Total bandwidth for the Harper equation. III. Corrections to scaling

1991 J. Phys. A: Math. Gen. 24 4055

(http://iopscience.iop.org/0305-4470/24/17/022)

J. Phys. A: Math. Gen. **24** (1991) 4055–4066. Printed in the UK

Total bandwidth for the Harper equation: III. Corrections to scaling

D J Thouless and Yong Tan

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

Received 21 January 1991

Abstract. Earlier work on the spectral measure for the Harper equation (discrete Mathieu equation) showed that, if there is a large common period p for the lattice and the sinusoidal potential, the spectral measure scaled as C/p, where C is a constant independent of the number of oscillations q of the potential in p lattice spacings. In the present work the corrections to this scaling law are found for $q=1$ and $q=2$. For $q=1$ it is shown that the corrections to scaling are logarithmic functions of p, while for $q=2$ the leading corrections are reduced by a factor p^{-2}. Analytical and numerical support is given for the assertion that this difference between odd and even values of q persists for higher values of q provided they are substantially less than $\sqrt{p}$.

1. Introduction

The Harper equation (Harper 1955)

$$V_1 c_{n-1} + V_1 c_{n+1} + 2V_2 \cos(2\pi\phi n + \nu) c_n = E c_n \tag{1.1}$$

represents a discrete wave equation (tight binding model) with two competing periods, such as could arise in a one-dimensional quasicrystal. It was originally derived as the equation for an electron in a two-dimensional tight-binding crystal perturbed by a weak magnetic field. In fact the same equation is obtained also for free electrons in a uniform magnetic field perturbed by a two-dimensional sinusoidal potential. For these latter two cases rectangular symmetry of the lattice gives $V_1 \neq V_2$, whereas the case of square symmetry leads to $V_1 = V_2$.

This equation has some very interesting properties (Sokoloff 1985, Thouless 1990a). For irrational ϕ the energy gaps are everywhere dense, so that the spectrum is a Cantor set (Azbel 1964). Numerical calculation of the spectrum for the case $V_1 = V_2$ by Hofstadter (1976) has shown a beautiful self-similar structure, with a spectral measure that shrinks to zero for irrational ϕ. Aubry and André (1980) have shown that for $V_1 > V_2$ all eigenstates are extended, while for $V_1 < V_2$ all eigenstates are localized with an energy-independent localization length for irrational ϕ. Furthermore the measure of the spectrum is $4|V_1 - V_2|$, and shrinks to zero at the critical point $V_1 = V_2$ where the localization length becomes infinite.

There are a number of papers in which scaling, perturbative, or semiclassical arguments have been applied to this equation (Sokoloff 1985, Wilkinson 1987, Bell and Stinchcombe 1989). In an earlier paper by one of us (Thouless 1983), which we call BW1, the way in which the measure of the spectrum approaches its limiting form for irrational ϕ when $\phi = q/p$ and p is large was studied by a combination of finite size scaling arguments and numerical studies of the spectrum. The finite size scaling form of the measure of the spectrum is

$$W(V_1, V_2, q/p) \sim \frac{V_2}{p} f\left(\frac{p(V_2 - V_1)}{V_2}\right) \tag{1.2}$$

where the scaling function $f(x)$ was found to have a numerical value close to 9.33 at the origin. In a recent paper (Thouless 1990b), which we call BW2, an analytic expression for the scaling function was obtained for the particular case $q = 1$, p odd, namely

$$f(x) = 4x + \frac{8}{\pi} \sum_{n=0}^{\infty} (-1)^n \frac{\mathrm{e}^{-(n+1/2)x}}{(n+\frac{1}{2})^2}. \tag{1.3}$$

Two questions raised by the results of BW1 which were not answered in BW2 are why this scaling function depends only on the denominator p and not on the numerator q, and why the corrections to scaling are so sensitive to the value of q.

In this paper, section 2 contains a brief summary of some of the numerical results on the sum of the widths of the bands (measure of the spectrum) for different q/p. In particular there are graphs showing how for fixed denominator the corrections to scaling vary by several orders of magnitude as the numerator is changed.

In section 3 the argument of BW2, for the case $q = 1$, $V_1 = V_2$, is repeated more carefully using the WKB approximation to improve the approximation for the Green function. This improvement gives corrections to the scaling formula for pW which, for very large p, are of order $(\log p)^{-2}$. In fact the expansion in inverse powers of the logarithm converges rather slowly, but we have an explicit expression for these corrections to scaling which, when it is subtracted from the numerical results, leaves only corrections of order p^{-2}.

In section 4 a somewhat more general method is introduced. This method relies on the fact that only bands close to $E = 0$ make significant contributions, and in that neighbourhood the WKB method can be used to study the wavefunctions. In this way we get explicit approximations for the eigenvalues, which can be expressed as the zeros of a small number of explicit functions of analytic form. We take the logarithm of the ratio of the function whose zeros give the tops of the sub-bands to the function whose zeros give the bottoms and integrate this round a closed curve to get an explicit expression for the measure of the spectrum. This method is used first for the case $q = 1$, where the same result is obtained that was obtained in section 3, and is then used for a detailed examination of the case $q = 2$, p odd. We find that the corrections to scaling vanish in this approximation for $q = 2$, leaving only corrections of order p^{-2}, which we have neglected throughout. These results are in accord with our numerical work, which showed that the corrections to scaling were particularly large for $q = 1$, and were proportional to p^{-2} for $q = 2$.

In section 5 there is some discussion of how these results can be generalized to other small values of the numerator. We do not carry out the calculations for any other case in detail, as the set of equations which must be derived gets larger as q gets larger, but we show that the qualitative difference between odd and even values of q should persist so long as q is sufficiently small that the WKB method is valid throughout the region of interest.

2. Numerical results

In BW1 numerical work appeared to show that, for the case $V_1 = V_2$,

$$\lim_{p\to\infty} pW(q/p) = C \tag{2.1}$$

where C is a constant which does not depend on how q is related to p. It was found in BW2 that this common limit C is $32G/\pi$, where G is the Catalan constant. For odd p the convergence to the limit is always from below, while for even p the convergence is from above. The speed of convergence to this limit does depend strongly on the form of q. For $q=2$ and p odd, and for p even and $q=p/2-1$, the correction terms were of order p^{-2}, while for $q=1$ the correction terms were found to fall off slower than $p^{-1/3}$. For some other cases examined in BW1, such as the Fibonacci sequence, the convergence also seemed to be fairly fast.

To illustrate this dependence on q we show in figure 1 the value of $(-1)^p(pW-C)$ for different q with p fixed. This figure shows the results at $V_1 = V_2$ for $p=199$ and $p=194$, but the results for other values of p and for $V_1 \neq V_2$ are similar. The convergence is slow for small odd values of q, both for odd and even p, and it is particularly slow for $q=1$. For even p the convergence is fast for q/p close to $\frac{1}{2}$, while it is slow for odd p. In general there is slow convergence for p/q close to any integer for odd p or to an odd integer for even p, and other peaks in $|pW-C|$ come for values of q/p close to other simple fractions.

3. Correction for $q=1$

The case $q=1$ is anomalous in two respects. The value of $|pW-C|$ is much larger than for any other value of q, and an exact expression was derived for this expression in terms of the Green function in BW2. The total bandwidth for the Harper equation with $\phi = 1/p$ for p odd is calculated by integrating

$$\ln[zG_{00}^{++}(z)G_{ss}^{-+}(z)/G_{ss}^{--}(z)] \tag{3.1}$$

along the imaginary axis. Here $s=(p-1)/2$ and the superscripts of the Green functions $G(z)$ denote the odd and even boundary conditions at 0 and $p/2$. The Green functions can be written in terms of solutions of (1.1) for $E=z$. Then the integral is

$$-\frac{2}{\pi \mathrm{i}}\int_{-\mathrm{i}\infty}^{\mathrm{i}\infty} \ln\frac{za_0(b_{s+1}+b_s)}{(a_{-1}-a_1)(b_{s+1}-b_s)}\,\mathrm{d}z \tag{3.2}$$

where a_n and b_n are solutions which satisfy the boundary conditions $a_{s+1}=a_s$, $b_0=0$ respectively.

In BW2 these functions were evaluated by approximating the difference equation by a differential equation, and replacing the cosine in (1.1) by parabolas centred at the maximum and minimum (for a_n and b_n respectively). Most of the errors introduced into the evaluation of pW by these approximations are of order p^{-2}, but there is a non-perturbative correction due to the use of the wrong boundary conditions when the cosine curve is replaced by a parabola. Figure 2 shows the results of numerical evaluations of the Green function $G_{00}^{++}(E)$ with $E=\mathrm{i}z$ and $V_1 = V_2$. The discrepancy at small z between the exact $G_{00}^{++}(\mathrm{i}z)$ and the approximation used in BW2 can be observed.

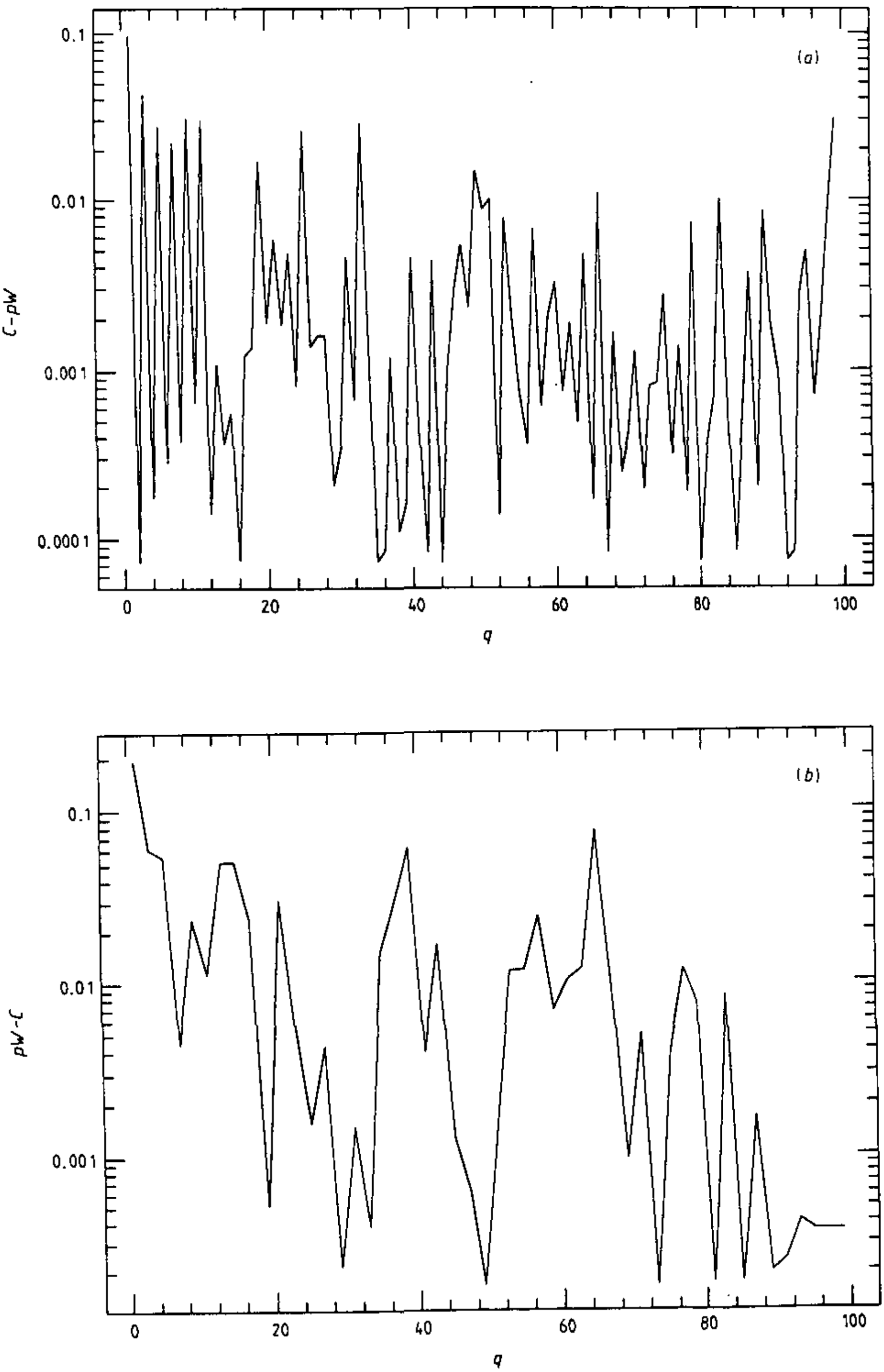

Figure 1. The correction to the scaling $(-1)^p(pW-C)$ at the critical point $V_1=V_2=1$ is plotted on the logarithmic scale as a function of the the numerator q with a common odd denominator ($p=199$) in (a) and even denominator ($p=194$) in (b) respectively.

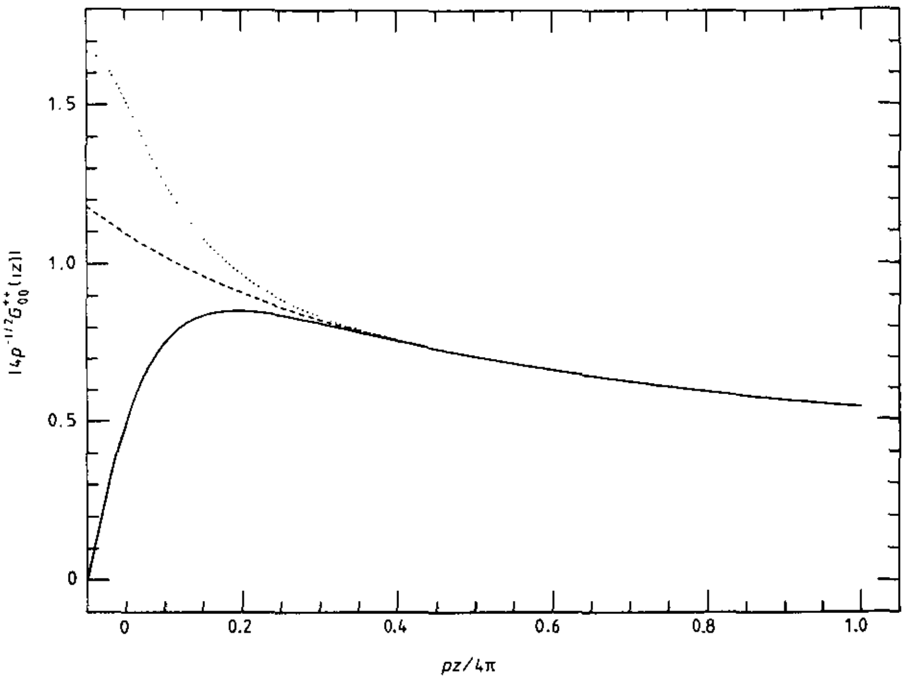

Figure 2. The modulus of the rescaled Green function $4p^{-1/2}G_{00}^{++}(E)$ for the imaginary energy $E=\mathrm{i}z$ is plotted against the rescaled energy $pz/4\pi$. The dsahed line is the result in BW2, the solid line is for $p=4n+3$ ($p=99$), and the dotted line is for $p=4n+1$ ($p=101$). The Green function is calculated as the inverse of $(\mathrm{i}z-H^{++})$ where H^{++} is the matrix obeying even–even boundary condition at 0 and $p/2$. The explicit expression for H^{++} is given in BW1. Since H^{++} is tridiagonal, the numerical value of its determinant, and therefore the diagonal elements of the Green function, is very easy to obtain.

In order to obtain boundary conditions which are better than those used in BW2 we use the WKB approximation for the solutions to (1.1) between the two turning points at 0 and $p/2$, and use the approximation of the cosine curve by a parabola to obtain a connection formula. At the turning points the two independent solutions, c_n^+ and c_n^-, for the imaginary energy $E=\mathrm{i}z$, are given in BW2 as

$$\begin{aligned} c_n^+ &= (-1)^n D_{-(1/2)-(pz/4\pi)}\left(\mathrm{e}^{-\mathrm{i}(\pi/4)}2\sqrt{\frac{\pi}{p}}\,n\right) \\ c_n^- &= (-1)^n D_{-(1/2)+(pz/4\pi)}\left(\mathrm{e}^{\mathrm{i}(\pi/4)}2\sqrt{\frac{\pi}{p}}\,n\right) \end{aligned} \tag{3.3}$$

for n around zero and

$$\begin{aligned} c_n^+ &= \alpha D_{-(1/2)+(pz/4\pi)}\left(\mathrm{e}^{-\mathrm{i}(\pi/4)}2\sqrt{\frac{\pi}{p}}\left(\frac{p}{2}-n\right)\right) \\ c_n^- &= \beta D_{-(1/2)-(pz/4\pi)}\left(\mathrm{e}^{\mathrm{i}(\pi/4)}2\sqrt{\frac{\pi}{p}}\left(\frac{p}{2}-n\right)\right) \end{aligned} \tag{3.4}$$

for n around $p/2$, where α, β are constants which can be calculated by WKB approximation when the two points at $n=0$ and $n=p/2$ are connected, and $D_\nu(z)$ is the parabolic cylinder function (Whittaker and Watson 1952).

In the following the WKB solution $c_n = e^{\phi(n)}$ is used in (1.1) which then becomes

$$2\exp\left(\frac{1}{2!}\phi''(n)+\frac{1}{4!}\phi^{(4)}(n)+\ldots\right)\cosh\left(\phi'(n)+\frac{1}{3!}\phi^{(3)}(n)+\ldots\right)$$
$$= iz - 2\cos\left(\frac{2\pi}{p}n\right). \tag{3.5}$$

If we assume $\phi'(n)$ is dominant, the solutions of first order are

$$\phi'(n) \approx \cosh^{-1}\left(\frac{iz}{2} - \cos\left(\frac{2\pi}{p}n\right)\right). \tag{3.6}$$

We are interested in the case of z small, p large, and $\sin(2\pi n/p)$ not too small (not too close to the turning points). To first order in z we have

$$\phi'(n) \approx \mp i\pi \pm i\frac{2\pi n}{p} \mp \frac{z}{2\sin(2\pi n/p)} - \frac{\pi}{p}\cot\left(\frac{2\pi n}{p}\right) \tag{3.7}$$

where the last term comes from the second derivative $\phi''(n)$. Furthermore it can be shown that the higher-order terms in the two small parameters, z and $1/p$, can be neglected in comparison with the terms which we have kept. Integrating (3.7), we have the WKB solutions

$$c_n^{\pm} = c_{\pm}\exp\left[\mp i\pi n \pm i\frac{\pi}{p}n^2 \mp \frac{pz}{4\pi}\ln\left(\tan\frac{\pi n}{p}\right) - \tfrac{1}{2}\ln\left(\sin\frac{2\pi n}{p}\right)\right]. \tag{3.8}$$

The asymptotic expansions of the parabolic cylinder functions of (3.3) and (3.4) are given in essentially the same form by Whittaker and Watson (1952), so one can match the coefficients at the two turning points to get the results

$$\alpha = \exp\left(i\frac{pz}{8} - \frac{pz}{4\pi}\ln\frac{4p}{\pi} - i\frac{\pi p}{4}\right) \qquad \beta = \exp\left(i\frac{pz}{8} + \frac{pz}{4\pi}\ln\frac{4p}{\pi} + i\frac{\pi p}{4}\right). \tag{3.9}$$

From these two solutions we can construct solutions with the required symmetry at 0 and $p/2$. For example, we can formulate a symmetric solution about $n = p/2$ as $a_n = c_n^+ + c_{p-n}^+$; explicitly, for n close to $p/2$,

$$\alpha D_{-(1/2)+(pz/4\pi)}(m) + \alpha D_{-(1/2)+(pz/4\pi)}(-m) \tag{3.10}$$

where $m = e^{-i(\pi/4)}2\sqrt{\pi/p}(p/2-n)$. $D_{-(1/2)-(pz/4\pi)}(-m)$ is also a solution which can be expressed in terms of the parabolic cylinder functions with positive variable, by using the connection formula at the turning point,

$$D_\nu(x) = e^{\nu\pi i}D_\nu(-x) + \frac{\sqrt{2\pi}}{\Gamma(-\nu)}e^{(\nu+1)\pi i/2}D_{-\nu-1}(-ix) \tag{3.11}$$

(Whittaker and Watson 1952). In this way one should be able to evaluate the coefficients in front of two independent solutions, and the solution which satisfies the boundary condition $a_{s+1} = a_s$ can be written as

$$a_n = (1 + e^{(y-(1/2))\pi i})c_n^+ + \left(\frac{\sqrt{2\pi}}{\Gamma(\frac{1}{2}-y)}e^{(y+(1/2))\pi i/2}\frac{\alpha}{\beta}\right)c_n^- \tag{3.12}$$

where $y = pz/4\pi$. Similarly the solution which satisfies $b_0 = 0$ can also be written as

$$b_n = (1 - e^{-(y+(1/2))\pi i})c_n^+ - \left(\frac{\sqrt{2\pi}}{\Gamma(\frac{1}{2}+y)} e^{-(y-(1/2))\pi i/2}\right) c_n^- . \tag{3.13}$$

When these solutions are used to construct the Green functions, they are found to fit the numerical calculations described in the last section very well.

To obtain the total bandwidth, one just substitutes the solutions into (3.2). The value of the parabolic cylinder function at zero is used, we also need the relation between $D_{-1/2-y}(0)$ and $D_{-1/2+y}(0)$, and between their first derivative, which can all be obtained from (3.11). After some manipulations, the rescaled total bandwidth pW is found to be C with an additional integral,

$$-8\int_0^{\Lambda} dy \ln \frac{\dfrac{\pi^2}{[\Gamma(\frac{1}{2}+y)]^4} + (-1)^{p+1}(1+\sin \pi y)^2 \exp\left(-4y \ln \dfrac{4p}{\pi}\right)}{\dfrac{\pi^2}{[\Gamma(\frac{1}{2}+y)]^4} + (-1)^{p+1}(1-\sin \pi y)^2 \exp\left(-4y \ln \dfrac{4p}{\pi}\right)} \tag{3.14}$$

where a factor $(-1)^{p+1}$ is added for the convenience of our following calculations and Λ is a cutoff for the reason that our approximation is valid only for the small y. This integral can be expanded as a series in $1/(\ln 4p/\pi)^2$, with the leading term equal to zero, but the series converges very slowly. However, its numerical value can be found very easily since the integrand peaks sharply somewhere near $y = 0$ and is very small for quite a long range of y; Λ is chosen to lie in this range. Where $y \geqslant 4ep/\pi$ the integrand starts oscillating, but the WKB approximation used to derive the form of the integrand is no longer valid. This integral is in agreement with the numerical results for $C - pW$ to one part in 10^4 generally, in the range $99 \leqslant p \leqslant 299$, and the difference between this result and the numerical one is proportional to p^{-2}, as one would expect from perturbation theory.

4. Explicit approximation for band edges

In this section, we introduce a method which can be applied to more general cases. The method introduced in BW2 and exploited in section 3 of this paper relies on the particular ordering of eigenstates of (1.1) with different symmetries. For $q \neq 1$ the eigenvalues are differently ordered, and the method based on the integration of the Green function along the imaginary axis can no longer be used. Instead, explicit approximate equations are derived for the eigenvalues, and these equations are then used in conjunction with a contour integral to perform the sum which gives the total bandwidth. The difficulty is that the larger the numerator q is the more separate equations have to be treated. However, some general features can already be seen from the treatments of the cases $q = 1$ and $q = 2$ that some tentative generalizations can be made.

4.1. *General formulation*

As was discussed in BW1, the band edges can be found as eigenvalues of (1.1) with particular symmetries, and they form four families. There are those with $\nu = 0$ which are even under reflection about both $n = 0$ and $n = p/2$, or odd under reflection about both points, and we denote these eigenvalues as E_m^{++} and E_m^{--} respectively. There are also those with $\nu=\pi\phi$ which are even under reflection about one point and odd under reflection about the other point, and we denote these as $\mathscr{E}_m^{+-}$ and $\mathscr{E}_m^{-+}$. For odd values of p there are simple relations such as $E_m^{++}=-\mathscr{E}_m^{+-}$, so that only two of the four sets of eigenvalues needs to be considered explicitly. For even p all four classes have to be considered, as negative and positive pairs lie within the same class. This feature makes the case p odd somewhat less cumbersome to deal with, which is why we have concentrated attention on p odd.

It is important that, in the case of p large and q small, the bandwidth can be written in such a way that only eigenvalues near $E=0$ contribute significantly. For $V_1=V_2$ no trick is required for this, as only the central bands have significant width, but for $V_1\neq V_2$ the technique used in BW1 to express $W-4|V_2-V_1|$ in terms of the E^{-+} as well as of the E^{++} and E^{--} must be used. If the eigenvalues are numbered in such a way that the extremal band edges are given by $\pm E_0^{++}$, while the intermediate band gaps are bounded by $E_n^{--}>E_n^{++}$ for $n\leqslant r$ and by $E_n^{--}<E_n^{++}$ for $n>r$, then the results of BW1 imply

$$W=4|V_2-V_1|+4\sum_{n=1}^{r}(E_n^{-+}-E_n^{--})+4\sum_{n=r+1}^{s}(E_n^{-+}-E_n^{++}). \tag{4.1}$$

In section 3, we have obtained the solutions to (1.1) obeying one-point boundary conditions, and application of a second boundary condition to such a solution gives the equation whose solution gives an approximation for the eigenvalues. From such equations we construct an analytic function $\phi(z)$ which has simple poles and zeros at the required eigenvalues b_i and a_i, so that the expression for the bandwidth is written in the form

$$\frac{1}{2\pi\mathrm{i}}\oint z\frac{\phi'(z)}{\phi(z)}\,\mathrm{d}z=\sum(-a_i+b_i). \tag{4.2}$$

4.2. Numerator unity and odd p

In order to illustrate how this method works and examine its validity, we start with the case $\phi=1/p$ with odd p, for which the bandwidth was obtained in section 3. The solutions for imaginary energy with one-point boundary conditions are given in (3.12) or (3.13). If we impose a second boundary condition at the other turning point, namely $a_0=0$ or $a_{-1}=a_1$ for a_n, or $b_{s+1}=-b_s$ for b_n, and make an analytic continuation to the real axis, approximation for the quantized eigenvalues are obtained:

$$\operatorname{Im}\ln\Gamma(\tfrac{1}{2}-\mathrm{i}x)+x\ln\frac{4p}{\pi}-\tfrac{1}{2}\tan^{-1}\sinh(\pi x)=(m-\tfrac{1}{4})\pi \qquad \text{for } E^{--} \tag{4.3}$$

$$\operatorname{Im}\ln\Gamma(\tfrac{1}{2}-\mathrm{i}x)+x\ln\frac{4p}{\pi}+\tfrac{1}{2}\tan^{-1}\sinh(\pi x)=(m-\tfrac{1}{4})\pi \qquad \text{for } E^{++} \tag{4.4}$$

where $x=pE/4\pi$, m is integer, and E^{--} and E^{++} are band edges.

Equations (4.3) and (4.4) give one set of band edges while the other set, denoted by $\mathscr{E}^{+-}$ and $\mathscr{E}^{-+}$, is found by changing the sign of x in these equations. Equivalently the other set is found by replacing $m-\frac{1}{4}$ by $m+\frac{1}{4}$ on the right-hand side of these equations. We name the equations after the substitution as (4.3^-) and (4.4^-) for convenience. The total bandwidth is the sum over all the intervals between two successive band edges. By comparing (4.3) and (4.4) with each other, one may conclude that E_m^{--} is higher than E_m^{++} on the positive real axis, and the ordering is reversed on the negative real axis. If (4.3^-) and (4.4^-) are included, then the spectrum is complete. The conclusion can be drawn that only those bands around the spectrum centre (small x) are appreciably wide; for x large the bands are exponentially narrow and these narrow bands are almost equally spaced. The above arguments have been checked to agree with numerical results.

Now we calculate the total bandwidth by Cauchy integral with the contour from $\mathrm{i}\infty$ to $-\mathrm{i}\infty$ along the imaginary axis and along a semicircle $R\,\mathrm{e}^{\mathrm{i}\theta}$ with $R\to\infty$ and θ from $-\pi/2$ to $\pi/2$, so that the positive zeros are included. Assuming that the contribution from the semicircle vanishes if the integrand becomes zero as $R\to\infty$, the integral along the imaginary axis is left, after integrating in parts, which becomes

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\ln\phi(\mathrm{i}y)\,\mathrm{d}y=\sum(a_i-b_i). \tag{4.5}$$

To construct $\phi(z)$ in (4.2), one nseeds to rewrite (4.3) and (4.4) in such a way that the explicit dependence on the integer m is removed. After some algebra they become

$$f^{--}(x)=\mathrm{i}+\sinh\pi x-\frac{\pi}{[\Gamma(\frac{1}{2}-\mathrm{i}x)]^2}\exp\left(-2\mathrm{i}x\ln\frac{4p}{\pi}\right)=0 \tag{4.6}$$

$$f^{++}(x)=\mathrm{i}-\sinh\pi x-\frac{\pi}{[\Gamma(\frac{1}{2}-\mathrm{i}x)]^2}\exp\left(-2\mathrm{i}x\ln\frac{4p}{\pi}\right)=0 \tag{4.7}$$

while the other set of band edges is given by $f^{--}(-x)=0$, $f^{++}(-x)=0$. Before evaluating the integral, one should be aware that these two equations have zeros of order 2 on the imaginary axis at half-integer multiples of i. To get rid of these we multiply $f^{--}(\mathrm{i}x)$ and $f^{++}(\mathrm{i}x)$ by $[\Gamma(\frac{1}{4}+x/2)]^2$ and $[\Gamma(\frac{3}{4}+x/2)]^2$ respectively. Another factor of $(2/y)^2$ is introduced to ensure that the argument of the logarithm tends to unity at infinity. Provided that there are then no zeros and poles away from the real axis, the total bandwidth pW, for $p=4n+3$, is

$$4\int_{-\infty}^{\infty}\ln\left\{\left(\frac{2}{y}\right)^2\left[\frac{\Gamma(\frac{3}{4}+y/2)\Gamma(\frac{3}{4}-y/2)}{\Gamma(\frac{1}{4}-y/2)\Gamma(\frac{1}{4}+y/2)}\right]^2\frac{f^{++}(\mathrm{i}y)f^{++}(-\mathrm{i}y)}{f^{--}(-\mathrm{i}y)f^{--}(\mathrm{i}y)}\right\}\mathrm{d}y. \tag{4.8}$$

Since the argument of the logarithm has a double pole at the origin, the integration must be displaced slightly from the origin. After some manipulations, we find that the integration gives $32G/\pi$ with the correction we had before in (3.14). For $p=4n+1$ there is a similar result where the two sets of band edges are exchanged.

4.3. Numerator unity and even p

The total bandwidth for $q=1$ with even p can be evaluated in the same manner. Again one set of the band edges is described by (4.3) and (4.4) with $(m-\frac{1}{4})\pi$ replaced by $m\pi$ for $\nu=0$ (in the case $p=4n$; it is quite similar for $p=4n+2$). It is understood now that $x^{++}=x^{--}=0$ is a degenerate solution, so two bands are touching at $E=0$. If on the right-hand side of (4.3) and (4.4) $(m+\frac{1}{2})\pi$ are substituted instead, the other set of band edges, for $\nu=\pi/p$, can be obtained. The reason is that $\nu=\pi/p$ is equivalent to the translational shift of cosine potential from n to $n+\frac{1}{2}$, so we connect two turning points from $\frac{1}{2}$ to $p/2-\frac{1}{2}$ rather than from 0 to $p/2$ as in $\nu=0$. Using the WKB solutions $c_n^{\pm}$, we find that an additional phase $\exp(\mp\mathrm{i}(\pi/2))$ should be added to α and β respectively.

The band edges given by x^{++} and x^{--} can be substituted into the expression for the total bandwidth. Again pW is found to be $32G/\pi$ and the correction is given by (3.14), with a positive coefficient twice as large as is found for odd p. This is in agreement with what numerical results give.

4.4. Numerator 2 and odd p

For $\phi = 2/p$ with p odd, the same procedure is followed as for $1/p$. We start from two independent solutions for n around 0, find their corresponding forms around $p/2$, and impose odd or even boundary conditions at 0 and $p/2$ in order to obtain corresponding eigenvalues. But p should be replaced by $p/2$ in α and β before they can be used again to connect two turning points, namely from 0 to $p/4$ and from $p/4$ to $p/2$, and the connection formula at one turning point $p/4$, equation (3.11), should be included. Here we omit the laborious derivation and give directly the equations determining the eigenvalues for different boundary conditions. These equations are

$$\operatorname{Im}\ln\Gamma(\tfrac{1}{2}-\mathrm{i}x)+x\ln\frac{2p}{\pi}-\tfrac{1}{2}\tan^{-1}\mathrm{e}^{\pi x}=\tfrac{1}{2}(m-\tfrac{1}{4})\pi \qquad \text{for } E^{-+} \tag{4.9}$$

$$\operatorname{Im}\ln\Gamma(\tfrac{1}{2}-\mathrm{i}x)+x\ln\frac{2p}{\pi}-\tfrac{1}{2}\tan^{-1}\frac{\mathrm{e}^{\pi x}}{\sqrt{2+\mathrm{e}^{-2\pi x}}}=(m-\tfrac{1}{8})\pi \tag{4.10a}$$

$$\operatorname{Im}\ln\Gamma(\tfrac{1}{2}-\mathrm{i}x)+x\ln\frac{2p}{\pi}+\tfrac{1}{2}\tan^{-1}\frac{\mathrm{e}^{\pi x}}{\sqrt{2+\mathrm{e}^{-2\pi x}}}=(m+\tfrac{3}{8})\pi \qquad \text{for } E^{++} \tag{4.10b}$$

$$\operatorname{Im}\ln\Gamma(\tfrac{1}{2}-\mathrm{i}x)+x\ln\frac{2p}{\pi}+\tfrac{1}{2}\tan^{-1}\frac{\mathrm{e}^{\pi x}}{\sqrt{2+\mathrm{e}^{-2\pi x}}}=(m-\tfrac{1}{8})\pi \tag{4.11a}$$

$$\operatorname{Im}\ln\Gamma(\tfrac{1}{2}-\mathrm{i}x)+x\ln\frac{2p}{\pi}-\tfrac{1}{2}\tan^{-1}\frac{\mathrm{e}^{\pi x}}{\sqrt{2+\mathrm{e}^{-2\pi x}}}=(m+\tfrac{3}{8})\pi \qquad \text{for } E^{--}. \tag{4.11b}$$

Here $x = pE/8\pi$, m is any integer and $p = 4n+3$. The equations for $p = 4n+1$ can be derived similarly.

For present case, two adjacent edges given by E^{++}, which are getting closer as E goes large, alternate with those by E^{--}, in other words the edges in E^{++} or E^{--} act as band top and bottom alternatively. The best way to deal with it is to classify band edges into four groups, namely given by (4.8a, b) and (4.9a, b), such that in one group, (4.8b) for instance, all edges are band tops for all positive and negative values. For a complete spectrum, the band tops are given by (4.9b^-), (4.8b), (4.8a^-) and (4.9a) while the rest gives the band bottom, where the superscript minus sign stands for the other set of band edges which is given with sign of x changed similar to $q=1$ case. An adaptation of (4.1) to this situation gives the total band width as

$$4\sum_{m(\text{odd})}(E_m^{-+}-E_m^{++})+4\sum_{m(\text{even})}(E_m^{-+}-E_m^{--}). \tag{4.12}$$

For the contour integral of (4.2) we need

$$\phi(z)=\frac{(4.9a)(4.9b)}{(4.10a)(4.11b)} \tag{4.12}$$

where we divide (4.9) as (4.9*a*, *b*) for even and odd m respectively. Equation (4.7) can be written without the parameter m in the form

$$(1-\mathrm{i}\,\mathrm{e}^{\pi x})\sqrt{1+\mathrm{e}^{-2\pi x}} \mp 2\mathrm{e}^{-\mathrm{i}(\pi/4)} \frac{\pi}{[\Gamma(\frac{1}{2}-\mathrm{i}x)]^2} \exp\left(-2\mathrm{i}x \ln \frac{2p}{\pi}\right) = 0 \tag{4.14}$$

while (4.10*a*) and (4.11*b*) have the form

$$\sqrt{2+\mathrm{e}^{-2\pi x}} - \mathrm{i}\,\mathrm{e}^{\pi x} \mp 2\mathrm{e}^{-\mathrm{i}(\pi/4)} \frac{\pi}{[\Gamma(\frac{1}{2}-\mathrm{i}x)]^2} \exp\left(-2\mathrm{i}x \ln \frac{2p}{\pi}\right) = 0. \tag{4.15}$$

To get rid of zeros on the imaginary axis we should multiply (4.15) by $[\Gamma(\frac{1}{4}-\mathrm{i}(z/2))]^2$, and (4.14) by $[\Gamma(\frac{1}{4}-\mathrm{i}(z/2))]^{3/2}[\Gamma(\frac{3}{4}-\mathrm{i}(z/2))]^{1/2}$.

Again (4.2), together with (4.13), gives the rescaled total bandwidth pW as

$$-16\mathrm{i} \oint z \frac{\mathrm{d}}{\mathrm{d}z} \ln \left[\left(\frac{2\mathrm{i}}{z}\right)^{1/2} \frac{\Gamma\left(\frac{3}{4}-\mathrm{i}\frac{z}{2}\right) (1-\mathrm{i}\,\mathrm{e}^{\pi z})^2(1+\mathrm{e}^{-2\pi z}) + \frac{4\mathrm{i}\pi^2}{[\Gamma(\frac{1}{2}-\mathrm{i}z)]^4} \exp\left(-4\mathrm{i}z \ln \frac{2p}{\pi}\right)}{\Gamma\left(\frac{1}{4}-\mathrm{i}\frac{z}{2}\right) (\sqrt{2+\mathrm{e}^{-2\pi z}} - \mathrm{i}\,\mathrm{e}^{\pi z})^2 + \frac{4\mathrm{i}\pi^2}{[\Gamma(\frac{1}{2}-\mathrm{i}z)]^4} \exp\left(-4\mathrm{i}z \ln \frac{2p}{\pi}\right)} \right] dz \tag{4.16}$$

where the contour is $z = R\,\mathrm{e}^{i\theta}$ with R sufficiently large. For the same reason as in the calculation of correction to $q=1$ case, R can not be taken to infinity since after a certain value of R, of the order of $2\,\mathrm{e}p/\pi$, the right order of band edges is destroyed, and the WKB approximation is no longer valid. This gives only an exponentially small correction, and the rescaled total bandwidth is

$$-16\mathrm{i} \oint z \frac{\mathrm{d}}{\mathrm{d}z} \ln \left[\left(\frac{2\mathrm{i}}{z}\right)^{1/2} \frac{\Gamma(\frac{3}{4}-\mathrm{i}(z/2))}{\Gamma(\frac{1}{4}-\mathrm{i}(z/2))} \right] \mathrm{d}z - 16\mathrm{i} \int_C z \frac{\mathrm{d}}{\mathrm{d}z} \ln \left[\frac{(1-\mathrm{i}\,\mathrm{e}^{\pi z})^2(1+\mathrm{e}^{-2\pi z})}{(\sqrt{2+\mathrm{e}^{-2\pi z}} - \mathrm{i}\,\mathrm{e}^{\pi z})^2} \right] \mathrm{d}z \tag{4.17}$$

where contour C is $\pi \leqslant \theta \leqslant 2\pi$. The integral is equivalent to

$$\begin{aligned} -16\mathrm{i} \int_{-\infty}^{\infty} y \frac{\mathrm{d}}{\mathrm{d}y} \ln \left[\frac{(1-\mathrm{i}\,\mathrm{e}^{\pi y})^2(1+\mathrm{e}^{-2\pi y})}{(\sqrt{2+\mathrm{e}^{-2\pi y}} - \mathrm{i}\,\mathrm{e}^{\pi y})^2} \right] \mathrm{d}y &= 32 \int_{-1}^{\infty} \left[\tan^{-1} \mathrm{e}^{\pi y} - \tan^{-1} \frac{\mathrm{e}^{\pi y}}{\sqrt{2+\mathrm{e}^{-2\pi y}}} \right] \mathrm{d}y \\ &\doteq \frac{32}{\pi} \int_0^1 \mathrm{d}y \frac{1}{y} \tan^{-1} y \frac{\sqrt{1+2y^2}\sqrt{2+y^2} - y}{\sqrt{1+2y^2} + y^3\sqrt{2+y^2}} \end{aligned} \tag{4.18}$$

which is $32G/\pi$, with no logarithmic corrections. The limit of the integral could be taken to infinity because the integrand is exponentially decaying, so only an exponentially small correction is ignored.

5. Discussion

In sections 3 and 4.2 we found logarithmic corrections to $|pW - C|$ for fractions with numerator unity, similar to the corrections found in the theory of critical phenomena when there are confluent singularities. In contrast, in section 4.4 we found that for numerator 2 there are no such logarithmic corrections, so that the leading corrections are the perturbative terms of order p^{-2} that we have ignored. The structural difference that leads to this outcome is clear. In the case of numerator unity there is a contour integral that surrounds the positive real axis, and it is the contribution from close to the origin that gives the logarithmic terms. In the case of numerator 2 the contour integral (4.16) surrounds the entire real axis, and there are no contributions from close to the origin. In turn, the reason for this difference is for $q=1$ the ordering of the band edges E^{++} and E^{--} is reversed between the negative real axis and the positive real axis, whereas for $q=2$ there is no such reversal at the origin, but the reversal occurs for large energies where the contribution to the bandwidth is negligible. It is this reversal of the order of the eigenvalues of different symmetries that necessitates a contour integral passing through the origin and giving logarithmic corrections.

This observation allows us to generalize the argument to other fixed values of the numerator q when the denominator p is large. Van Mouche (1989) has proved that the ordering of eigenvalues does not change as a function of V_1/V_2, so we can use perturbative arguments valid for very small values of this ratio to work out the ordering. In this limit the gaps occur at $2V_2\cos(2\pi nq/p)$, and the corresponding band edges are ordered with $E^{++}<E^{--}$ for $0<n<p/4$, and with $E^{++}>E^{--}$ for $p/4<n<p/2$. The reversal therefore occurs when n is close to $p/4$, which corresponds to gaps close to zero energy for odd q, but close to the extremals at $\pm 2V_2$ for even q. This allows two features of the numerical results shown in figure 1 to be understood. There are no logarithmic corrections to pW for small even values of q, because there is no change in the ordering of the band edges close to the origin. For small odd values of q there are logarithmic corrections, but their magnitude decreases as q increases, since only two of the $2q$ separate groups of eigenvalues contributes to the logarithmic terms.

The methods we have used assume $q \ll \sqrt{p}$, since it is only in that case that the regions in which the WKB approximation can be used and those in which the parabolic cylinder functions are a good approximation cover the entire space. To understand other features of figure 1, such as those we drew attention to in section 2, it is necessary to combine Azbel's (1964) continued fraction analysis of the spectrum with the methods we have used here. A plausible argument can be constructed for some of the features of the numerical results, but we have not yet pursued these arguments in detail.

Acknowledgment

This work was supported in part by the US National Science Foundation under Grant No. DMR 89-16052.

References

Aubry S and André G 1980 *Ann. Israel Phys. Soc.* vol 3 (Bristol: Adam Hilger) p 133
Azbel M Ya 1964 *Zh. Eksp. Teor. Fiz.* **46** 929
Bell S C and Stinchcombe R B 1989 *J. Phys. A: Math. Gen.* **22** 717
Harper P G 1955 *Proc. Phys. Soc.* A **68** 874
Hofstadter D R 1976 *Phys. Rev.* B **14** 2239
Sokoloff J B 1985 *Phys. Rep.* **126** 189
Thouless D J 1983 *Phys. Rev.* **28** 4272
—— 1990a *Number Theory and Physics* ed J M Luck, P Moussa and M Waldschmidt (Berlin: Springer) p 170–6
—— 1990b *Commun. Math. Phys.* **127** 187
Van Mouche P 1989 *Commun. Math. Phys.* **122** 23
Whittaker E T and Watson G N 1952 *Course of Modern Analysis* (Cambridge: Cambridge University Press) p 348–9
Wilkinson M 1987 *J. Phys. A: Math. Gen.* **20** 4337