Total bandwidth for Harper equation: correction to renormalization analysis

1995 J. Phys. A: Math. Gen. 28 4163

(http://iopscience.iop.org/0305-4470/28/14/031)

J. Phys. A: Math. Gen. 28 (1995) 4163–4173. Printed in the UK

Total bandwidth for Harper equation: correction to renormalization analysis

Yong Tan

Department of Physics and Applied Physics, John Anderson Building, University of Strathclyde, Glasgow G4 0NG, UK

Received 7 February 1995

Abstract. This paper studies the correction to the effective Hamiltonian of the Harper equation in respect of the total bandwidth. The effective Hamiltonian describes the further splitting from the spectrum of a lower-order fraction when the spectrum of a higher-order fraction is considered. The total bandwidth for a split sub-band is given by the Thouless scaling law modified by the curvature of the effective Hamiltonian at its saddle points. In this paper, it is shown that the leading term of the effective Hamiltonian is exactly solvable. Moreover the correction to the curvature of the effective Hamiltonian is obtained numerically, and the explicit analytical expression is also found. This gives the second sum rule, which is of importance in deriving the Thouless scaling law for a generic fraction.

1. Introduction

In this paper, I consider the spectrum of the rational Harper equation [1]

$$e^{-ik_1}c_{n-1} + 2\cos(2\pi\beta n + k_2)c_n + e^{ik_1}c_{n+1} = Ec_n \tag{1.1}$$

where $\beta = p/q$ is a rational number, and q and p are coprime integers. The choice of the boundary condition, $c_{n+q} = c_n$, reduces the Harper equation to a $q \times q$ finite matrix. Among its many interesting properties, the Harper equation has a very rich spectrum which was first depicted by Hofstadter [2] and now is well known as the Hofstadter butterfly. This equation has attracted extensive studies for many years (see, e.g., [3] and references therein). Apart from its pure mathematical interests, some important applications have been found quite recently in solid state physics, such as quantum Hall effect [4, 5] and flux phase of high T_c superconductivity [6, 7]. As a matter of fact, the Harper equation describes the systems of either Landau levels perturbed by a weak sinusoidal potential or the tight binding electrons subject to a weak magnetic field.

The emphasis of this paper is on the spectral properties, in particular the total bandwidth. It was first observed numerically by Thouless [8] that the total bandwidth of the Harper equation scales like

$$\lim_{q\to\infty} qW(q, p) = 32C/\pi \tag{1.2}$$

where C is Catalan's constant. This scaling appears to be universal as the limiting value is independent of the numerators p. However, the corrections to this scaling law depend strongly on the choice of p. This scaling law was later derived analytically [9–11], but the derivation relies on the fact that β is small.

The generalization to a generic fraction was made by Last and Wilkinson [12]. It is based on the renormalization analysis of the Harper equation [13, 14]. For convenience, I introduce an alternative representation of the Harper equation, explicitly

$$H_0(\hat{p}, \hat{x}) = 2\cos\hat{p} + 2\cos\hat{x} \tag{1.3}$$

with the commutation relation given by

$$[\hat{x}, \hat{p}] = \mathrm{i}2\pi\beta. \tag{1.4}$$

Now consider a sequence of $\beta = p/q$ which approaches the simple fraction $\beta_0 = p_0/q_0$. I denote the difference by $\Delta\beta = \beta - \beta_0$. The spectrum for β takes the shape of the spectrum for β_0 which has q_0 sub-bands, and the fine structure, determined by $\Delta\beta$, can be found by the renormalization analysis. Wilkinson [13] pointed out that the further splitting of the lth sub-band is described by the effective Hamiltonian

$$H_l = \mathcal{E}_l(\hat{p}'/q_0, \hat{x}'/q_0) + \mathrm{O}(\Delta\beta) \tag{1.5}$$

where $\mathcal{E}_l$ is the energy dispersion of the lth sub-band. The commutation relation is given instead by

$$[\hat{x}', \hat{p}'] = \mathrm{i}2\pi\beta_{\mathrm{eff}} \tag{1.6}$$

where the effective fraction is

$$\beta_{\mathrm{eff}} = \frac{q_0\beta - p_0}{\beta(1 - q_0 M_l)/p_0 + M_l} = \frac{p_{\mathrm{eff}}}{q_{\mathrm{eff}}}. \tag{1.7}$$

M_l is the Chern integer which can be obtained by solving the Diophantine equation [4, 5]

$$q_0 t_l + p_0 s_l = l \tag{1.8}$$

with $|s_l| \leqslant q_0/2$; $M_l = t_l - t_{l-1}$ and $N_l = s_l - s_{l-1}$.

Note that if the effective fraction for the lth sub-band is small, it is plausible that the analytical results obtained by Thouless [9] are applicable to the effective Hamiltonian. This observation allowed Last and Wilkinson [12] to obtain an explicit expression for the total bandwidth of the split lth sub-band, which is given by (1.2) with q being replaced by q_{eff} and the constant being modified by the curvature of the dispersion at its saddle points. They were also able to derive a sum rule for the derivative of the characteristic polynomial. The combination of these two results appeared to give the same scaling law as (1.2) for the entire spectrum of β, however, with an error of order $\Delta\beta$. This error term must vanish in order to agree with both numerical [8, 10] and analytical results [11]. Therefore it is necessary to take into account the correction term (of order $\Delta\beta$) in the effective Hamiltonian.

In this paper, I investigate the correction term in the effective Hamiltonian, as appearing in (1.5) to first order in $\Delta\beta$. In particular, I am interested in how the total bandwidth for a split sub-band is affected by the inclusion of this correction term. In section 2, I show that the leading term in (1.5), i.e. the energy dispersion, can be solved exactly in terms of solutions of the Harper Hamiltonian, and consequently the band structure and the total bandwidth can be obtained explicitly. In section 3, the effective Hamiltonian is rederived by using the theory of Bellissard and Rammal [15] with necessary modifications. Part of the correction term is obtained explicitly, and some properties of this correction are discussed. In section 4, the total bandwidth for a cluster of bands resulting from the splitting of a sub-band of β_0 is calculated numerically. The results are used to extract the correction to the curvature of the effective Hamiltonian. Section 5 summarizes the numerical results obtained in section 4 in an analytical expression, which is further manipulated to give a second sum rule.

2. Effect of higher orders in dispersion

In this section, I show that the leading term of the effective Hamiltonian for a sub-band, described explicitly by (1.5), is exactly solvable in terms of solutions of the Harper Hamiltonian with small fractions (i.e. $\beta \to 0$ in (1.3) and (1.4)). As a result, the total bandwidth for a split sub-band of β_0 is obtained explicitly. It is further argued that the quadratic correction, of order $q_{\rm eff}^{-2}$, to the total bandwidth of this split sub-band is absent.

It is well known that the energy dispersion of a sub-band of β_0, denoted by $\mathcal{E}_l(k_1, k_2)$, satisfies the characteristic polynomial [16]:

$$P_{q_0}(E) = 2(\cos q_0 k_1 + \cos q_0 k_2) \tag{2.1}$$

where $P_{q_0}(E)$ is a q_0th-order polynomial independent of wavenumbers k_1 and k_2. It is useful to define the critical energy [7, 11] for each sub-band, E_l, which is given by the solution of (2.1) with its right-hand side being zero. This enables us to expand (2.1) in terms of its right-hand side, and consequently write

$$\mathcal{E}_l(k_1, k_2) = E_l + \sum_{n=1} a_n^{(l)} [2(\cos q_0 k_1 + \cos q_0 k_2)]^n. \tag{2.2}$$

The first two coefficients are given by

$$a_1^{(l)} = \frac{1}{P'_{q_0}(E_l)} \qquad a_2^{(l)} = -\frac{P''_{q_0}(E_l)}{2[P'_{q_0}(E_l)]^3}. \tag{2.3}$$

Therefore the leading term of effective Hamiltonian can be expressed as

$$\begin{aligned}\mathcal{E}_l(\hat{x}'/q_0, \hat{p}'/q_0) &= E_l + a_1^{(l)} \cdot 2(\cos \hat{p}' + \cos \hat{x}') + a_2^{(l)} \cdot [2(\cos \hat{p}' + \cos \hat{x}')]^2 + \cdots \\ &= E_l + a_1^{(l)} \cdot H_0 + a_2^{(l)} \cdot H_0^2 + \cdots \end{aligned} \tag{2.4}$$

where H_0 is the Harper Hamiltonian with effective fraction $\beta_{\rm eff}$ which is small. One can conclude that (2.4) can be solved exactly in terms of eigenvalues of H_0, namely e_j. The eigenvalue of $\mathcal{E}_l(\hat{x}'/q_0, \hat{p}'/q_0)$ is written explicitly as

$$\tilde{e}_j = E_l + a_1^{(l)} \cdot e_j + a_2^{(l)} \cdot e_j^2 + \cdots . \tag{2.5}$$

Because of the quadratic term, the spectrum becomes asymmetric about the critical energy, except the central sub-band for odd q_0, as is shown in the Hofstadter diagram.

Now I consider the total bandwidth for a split sub-band. The main contribution comes from the leading term of the effective Hamiltonian, which is given by

$$\begin{aligned} W_l &= \sum_j |\tilde{e}_j^+ - \tilde{e}_j^-| \\ &= |a_1^{(l)}| \sum_j \left| \left[e_j^+ - e_j^-\right] \left[1 + \frac{a_2^{(l)}}{a_1^{(l)}} \left(e_j^+ + e_j^-\right)\right] \right| + \mathrm{O}(q_{\rm eff}^{-3}) \end{aligned} \tag{2.6}$$

where the superscripts $+$ and $-$ denote the band edges. When the full effective Hamiltonian is considered, there is also a contribution (of order $\Delta\beta/q_{\rm eff}$) from the correction term. Most spectral properties of the Harper Hamiltonian with small fractions, such as ordering of band edges and corrections to the scaling, have been understood analytically due to the work of Thouless and Tan [10, 11]. It can be estimated by scaling arguments [9] that the linear term in e_j gives a contribution of order $q_{\rm eff}^{-1}$ to the total bandwidth, while the quadratic term in (2.5) renders a contribution of order $q_{\rm eff}^{-2}$. Since $|a_2^{(l)}/a_1^{(l)}|$ is much smaller than 1/8, the ordering of band edges does not change because of the inclusion of the second term. In addition, since e_j is symmetric about zero, one has

$$W_l = |a_1^{(l)}| \sum_j |e_j^+ - e_j^-| + \mathrm{O}(q_{\rm eff}^{-3}). \tag{2.7}$$

This indicates that the quadratic term in (2.5) does not affect the total bandwidth for this sub-band. This comes out of the fact that, although the bands on one side of critical energy are stretched while those on the other side are shrunk, they cancel out exactly with each other.

The analogy of the Harper equation immediately gives

$$W_l = \frac{32C}{\pi} f_0^{(l)} \frac{1}{q_{\rm eff}} \tag{2.8}$$

where $f_0^{(l)} = |a_1^{(l)}| = 1/|P'_{q_0}(E_l)|$. The correction to (2.8) is of order $q_{\rm eff}^{-3}$ if $p_{\rm eff}$ is even, while if $p_{\rm eff}$ is odd there are logarithmic corrections [11].

Alternatively, Last and Wilkinson [12] started directly from (1.5), which can be expanded in the vicinities of the saddle points in a form

$$H_l \approx \pm f_0^{(l)} \cdot (\hat{p}'^2 - \hat{x}'^2)/q_0^2 \tag{2.9}$$

where $f_0^{(l)} = (1/2)\partial^2\mathcal{E}_l/\partial k_\mu^2$, and $\mu = 1, 2$. This is because that the behaviour of the Hamiltonian around its saddle points tends to describe the main contribution to the total bandwidth. Furthermore, by use of the identity

$$\partial^2\mathcal{E}_l/\partial k_\mu^2 = \pm 2q_0^2/P'_{q_0}(E_l) \tag{2.10}$$

equation (2.8) can be obtained.

3. Correction to the effective Hamiltonian

In this section, I follow Bellissard and Rammal [15] to investigate the properties of correction term in the effective Hamiltonian. The following calculation is carried out for all the values of wavenumber (k_1, k_2), rather than the specific ones giving the band edges. This allows one to obtain the complete effective Hamiltonian.

It is pointed out by Bellissard and Rammal [15] that the Hamiltonian for a higher fraction $\beta = p/q$ can be effectively represented by that of a lower fraction $\beta_0 = p_0/q_0$. However, further quantization according to the difference between these two fractions must be included. Similarly I write the Hamiltonian as

$$\mathcal{H} = H(k + \gamma^{1/2}K) + \gamma H_1(k + \gamma^{1/2}K) + \mathrm{O}(\gamma^2) \tag{3.1}$$

where $\gamma = 2\pi q_0^{-2}\beta_{\rm eff} = 2\pi\Delta\beta + \mathrm{O}[(\Delta\beta)^2]$. K_1 and K_2 are normalized canonical operators satisfying the commutation $[K_1, K_2] = -\mathrm{i}$. H is given by (1.1), which is represented by a finite $q_0 \times q_0$ matrix. The correction term H_1 remains unknown; however, its second derivative, which is the main interest of this paper, is found in the sections that follow.

To project $\mathcal{H}$ into a sub-band of β_0, similarly I use the Schur formula [17]

$$H_{\rm eff}(E) = P\mathcal{H}P + P\mathcal{H}Q \cdot (E - Q\mathcal{H}Q)^{-1} \cdot Q\mathcal{H}P \tag{3.2}$$

where P and Q are projection operators. Specifically the projection operator for the lth sub-band of H is constructed as

$$P = |l\rangle\langle l| \qquad Q = 1 - P = \sum_{l' \neq l} |l'\rangle\langle l'| \tag{3.3}$$

where $|l\rangle$ satisfies

$$H(k_1, k_2)|l\rangle = \mathcal{E}_l(k_1, k_2)|l\rangle. \tag{3.4}$$

By the choice of projection operators, E in (3.2) is a function of K_1 and K_2 and therefore is an operator generated by $\Delta\beta$. Furthermore, the effective Hamiltonian is given by,

$$H_l = E(K_1, K_2) = \langle l|H_{\rm eff}(E)|l\rangle. \tag{3.5}$$

Equations (3.2) and (3.5) form a pair of self-consistent equations.

3.1. Projection of H

To calculate the effective Hamiltonian in orders of γ, it is necessary to expand

$$H(k + \gamma^{1/2}K) = H(k) + \gamma^{1/2}\sum_{\mu} \partial_\mu H(k) \cdot K_\mu + \frac{\gamma}{2!}\sum_{\mu,\nu} \partial_\mu\partial_\nu H(k) \cdot K_\mu K_\nu + \cdots \tag{3.6}$$

where $\partial_\mu \equiv \partial/\partial k_\mu$ and $\mu = 1, 2$. Then the projection given by (3.2) can be applied iteratively, the method has been described in detail in [15].

The following formula:

$$\langle l'|\partial_\mu H|l\rangle = \partial_\mu\mathcal{E}_l \cdot \delta_{ll'} + (\mathcal{E}_l - \mathcal{E}_{l'}) \cdot \langle l'|\partial_\mu l\rangle = \partial_\mu\mathcal{E}_l \cdot \delta_{ll'} + \langle l'|\mathcal{E}_l - H|\partial_\mu l\rangle \tag{3.7}$$

has been used to deal with the derivatives of H. Here I omit laborious calculations and give the main results, which are greatly simplified compared with those of Bellissard and Rammal. Up to first order in γ, the effective Hamiltonian is given by

$$H_l = \mathcal{E}_l(k_1, k_2) + \gamma^{1/2}\sum_{\mu}(\partial_\mu\mathcal{E}_l) \cdot K_\mu + \frac{\gamma}{2}\sum_{\mu,\nu}(\partial_\mu\partial_\nu\mathcal{E}_l) \cdot K_\mu K_\nu - \gamma\sigma_E^{(0)} \tag{3.8}$$

where

$$\sigma_E^{(0)} = \frac{\mathrm{i}}{2}\Big[\langle\partial_1 l|\mathcal{E}_l - H(k)|\partial_2 l\rangle - \langle\partial_2 l|\mathcal{E}_l - H(k)|\partial_1 l\rangle\Big]. \tag{3.9}$$

Here $\sigma_E^{(0)}$ is a sort of Berry two-form, often called the Wilkinson–Rammal term [15, 18], which has been found previously at the band edges.

It is necessary to obtain the terms of higher orders in γ, as they indicate the proper quantization rules for $\sigma_E^{(0)}$. The terms of order $\gamma^{3/2}$ read

$$\frac{1}{3!}\sum_{\mu,\nu,\rho}(\partial_\mu\partial_\nu\partial_\rho\mathcal{E}_l)\cdot K_\mu K_\nu K_\rho - \sum_\mu(\partial_\mu\sigma_E^{(0)})\cdot K_\mu \tag{3.10}$$

while the terms of order γ^2 are

$$\frac{1}{4!}\sum_{\mu,\nu,\rho,\sigma}(\partial_\mu\partial_\nu\partial_\rho\partial_\sigma\mathcal{E}_l)\cdot K_\mu K_\nu K_\rho K_\sigma - \frac{1}{2!}\sum_{\mu,\nu}(\partial_\mu\partial_\nu\sigma_E^{(0)})\cdot K_\mu K_\nu + \text{constant} \tag{3.11}$$

where the constant term is a function of $\boldsymbol{k}$ and is irrelevant to the following discussions. It is clear that equations (3.8), (3.10) and (3.11) can be combined and consequently written as

$$H_l = \mathcal{E}_l(k_1+\gamma^{1/2}K_1, k_2+\gamma^{1/2}K_2) - \gamma\sigma_E^{(0)}(k_1+\gamma^{1/2}K_1, k_2+\gamma^{1/2}K_2). \tag{3.12}$$

This can be further expressed as

$$H_l = \mathcal{E}_l(\hat{p}/q_0, \hat{x}/q_0) - \gamma\sigma_E^{(0)}(\hat{p}/q_0, \hat{x}/q_0) \tag{3.13}$$

since one can use the identities

$$\hat{p} = q_0(k_1+\gamma^{1/2}K_1) \qquad \hat{x} = q_0(k_2+\gamma^{1/2}K_2). \tag{3.14}$$

Equation (3.13) recovers the result of (1.5) obtained by Wilkinson [13]. In addition there is a correction term. Since the correction term comes from the standard projection of H, it is most likely that it can be expanded in terms of the Harper Hamiltonian.

3.2. *Projection of H_1*

Although the explicit expression of the correction term H_1 in (3.1) is unknown, it is useful to investigate how it behaves under the same projection. Similarly I write

$$H_1(\boldsymbol{k}+\gamma^{1/2}\boldsymbol{K}) = H_1(\boldsymbol{k}) + \gamma^{1/2}\sum_\mu \partial_\mu H_1(\boldsymbol{k})\cdot K_\mu + \frac{\gamma}{2!}\sum_{\mu,\nu}\partial_\mu\partial_\nu H_1(\boldsymbol{k})\cdot K_\mu K_\nu + \cdots . \tag{3.15}$$

The leading-order contribution is given by $\gamma\langle l|H_1(\boldsymbol{k})|l\rangle$, which I define

$$\sigma_E^{(1)}(k_1,k_2) = \langle l|H_1(\boldsymbol{k})|l\rangle. \tag{3.16}$$

Similarly the quantization of this term is determined by higher-order terms, which can be obtained explicitly

$$\gamma^{3/2}\sum_\mu \partial_\mu\sigma_E^{(1)}(k_1,k_2)\cdot K_\mu + \frac{\gamma^2}{2}\sum_{\mu,\nu}\partial_\mu\partial_\nu\sigma_E^{(1)}(k_1,k_2)\cdot K_\mu K_\nu + \gamma^2\cdot\text{constant}. \tag{3.17}$$

Therefore equations (3.16) and (3.17) can be combined to yield an additional correction term to (3.13), namely

$$\gamma\sigma_E^{(1)}(\hat{p}/q_0, \hat{x}/q_0). \tag{3.18}$$

The second derivative of this term at saddle points is given by (5.1). It is plausible that (3.18) cannot be expressed in power series of the Harper Hamiltonian. This is supported by numerical observation. However, one expects that this term has a similar shape of separatrix in the neighbourhood of the saddle points. The correction to the total bandwidth can be estimated by using its curvature at the saddle points.

4. Numerical results

It has been shown in section 2 that the total bandwidth for the split lth sub-band, W_l, is proportional to q_{eff}^{-1}, with the coefficient rescaled by the curvature of the dispersion at the saddle points. In this section, I investigate numerically the effect of the correction terms, given by (3.13) and (3.18), on W_l.

It has been pointed out in section 2 that the leading term of the effective Hamiltonian given by the dispersion holds all the properties of the Harper Hamiltonian, as long as the total bandwidth is concerned. It is useful to recall the results of the Harper equation with small fractions. For a higher-order fraction p/q which approaches the lower-order one p_0/q_0, the correction to the scaling (1.2) contains logarithmic terms if the parity of

$$q_0 p - p_0 q \tag{4.1}$$

is odd; and the leading correction to (1.2) is proportional to q^{-2} if the parity is even [11]. It is apparent that the corrections to W_l from (3.13) and (3.18) are of order $\Delta\beta/q_{\text{eff}}$.

To avoid the logarithmic corrections, I choose a sequence which converges to a simple fraction $\beta_0 = p_0/q_0$ in a manner

$$\beta_N = \frac{p_0}{q_0 + n_0/N} \tag{4.2}$$

where $n_0 = 2, 4, \ldots$ is an even integer, and N is an odd integer which goes to infinity. The difference is

$$\Delta\beta = \beta_N - \beta_0 = -\frac{n_0 p_0}{q_0(q_0 N + n_0)}. \tag{4.3}$$

The effective fraction is given by (1.7), explicitly

$$\beta_{\text{eff}} = -\frac{p_0 n_0}{N + n_0(1 - p_0 N_l)/q_0} \tag{4.4}$$

where $(1 - p_0 N_l)/q_0 = M_l$ is an integer. One should be able to check that neither the total bandwidth for fraction β_N nor the total bandwidth of the split lth sub-band contains logarithmic corrections.

Take $\beta_0 = 1/3$ and $n_0 = 2$ as an example, there are three bands with effective

$$\beta_{\text{eff}} = -\frac{2}{q_{\text{eff}}^{(l)}} \tag{4.5}$$

where $q_{\text{eff}}^{(1)} = q_{\text{eff}}^{(3)} = N$ and $q_{\text{eff}}^{(2)} = N + 2$, respectively.

The idea of this numerical calculation is that, provided that the sequence like β_N is chosen, the correction from the dispersion to the scaling given by (2.8) must be of the order q_{eff}^{-3}. If numerically the correction term is found to be of lower order, it must come from the correction terms (3.13) and (3.18). To include the correction to the curvature, I express the total bandwidth for the split lth sub-band in a form

$$W_l = \frac{32C}{\pi} f_{\text{eff}}^{(l)} \frac{1}{q_{\text{eff}}^{(l)}} + \mathrm{O}\left(\frac{1}{N^3}\right) \tag{4.6}$$

where $f_{\rm eff}^{(l)}$ is the effective curvature. For convenience of later discussion, I adopt a slightly different expression,

$$W_l = \frac{32C}{\pi} f_{\rm eff}^{(l)} \frac{q_0}{q} + \mathrm{O}\left(\frac{1}{N^3}\right). \tag{4.7}$$

The relation between (4.6) and (4.7) will be discussed in section 6. Furthermore, the effective curvature can be written as

$$f_{\rm eff}^{(l)} = \frac{1}{|P'_{q_0}(E_l)|} + \alpha_l \cdot \Delta\beta. \tag{4.8}$$

The coefficient α_l is of principal interest, as it is related to the curvature of the effective Hamiltonian. The numerical value of $f_{\rm eff}^{(l)}$ is found from the total bandwidth, and it is further inserted into (4.8) to find the value of coefficient α_l. The results are presented in table 1, which are independent of the choice of n_0 as expected. Moreover the symmetry shows that $\alpha_{(q_0+1-l)} = \alpha_l$.

Table 1. Numerical results of coefficients α_l.

p_0/q_0	α_1	α_2	α_3	α_4
1/3	0.4534	−0.9069	—	—
2/3	−0.4534	0.9069	—	—
1/4	0.3778	−0.3778	—	—
1/5	0.2000	0.4128	−1.2258	—
2/5	−0.1798	0.6678	−0.9759	—
1/6	0.0898	0.5255	−0.6154	—
1/7	0.0373	0.3408	0.3653	−1.4872
2/7	0.0116	0.2779	−0.9164	1.2538
4/7	0.3569	−0.7901	0.1448	0.5776
1/8	0.0148	0.1770	0.6238	−0.8156
5/8	−0.2775	0.6749	−0.7191	0.3217

5. Correction to curvature: sum rules

As has already been mentioned, the total bandwidth is described by the curvature of the effective Hamiltonian [9, 12]. The curvature is obtained by taking the second derivative of the terms in (3.13) and (3.18). Although the explicit form of $\sigma_E^{(1)}$ is unknown, the numerical results strongly suggest its existence. In fact, it is found that the numerical results presented in table 1 can be reproduced exactly if the curvature of the effective Hamiltonian takes the form

$$\frac{q_0^{-2}}{2}\left[\partial^2\mathcal{E}_l - \gamma\partial^2\sigma_E^0 - \gamma\partial^2\mathcal{E}_l\cdot(\sigma_H - \kappa^{-1}N_l)\right] \tag{5.1}$$

where ∂^2 is either $\partial^2/\partial k_1^2$ or $\partial^2/\partial k_2^2$, and (k_1, k_2) is taken at one of the saddle points, namely at $(0, \kappa/2)$ or $(\kappa/2, 0)$. The last term comes from the contribution of $\gamma\sigma_E^{(1)}$. In the above expression, σ_H is the well known Berry two-form defined by

$$\sigma_H = \mathrm{i}\left[\langle\partial_1 l|\partial_2 l\rangle - \langle\partial_2 l|\partial_1 l\rangle\right]. \tag{5.2}$$

Its average over the unit cell gives the Chern integer [4], namely

$$\kappa^{-1}\int_0^{\kappa}\int_0^{\kappa}\sigma_H(k_1,k_2)\,\mathrm{d}k_1\,\mathrm{d}k_2 = N_l. \tag{5.3}$$

The size of the Brillouin zone is given by $\kappa = 2\pi/q_0$. As mentioned in the introduction, N_l for the Harper equation can also be obtained by (1.8).

One can further use the identity (2.10) to write (5.1) in a form

$$\frac{\mathrm{sign}(\partial^2\mathcal{E}_l)}{P'_{q_0}(E_l)}(1+q_0\Delta\beta N_l) - \Delta\beta\cdot\pi q_0^{-2}\left[(\partial^2\sigma_E^0)+\sigma_H\cdot(\partial^2\mathcal{E}_l)\right]. \tag{5.4}$$

The total bandwidth of the split lth sub-band is then given by

$$W_l = \frac{32C}{\pi}\frac{1}{q_{\mathrm{eff}}^{(l)}}\left\{\frac{1}{|P'_{q_0}(E_l)|}(1+q_0\Delta\beta N_l) - \Delta\beta\cdot\pi q_0^{-2}\mathrm{sign}\left[\partial^2\mathcal{E}_l\right]\cdot\left[(\partial^2\sigma_E^0)+\sigma_H\cdot(\partial^2\mathcal{E}_l)\right]\right\}. \tag{5.5}$$

It can be shown that

$$\frac{1}{q_{\mathrm{eff}}^{(l)}}(1+q_0\Delta\beta N_l) = \frac{q_0}{q}. \tag{5.6}$$

This enables us to identify (5.5) with (4.7). Therefore the coefficient α_l can be expressed as

$$\alpha_l = -\pi q_0^{-2}\mathrm{sign}\left[\partial^2\mathcal{E}_l\right]\cdot\left[(\partial^2\sigma_E^0)+\sigma_H\cdot(\partial^2\mathcal{E}_l)\right]. \tag{5.7}$$

Table 2 presents the results calculated by this formula. It can also be shown that $\alpha_l(1-\beta_0) = -\alpha_l(\beta_0)$. The exact agreement is found between tables 1 and 2, with the difference of order N^{-1} or $\Delta\beta$.

Table 2. Coefficients α_l calculated by equation (5.7).

p_0/q_0	α_1	α_2	α_3	α_4
1/3	0.453 449[a]	−0.906 899[b]	—	—
2/3	−0.453 449	0.906 899	—	—
1/4	0.377 875	−0.377 875	—	—
1/5	0.200 024	0.412 908	−1.225 865	—
2/5	−0.180 087	0.668 099	−0.976 024	—
1/6	0.089 964	0.525 508	−0.615 472	—
1/7	0.037 389	0.340 894	0.365 409	−1.487 384
2/7	0.011 534	0.277 929	−0.916 613	1.254 300
4/7	0.356 729	−0.790 963	0.145 552	0.577 363
1/8	0.014 805	0.177 006	0.623 886	−0.815 697
5/8	−0.277 882	0.674 620	−0.718 393	0.321 654

[a] $\alpha_1 = \sqrt{3}\pi/12$.
[b] $\alpha_2 = -\sqrt{3}\pi/6$.

There are two sum rules associated with above expression. The first was derived explicitly by Wilkinson and Last [12]

$$\sum_{l=1}^{q_0} \frac{1}{|P'_{q_0}(E_l)|} = \frac{1}{q_0}. \tag{5.8}$$

The derivation follows the exact result obtained by Avron, van Mouche and Simon [19] for the intersection spectrum. The application of this sum rule to the total bandwidth leads (4.7) to the scaling law (1.2). Since there is no correction of order q^{-2} for the total bandwidth of the entire spectrum of β, it is expected that

$$\sum_{l=1}^{q_0} \alpha_l = \sum_{l=1}^{q_0} \operatorname{sign}\left[\partial^2 \mathcal{E}_l\right] \cdot \left[(\partial^2 \sigma_E^0) + \sigma_H \cdot (\partial^2 \mathcal{E}_l)\right] = 0. \tag{5.9}$$

This second sum rule has been checked to be valid for all the simple fractions.

Due to the inclusion of the correction terms $\sigma_E^{(0)}$ and $\sigma_E^{(1)}$, the positions of band edges and the critical energy will be shifted by an order of γ. The numerical observation on the shift of the band edges proves that the correction term $\gamma\sigma_E^{(0)}$ is correct [15, 18]. It also shows that $\sigma_E^{(1)}$ must be zero at the band edges. If the effective Hamiltonian is completely Harper-like, i.e. it can be expanded in powers of Harper Hamiltonian, the shift of the critical energy is given by the value of $\gamma(\sigma_E^{(0)} + \sigma_E^{(1)})$ for (k_1, k_2) being $(0, \kappa/2)$, *independent* of the choice of the sequence of β_N in (4.2). But the numerical calculation shows that it depends strongly on β_N (more specifically on n_0). This agrees with the fact that $\sigma_E^{(1)}$ is non-Harper-like, as strongly suggested by the curvature of $\sigma_E^{(1)}$ given by (5.1).

6. Conclusions and discussion

This paper studies the effective Hamiltonian of the Harper equation in the total bandwidth point of view. The renormalization analysis of the Harper equation has shown that the spectrum of a higher-order fraction of β can be obtained by that of a lower-order fraction β_0 with its further splitting described by the effective Hamiltonian. The total bandwidth for each split sub-band of β_0 has been found to be given by the Thouless scaling law modified by the curvature of the effective Hamiltonian. In this paper, the leading term in the effective Hamiltonian which is given by the dispersion is proved to be exactly solvable. Some properties of the correction terms in the effective Hamiltonian are also discussed. The main result of this paper is the correction to the curvature, which is obtained numerically, and its explicit analytical expression is also found. This yields a second sum rule in addition to the one obtained by Wilkinson and Last. The results for the curvature are essential to prove the Thouless scaling law of the total bandwidth for a generic fraction.

Acknowledgments

The author thanks M Wilkinson for suggesting the problem, and T Ziman for useful discussions and encouragement. I am very grateful to Professors J Bellissard and D J Thouless for many helpful discussions, and for their hospitality during my visit in Laboratoire de Physique Quantique in Toulouse and the University of Washington. I acknowledge the hospitality of ISI in Torino where part of the work was carried out. This work was supported by the EPSRC (UK).

References

[1] Harper P G 1955 *Proc. Phys. Soc.* A **68** 879
[2] Hofstadter D R 1976 *Phys. Rev.* B **14** 2239
[3] Sokoloff J B 1985 *Phys. Rep.* **126** 189
[4] Thouless D J, Kohmoto M, Nightingale M P and den Nijs M 1982 *Phys. Rev. Lett.* **49** 405
[5] Středa P 1982 *J. Phys. C: Solid State Phys.* **15** L717
[6] Affleck I and Marston J B 1988 *Phys. Rev.* B **37** 3774
[7] Tan Y and Thouless D J 1992 *Phys. Rev.* B **46** 2985
[8] Thouless D J 1983 *Phys. Rev.* B **28** 4272
[9] Thouless D J 1990 *Commun. Math. Phys.* **127** 187
[10] Thouless D J and Tan Y 1991 *J. Phys. A: Math. Gen.* **24** 4055
[11] Thouless D J and Tan Y 1991 *Physica* **177A** 567
[12] Last Y and Wilkinson M 1992 *J. Phys. A: Math. Gen.* **25** 6123
[13] Wilkinson M 1987 *J. Phys. A: Math. Gen.* **20** 4337
[14] Wilkinson M 1994 *J. Phys. A: Math. Gen.* **27** 8123
[15] Rammal R and Bellissard J 1990 *J. Physique* **51** 1803
[16] Bellissard J and Simon B 1982 *J. Funct. Anal.* **48** 408
[17] Feshbach H 1962 *Ann. Phys., NY* **19** 287
[18] Wilkinson M 1984 *J. Phys. A: Math. Gen.* **17** 3459
[19] Avron J, van Mouche P H M and Simon B 1990 *Commun. Math. Phys.* **132** 103