\cos^2(k+\pi\phi)\sin^2(\pi\phi) + \lambda^2]^{1/2} \tag{11}$$

which is valid near $k = -\pi\phi$ and $-\pi - \pi\phi$, where the states used are degenerate. There is a similar result valid near $k = \pi\phi$ and $\pi - \pi\phi$ which is obtained from the above by

replacing ϕ by $-\phi$. In both cases the gaps are of width 2λ. A sketch of the resulting spectrum is shown in figure 2. This, in effect, proves the first part of the clustering rules, namely that the spectrum is split up by two principal gaps into three sections. It strongly suggests that the origin of the rest of the rules can be explained in terms of perturbation theory. The splitting of the spectrum into three main sections is a feature of other 1D quasiperiodic potentials.

The next stage is to investigate the pattern of splittings of this first-order approximation to the spectrum to all orders of perturbation and hence to prove the clustering rules. Consider the upper side band which occupies $(-\pi\phi, \pi\phi)$ in k space. It is convenient to regard the set of k values given by (10), which locate the position in k space at which gaps first appear in mth, as steps in a walk in k space. There are then four walks all of step length $\pi\phi$, which move off to the left and to the right from the symmetry points at $k=0$ and $k=-\pi$ (which is equivalent to π) visiting the k values corresponding to the degenerate states linked by the perturbation. The mathematical problem involved in discussing the splittings of the sub-clusters is to find which walks (i.e. the m value) land in the region $(-\pi\phi, \pi\phi)$ and to show that this sequence of k values, after rescaling by $1/\phi$ to 'stretch' the region $(-\pi\phi, \pi\phi)$ to $(-\pi, \pi)$, is the same as that for the whole zone (given by the sequence in (10)) except that ϕ is replaced by a rescaled value ϕ'. The walk starting from $-\pi$ will take $m=[1/\phi]$ steps to reach $(-\pi\phi, \pi\phi)$ and the following step will also lie in this region as the step length is half its width. So, $m=[1/\phi]$ and $m=[1/\phi]+1$ are the lowest orders of perturbation theory which split the upper side cluster. Now consider the walk starting from the point $k=0$. It will reach the region $(-\pi\phi, \pi\phi)$ after $m=[2/\phi]$ steps so that $m=[2/\phi], [2/\phi]+1$ are the next orders of perturbation which split the upper side band. Continuing in this way it is clear that, in general

$$m=[n/\phi], [n/\phi]+1 \tag{12}$$

are the orders of perturbation splitting the upper side band where n is an integer $(n>1)$. These gaps will occur at $k=\pi(m\phi-n)$. Taking into account the fact that walks moving to the left produce gaps at k values of opposite sign and substituting into (10) gives

$$k=\pm\pi\phi\{n/\phi\} \qquad \text{or} \qquad \pm\pi\phi(1-\{n/\phi\}) \tag{13}$$

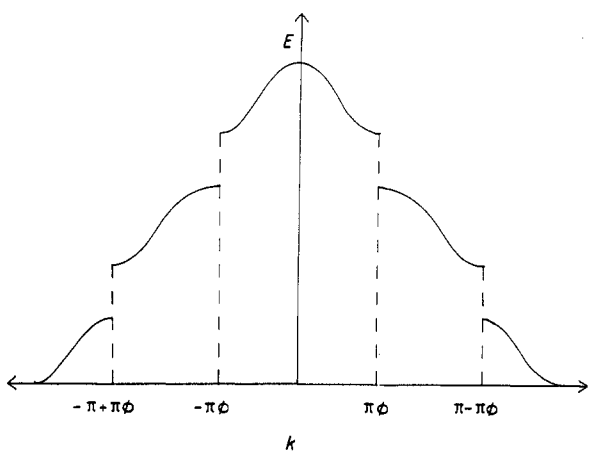

Figure 2. The spectrum of Harper's equation in the lowest order of perturbation theory.

(where the identity $\{x\} = x - [x]$ has been used). This can be written in the form

$$\tilde{k} = k/\phi = \pm\pi\left\{n\left\{\frac{1}{\phi}\right\}\right\} \qquad \text{or} \qquad \pm\pi\left(1 - \left\{n\left\{\frac{1}{\phi}\right\}\right\}\right) \tag{14}$$

where the following identity, valid for any integer n and any x, was used: $\{n\{x\}\} = \{nx\}$. This puts the equation (14) into exactly the same form as (10) but with m replaced by n and with a rescaled value of ϕ given by $\{1/\phi\}$. Now, in the description provided by (10) the gaps decrease with increasing m, which can take any integer value. In (14) n is also any positive integer and the associated gap also decreases with increasing n because by (12) increasing n implies increasing m. A similar argument carries through for the lower side cluster. It therefore appears that the description of how the gaps appear order by order in perturbation theory in the side clusters is identical to that for the full spectrum. The side clusters can thus, within this description, be regarded as rescaled full bands with ϕ transformed to $R(\phi)$ and k scaled by $1/\phi$. The hierarchical clustering is then implied by the recursive application of the above scheme. This proves the first part of Hofstadter's clustering rules.

To prove the rules for the central cluster is technically more difficult as the step length does not have a simple relation with the width of the k-space region occupied by the central cluster. A more compact way of doing this calculation, which will be necessary to derive the rule for the central cluster, is to write the condition for the walks to 'land' in $(-\pi\phi, \pi\phi)$ as an inequality $\pi n - \pi\phi < m < \pi\phi + \pi n$. Once this is simplified to $n/\pi - 1 < m < 1 + n/\phi$ it is clear that the solution is just (12) as found above. The condition for the k_m of (10) to lie in one of the two regions of k-space occupied by the central cluster is

$$\pi N(n) + \pi\phi < \pi m\phi < \pi N(n) + \pi - \pi\phi \tag{15}$$

where $N(n) = n$ takes into account the region $(\pi\phi, \pi - \pi\phi)$ and $N(n) = n - 1$ the region $(-\pi + \pi\phi, -\pi\phi)$. Equation (15) can be rearranged as

$$\frac{N(n)}{\phi} + 1 < m < \frac{N(n)+1}{\phi} - 1. \tag{16}$$

Making the substitution $x = \phi/(1-2\phi)$ and denoting $m - 2N - 1$ by p, this becomes $N < px < N + 1$ which has the solution $N = [px]$ and p can be any non-negative integer. It follows that $m = p + 2[px] + 1$ and hence the gaps occur at

$$k = \pm\pi(p\phi + 2\phi[px] + \phi - [px]) \tag{17}$$

$$\pm(p\phi + 2\phi[px] + \phi - [px] - 1) \tag{18}$$

where we have subtracted $N\pi$ to put these into the reduced zone. These equations simplify to

$$k = \pm\pi(\phi + (1-2\phi)\{p\{x\}\})$$
$$\pm\pi(\phi - 1 + (1-2\phi)\{p\{x\}\}). \tag{19}$$

The plus (minus) signs correspond to the right (left) sections of the central band.

We must now rescale k by a shift of $k_0 = \pm\pi\phi$ to bring the left and right sections of the central band together and then divide by $(1-2\phi)$ to dilate the k-space region of the resulting composite band into an equivalent full band. The central gaps then occur at

$$\tilde{k} = \frac{k - k_0}{1-2\phi} = \pm\pi\{p\{x\}\} \qquad \text{or} \qquad \pm\pi(1 - \{p\{x\}\}) \tag{20}$$

in the equivalent full band. Because m increases with increasing p, the gap magnitude decreases with increasing p. Comparing (20) with (10) shows that the rescaled value ϕ'' of the variable ϕ is

$$\phi'' = S(\phi) = \{x\} = \left\{\frac{\phi}{1-2\phi}\right\}. \tag{21}$$

By the same argument as for the side clusters this transformation can be applied recursively to generate the hierarchical structure of the spectrum. This completes the proof of the clustering rules.

4. Simple scaling theory

The form of the clustering rules suggests that they are the ϕ scaling equations of a renormalisation group (RG) transformation of Harper's equation. They have not yet been incorporated into such a transformation but by using a few simple ideas and some numerical results a much fuller description of the spectrum can be given.

Firstly, the fixed-point structure of the ϕ scaling equations can be investigated. It turns out that the only solutions of the equations $S(\phi^*) = \phi^*$ and $R(\phi^*) = \phi^*$ are

$$\phi_g = [1, 1, 1, \ldots] \qquad \text{and} \qquad \phi_s = [2, 2, 2, \ldots] \tag{22}$$

which are colloquially referred to as the golden and silver means respectively. Their mirror images about $\phi = \frac{1}{2}$ are also fixed points. For these values of ϕ it will be seen that the spectrum is exactly self-similar but this will be discussed later. In addition to the ϕ rescalings there are, of course, the k rescalings derived in the previous section and λ rescalings which have not been derived. It is possible to conjecture scaling equations for λ, by carrying out a decimation transformation on (1) with scale factor b. The leading term from such a transformation for λ is

$$\lambda' = \lambda^b \tag{23}$$

and the wavevector transformation coupled with this is

$$k' = bk. \tag{24}$$

Going back to the proof of the clustering rules we see that $b = 1/\phi$ or $1/(1-2\phi)$ must be chosen. These equations are of limited value as they stand because there is no simple way to obtain an energy scaling equation. However, for the special case where $E = 0$ and ϕ is the Golden Mean it is possible to derive one of the results of Ostlund and Pandit (1984). At $E = 0$ we are dealing with the central cluster so that $b = 1/(2\phi - 1) = 1/\phi^3$ since $\phi^2 + \phi = 1$. Defining $\tau = \lambda - 1$ as the 'distance' from criticality, (23) becomes

$$\tau' = b\tau \tag{25}$$

and also

$$n' = n/b. \tag{26}$$

This leads to the following scaling form of the wavefunction for ϕ_g at $E = 0$:

$$\psi_n(\tau) = \psi_{n/b}(b\tau) \tag{27}$$

which was observed numerically by Ostlund and Pandit (1984). Another result which can be derived with this simple theory is the expression for the correlation length. From (25) we can write $\xi' = \xi/b$. This allows us to write down the functional equation $\xi(\lambda^b) = \xi(\lambda)/b$ which has the solution $\xi = 1/\ln \lambda$. At the moment energy scalings have not been incorporated into the scaling theory of Harper's equation analytically but, by using numerical calculations, some interesting results can be obtained as we shall see in the next section.

5. Cantor set structure of the spectrum

Numerical diagonalisation using the Lanczos algorithm allows the spectrum of Harper's equation to be calculated with comparative ease for quite high orders of commensurability. Of particular interest are the results for ϕ_g and $\lambda = 1$ since this is a fixed point of the combined transformation for ϕ (the clustering rules) and λ (equation (23)). To do this we need to make use of the rational approximants ϕ_n of ϕ_g which are defined by $\phi_n = F_{n-1}/F_n$ where F_n are the Fibonacci numbers given by $F_{n+1} = F_n + F_{n-1}$, $F_0 = F_1 = 1$. For a given value of the denominator these numbers can be shown to be the 'best' approximants to ϕ_g. Applying the clustering rules to ϕ_n gives $R(\phi_n) = \phi_{n-2}$ and $S(\phi_n) = \phi_{n-3}$ which shows that in the incommensurate limit $n \to \infty$ both the central and side clusters will have the same structure as the spectrum as a whole. Our numerical calculations confirm this not only as a qualitative prediction but show that the central and side clusters are, in fact, rescaled versions of the spectrum as a whole. In detail they showed that the ratio of the bandwidth of the full spectrum for ϕ_n to the bandwidth of the side cluster for ϕ_{n+2} rapidly approaches a limiting value of $\alpha = 7.59$. There is a similar result for the ratio of the bandwidth of the full spectrum for ϕ_n to the bandwidth of the central cluster for ϕ_{n+3} but with a different scale factor $\beta = 13.74$. These results are suggestive of Cantor set behaviour, but are not sufficient. If now the ratios of the widths (maximum energy of the highest band less the minimum energy of the lowest band) of the full spectrum to those of the side and central clusters are calculated, they are found to be just α and β as for the bandwidths. This provides compelling evidence that the full spectrum can be described as being either a central cluster 'stretched' by β or a side cluster 'stretched' by α, and hence is a Cantor set. The contraction of the widths suggests that all the individual bands within the spectrum are contracted by the same factor and hence that the scale invariance is complete. This is a remarkable property which appears to be restricted to the fixed point values of the clustering rules. In the case of the 'silver' mean the scale factors α and β are 5.77 and 39.59 respectively.

The fractal dimension D_f of the spectrum can now be calculated in terms of the scale factors α and β. The 'length' or measure of the full spectrum will be designated $M(L)$. Using the scale invariance of the spectrum we obtain

$$M(L) = 2M(L/\alpha) + M(L/\beta) \tag{28}$$

and assuming that $M(L) \propto L^{D_f}$ we can deduce that the fractal dimension is given by the equation

$$1 = 2/\alpha^{D_f} + 1/\beta^{D_f}. \tag{29}$$

This equation can be solved for D_f using the scale factors quoted above to give 0.497 for ϕ_g and 0.496 for ϕ_s. The proximity of both of these values to 0.5 will be discussed later.

To complete this discussion of the spectrum it is interesting to write equations for the bandwidths. Let the bandwidth of the spectrum for ϕ_n be written as B_n. The previous work showed that

$$B_n = 2B_{n-2}/\alpha + B_{n-3}/\beta. \tag{30}$$

This can be solved by substituting $B_n \propto x^n$ and using the positive root of the resulting cubic equation. This gives, to sufficient accuracy,

$$B_n \propto \phi^n \propto 1/F_n = 1/q. \tag{31}$$

This shows that the algebraic fall-off of the bandwidth with q (Thouless 1983) is associated with the scaling of the spectrum.

An alternative approach to the calculation of the scale factors α and β, which is more in the spirit of the derivation of the clustering rules, is to use Brillouin-Wigner perturbation theory. Care has to be taken, however, because using equation (11) it is clear that the gap edges occur at $E = \pm 2\cos(\pi\phi) \pm \lambda$ and will overlap for $\lambda = 1$ which is the interesting case. By going to higher order in perturbation this problem can be eliminated. The expansion for the energy in Brillouin-Wigner perturbation is

$$E(k) = E_0(k) + V_{k,k} + \sum_{k'} \frac{|V_{kk'}|^2}{E(k) - E_0(k')} + \sum_{k',k'' \neq k} \frac{V_{kk'}V_{k'k''}V_{k''k}}{(E(k) - E_0(k'))(E(k) - E_0(k''))} + \ldots. \tag{32}$$

The essential difference between this and the usual Rayleigh-Schrödinger perturbation theory is that the exact energy $E(k)$ occurs in the denominators. This eliminates problems arising from vanishing denominators at the expense of producing a polynomial equation to solve in E. It is, however, a convenient framework for doing degenerate perturbation theory beyond lowest order. When applying this expansion to Harper's equation the potential will be taken to be $2\lambda\cos(2\pi\phi n + \nu)$, which only has two non-vanishing denominators: $V_{k,k\pm 2\pi\phi} = \lambda\exp(\pm i\nu)$. Notice that the squared moduli are independent of ν. This can be directly used to prove the result that for irrational ϕ the spectrum is independent of ν, as all odd-order matrix elements in the expansion vanish. Using the fact that the product of matrix elements for the even orders is λ^n and the expression $E_0(k) = 2\cos(k)$ for the unperturbed energy the second-order result for the dispersion near a gap is

$$E(k) = E_0(k) + \lambda^2\left(\frac{1}{E(k) - E_0(k + 2\pi\phi)} + \frac{1}{E(k) - E_0(k - 2\pi\phi)}\right). \tag{33}$$

This is a cubic equation for E and is, in fact, the same as that which could have been derived by diagonalising the sub-Hamiltonian formed by taking matrix elements of (8) in the sub-space of the states $|k\rangle$, $|k \pm 2\pi\phi\rangle$. It is the first term in the square brackets of (33) which is responsible for the formation of the gaps at $k = -\pi\phi$, $\pi - \pi\phi$ while the second is responsible for those at $k = \pi\phi$, $-\pi + \pi\phi$. Selecting any one of these terms yields the result $\pm 2\cos(\pi\phi) \pm \lambda$ for the gap widths. To calculate the location of the gap edges to higher order at $k = \pi\phi$, say, it is convenient to only retain those terms in the perturbation expansion which would have diverged had ordinary perturbation theory been used. These will clearly be the dominant terms in contributing to the

formation of the gaps. Up to fourth order in λ the expansion for E around $k=\pi\phi$ is, then,

$$E(k)=E_0(k)+\frac{\lambda^2}{E(k)-E_0(k-2\pi\phi)} + \frac{\lambda^4}{(E(k)-E_0(k-2\pi\phi))^2(E(k)-E_0(k-4\pi\phi))}. \tag{34}$$

This can be solved numerically at $k=\pi\phi$ quite simply as it is only a quartic equation and gives $E=0.201$, 1.841 for the gap edges if ϕ is the golden mean and $\lambda=1$. A similar calculation for $k=\pi-\pi\phi$ leads to the same answers but with the opposite sign, as expected. To find the width of the entire spectrum it is sufficient to use the full second-order result and solve for $E(k=0)$. This leads to the following result:

$$\Delta E=2(1+\cos(2\pi\phi)+2\sqrt{[(1-\cos(2\pi\phi))^2+\lambda^2]} \tag{35}$$

which gives 5.01 for ϕ_g and $\lambda=1$. This is in good agreement with our numerical results. Combining this result with the previous ones for the gap edges we can calculate the energy scale factors (for the golden mean) as $\alpha=7.57$ and $\beta=12.46$ which compare favourably with the numerical values. Similar calculations for the silver mean yield results of comparable accuracy.

It was pointed out earlier that the fractal dimension for both ϕ_s and ϕ_g turned out very close to one half. This raises the question whether the fractal dimension of the spectrum is equal to a half for a transcendental number, i.e. an irrational not having a periodic continued fraction representation. These numbers can be considered as being 'typical' and do not have a periodic continued fraction representation, and hence are not fixed points of either of the clustering formulae. To test this idea we considered $\sqrt[3]{2}-1$ which has the continued fraction representation

$$\sqrt[3]{2}-1=[3, 1, 5, 1, 1, 4, 1, 1, 8, \ldots]. \tag{36}$$

By calculating the spectrum numerically for $\phi=1120/4309$ and by using the definition

$$D_f=\lim_{\rho\to 0}-\frac{\ln N(\rho)}{\ln(\rho)} \tag{37}$$

(where $N(\rho)$ is the minimum number of rulers of length ρ to cover the fractal) we found that $D_f=0.518$. Again, this is very close to 0.5 and diagonalisation of a larger matrix may well yield a result even closer to a half. It is possible to argue that $D_f=0.5$ for all irrational values of ϕ. To do this remember that the bandwidth of the spectrum tends to zero as $1/q$. With $\phi=p/q$ there are q bands and their average width will be $1/q^2$. Going back to the definition (37) we can replace N with q and ρ with $1/q^2$ to give $D_f=0.5$. Of course this calculation is only suggestive but together with the numerical results it provides strong evidence that the fractal dimension is indeed equal to a half for all irrational values of ϕ.

6. Quasicrystal models

Equations similar to Harper's equation but with a different quasiperiodic potential have appeared recently in connection with the study of phonons in quasicrystals. Of

particular interest is the so called Fibonacci chain model (e.g. Ogadaki and Nguyen 1986). In its simplest form it is a generalisation of the diatomic chain with the atoms of masses m_a and m_b occurring in a quasiperiodic fashion in the non-stoichiometric ratio ϕ_g. The unit cell of the lattice is generated by iterating the following recurrence relation m times:

$$\begin{aligned} A &\to AB \\ B &\to A. \end{aligned} \tag{38}$$

Starting from B, say, this generates a word consisting of A and B which specify the location of the atoms of masses m_a and m_b in the chain. The periodic repetition of this word then generates the system. Notice that at the mth stage of the construction the repeating period in the system is just the Fibonacci number F_m, whence the name of the model. The quasiperiodic (incommensurate) limit is approached as $m \to \infty$.

The results of Kohmoto (1983) show that the Hofstadter clustering rules are valid for this model as well. Of course, these rules are of less interest in the Fibonacci chain model as the incommensurability is fixed at the golden mean, but nonetheless it is worthwhile to investigate their generality. For Harper's equation the essential ingredient was that the gaps appeared at the k values given by

$$k_l = \pm\pi\{l\pi\} \qquad \text{or} \qquad \pm\pi(1-\pi\{l\pi\}) \tag{39}$$

where l labels the gaps in order of decreasing magnitude. This must now be shown to be true for the Fibonacci chain model. In this model the potential can be written as

$$m_n = \sum_l V_l \exp(2\pi i \phi l n) \tag{40}$$

where all of the Fourier components indexed by l are present. Because of this, in lowest-order perturbation theory, gaps will open at all of the wavevectors given in (39) when put into a reduced zone scheme and with widths approximately $2|V_l|$ (Lifshitz and Pitaevski 1980). A basic point here is that the coefficients V_l depend upon the mass difference δ between the atoms (see appendix) so that for small δ the simple perturbative result for the gaps is valid. It remains then to show that the amplitudes decrease with increasing l. This can be done by explicitly Fourier transforming the potential. The result (which is given in an appendix) is, however, quite complicated and it is better to proceed by more intuitive arguments. After subtracting the average value of the potential it is clear that the dominant Fourier component is $\cos(2\pi\phi_g n)$ simply because the ratio of A to B sites is ϕ_g. It can then be argued that the higher harmonics $\cos(2\pi\phi_g l n)$, which will form gaps at k_l, will have amplitudes that decrease with increasing l. This is the basic result which is required to prove the clustering rules and the calculation proceeds as for Harper's equation. In conclusion, the result that the gap which forms at k_l is a decreasing function of l follows here because though the gap at k_l can be produced by an onward 'walk in k space' whose rth step is $\pi l_r \phi$ such that $\Sigma\, l_r = l$, the amplitude V_{l_r} of the Fourier component $2\pi\phi l_r$ of the potential decreases with l_r. Moreover, since each amplitude is proportional to δ, for small δ the dominant term comes from the one-step walk of length $\pi l \phi$ whose strength again decreases with l. This is quite distinct from the case of Harper's equation where the result is a consequence of the gap at k_l being formed by the lth order of perturbation with a single Fourier component.

7. Perspective

It is clear that the proof of the clustering rules can be regarded as part of a renormalisation group transformation of Harper's equation for which the energy scaling equations have to be calculated numerically. However, it seems likely that a calculation using the method of Wilkinson (1984) but using perturbation theory instead of WKB theory would provide a complete transformation and resolve, finally, the question of the fractal dimension of the spectrum. Work is in progress in this direction.

Acknowledgments

This work was supported in part by a postgraduate grant from the Science and Engineering Research Council (UK). The authors would like to thank B W Southern for a critical reading of the manuscript.

Appendix

In this appendix we calculate the discrete Fourier transform of the potential m_n occurring in the Fibonacci chain model and show that it is proportional to the mass difference $\delta = m_a - m_b$. This indicates that for small δ a perturbative approach is valid.

The quantity which we calculate is

$$m(q, k) = \sum_{n=1}^{N} m(q, n) \exp(\mathrm{i}kn) \tag{A1}$$

where q denotes the order of the hierarchy being considered and $N = F_q l$ (l = number of unit cells in the system). By splitting the summation up into l groups of F_q terms it becomes

$$m(q, k) = \tilde{m}(q, k)\left[\frac{1-\exp(\mathrm{i}kN)}{1-\exp(\mathrm{i}kF_q)}\right]$$

$$\tilde{m}(q, k) = \sum_{n=1}^{F_q} m(q, n) \exp(\mathrm{i}kn). \tag{A2}$$

Now $\tilde{m}$ will be calculated by deriving a difference equation for it. Define:

$$\begin{aligned} \theta(n) &= 1 \qquad n \geqslant 0 \\ &= 0 \qquad n < 0. \end{aligned} \tag{A3}$$

Then

$$\tilde{m}(q+1, n) = \tilde{m}(q, n)\theta(F_q - n) + \tilde{m}(q-1, n-F_q)\theta(n-1-F_q) \qquad n = 1, \ldots, F_{q+1} \tag{A4}$$

where the decomposition property of the potential has been used to write the qth word as a combination of the $(q-2)$th and $(q-1)$th word. The following equation can now be written for $\tilde{m}$:

$$\tilde{m}(q+1, k) = \tilde{m}(q, k) + \exp(\mathrm{i}kF_q)\tilde{m}(q-1, k). \tag{A5}$$

This has the particular solution

$$\tilde{m}(q, k) = 1 - \exp(\mathrm{i}kF_q) \tag{A6}$$

corresponding to the pure chain limit. This suggests substituting

$$\tilde{m}(q, k) = \mu(q, k)(1 - \exp(\mathrm{i}kF_q)) \tag{A7}$$

and if we define

$$\Delta(q, k) = \mu(q, k) - \mu(q-1, k) \tag{A8}$$

then

$$\Delta(q+1, k) = \Delta(q, k) f(q, k) \tag{A9}$$

where

$$f(q, k) = \left(\frac{\exp(\mathrm{i}kF_{q+1}) - \exp(\mathrm{i}kF_q)}{1 - \exp(\mathrm{i}kF_{q+1})} \right) \tag{A10}$$

so that

$$\mu(q, k) = \mu(0, k) + \sum_{j=1}^{q} \Delta(j, k)$$

$$\mu(q, k) = \mu(0, k) + \Delta_1 \left[1 + \sum_{j=1}^{m-1} \prod_{l=1}^{j} f(l, k) \right]. \tag{A11}$$

Δ_1 and μ_0 can be explicitly calculated to be

$$\Delta_1 = \frac{m_a - m_b}{\exp(-\mathrm{i}k) - 1} \qquad \mu_0 = \frac{m_b}{\exp(-\mathrm{i}k) - 1}. \tag{A12}$$

This solution shows that the coefficient of the quasiperiodic part of the potential depends upon the mass difference δ as stated in the text.

References

Aubry S and André G 1980 *Ann. Israel Phys. Soc.* **3** 133
Azbel M Ya 1964 *Zh. Eksp. Teor. Fiz.* **46** 929
Bell S C 1986 *D. Phil. Thesis* Oxford University
Harper P G 1955 *Proc. Phys. Soc.* A **68** 874
Hofstadter D R 1976 *Phys. Rev.* B **14** 2239
Kohmoto M 1983 *Phys. Rev. Lett.* **51** 1198
Lifshitz E M and Pitaevski L P 1980 *Statistical Physics* (New York: Pergamon)
Luck J M and Petritis D 1985 *Preprint* Saclay
Ogadaki T and Nguyen D 1986 *Phys. Rev.* B **33** 2184
Ostlund S and Pandit R 1984 *Phys. Rev.* B **29** 1394
Peierls R E 1933 *Z. Phys.* **80** 763
Stinchcombe R B and Bell S C 1987 *J. Phys. A: Math. Gen.* **20** L739
Thouless D J 1983 *Phys. Rev.* B **28** 4272
Thouless D J, Kohmoto M, Nightingale M P and denNijs M 1982 *Phys. Rev. Lett.* **49** 405
Wilkinson M 1984 *Proc. R. Soc.* A **391** 305
Ying S C 1981 *Phys. Rev.* B **3** 4160