Quantum group, Bethe ansatz equations, and Bloch wave functions in magnetic fields

Yasuhiro Hatsugai*
Department of Applied Physics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan

Mahito Kohmoto
Institute for Solid State Physics, University of Tokyo, 7-22-1 Roppongi Minato-ku, Tokyo 106, Japan

Yong-Shi Wu
Department of Physics, University of Utah, Salt Lake City, Utah 84112
(Received 14 August 1995)

The wave functions for a two-dimensional Bloch electron in uniform magnetic fields at the mid-band points are studied by exploiting a connection to the quantum group $U_q(sl_2)$: A linear combination of its generators gives the Hamiltonian. We apply both analytic and numerical methods to obtain and analyze the wave functions, by solving the functional Bethe ansatz equations first proposed by Wiegmann and Zabrodin on the basis of the above observation. The semiclassical case with the flux per plaquette $\phi = 1/Q$ is analyzed in detail, by exploring a structure of the Bethe ansatz equations. We also reveal the multifractal structure of the solutions to Bethe ansatz equations and corresponding wave functions when ϕ is irrational, such as the golden or silver mean.

I. INTRODUCTION

The effects of a magnetic field on electrons in the quantum regime are a fascinating problem with very rich physics. The path-dependent geometric phase of the wave function is known to play a fundamental role in the problem.[1] For example, the geometric phase and other geometric considerations, such as gauge invariance in topologically nontrivial systems[2] and the effect of edges,[3] are crucial to the quantum Hall effect in a planar electron system.[4,1] Moreover, the quantized Hall conductance has a topological character, both for a system without boundary (periodic boundary conditions)[5] and for one with boundaries (edges).[6] Thus geometry and topology are of central importance in this condensed-matter problem.

In this paper, we will show that a certain exotic algebraic structure plays an important role as well, if one includes the effects of a periodic potential in this problem. This is not too surprising, since there are two fundamental periods in the problem: One is the original period in the geometric phase, and the other is that of the periodic potential. Interplay of the two intrinsic periods brings more interesting structures and, therefore, richer physics into the problem.

Normally a solid with well localized atomic orbits is modeled by the tight-binding Hamiltonian and the effect of the magnetic field is included by the Peierls substitution. This procedure is usually motivated as an approximation. However, the gauge invariance is preserved by the substitution, i.e., both the original continuum and the lattice systems have the same local $U(1)$ symmetry. Thus in our opinion, the Peierls substitution has captured a fundamental feature of the problem and, therefore, has a meaning more fundamental than merely an approximation.

So in the problem of a planar Bloch electron in a uniform magnetic field, one can use the flux per plaquette, ϕ, to characterize the system. When ϕ is rational, the one-body Schrödinger equation can be reduced[7] to a difference equation on a linear lattice, called Harper's equation, which appears in many different physical contexts ranging from the quantum Hall effect[5,6] to quasiperiodic systems.[8] When ϕ is irrational, that is, the magnetic flux is incommensurate with the periodic potential, the spectrum is known to have an extremely rich structure like the Cantor set and to exhibit a multifractal behavior.[9,8] Another interesting case is the weak field limit. When the flux is small, a semiclassical treatment of the WKB type is justified.[10]

Recently, a connection is found between this physical problem and a mathematically new concept, the so-called quantum group.[11] The essential point is that, at least for some special points in the spectrum, e.g., at the so-called mid-band points, the corresponding Hamiltonian (in certain gauge) can be expressed as a linear combination of generators of a quantum group, called $U_q(sl_2)$, which is a deformation of usual Lie group sl_2. (For a self-contained presentation see below.) On the basis of this observation, Wiegmann and Zabrodin[11] have proposed a set of functional Bethe ansatz equations[12] for corresponding wave functions. This opened the possibility of solving the problem analytically. Indeed, we have turned the dream into reality: We have found explicit analytic solutions to this problem for the first time in the literature.[13] The Bethe ansatz equations associated with $U_q(sl_2)$ also provide us a method to solve the problem numerically. Especially when we study the irrational limit of a well-organized sequence of rational fluxes, the method is convenient and useful in revealing the multifractal behavior, as we briefly reported previously in Ref. 13.

In this way, it becomes clear that, in addition to the underlying geometric and topological structures, the problem of a planar Bloch electron in a magnetic field also possesses an interesting, fundamental algebraic structure hidden in it. Previous works in the literature on possible relevance of quantum groups to physics often deal with a sort of artificial q

deformation of an originally physical problem. However, in our present case, the Bloch electron in magnetic field, the quantum group $U_q(sl_2)$ is relevant to a realistic, physical problem: We are dealing with ordinary electrons, not q-deformed electrons. Therefore, we have shown that quantum groups can have a fundamental meaning for observable phenomena.

A clarification of terminology seems appropriate at this point. Usually Bethe ansatz refers to an ansatz for many-body wave functions in a one-dimensional exactly solvable problem. The pseudomomenta parametrizing the many-body wave function satisfy a set of coupled equations, called the Bethe ansatz equations. In this paper, the term ''Bethe ansatz equations'' is not used in its usual sense, since we are studying a one-body problem. The one-body wave function in question is shown to be related to a polynomial of an auxiliary complex variable, and it is the set of algebraic equations determining the roots of the polynomial that is of a form similar to usual Bethe ansatz equations.

In this paper, we are going to report on more details of our published results and to present a number of solutions and results. We will describe our analytic solutions in detail and in a self-contained manner, and present a great deal of numerical solutions to the Bethe ansatz equations when the analytic ones are not available. In the next section, Sec. II, we review the Bloch electron and magnetic translations. In Secs. III and IV, we summarize necessary stuff about the quantum group $U_q(sl_2)$, its representations, and the functional Bethe ansatz equations following Wiegmann and Zabrodin.[11] Then, in Sec. V, we explain in detail how to analytically solve the Bethe ansatz equations at a particular mid-band point $E=0$. Subsequently, Sec. VI is devoted to an analysis of the explicit wave functions obtained. In particular, a multifractal analysis of the wave functions for irrational flux, such as the golden and silver mean, is presented. Finally, in Sec. VII, we present numerical solutions to the Bethe ansatz equations at mid-band points other than $E=0$. Some technical details are put into several Appendices for the convenience of reader.

II. BLOCH ELECTRON AND MAGNETIC TRANSLATIONS

A tight-binding electron on the square lattice in a perpendiculer magnetic field is described by the Hamiltonian

$$H=T_x+T_y+T_x^\dagger+T_y^\dagger, \tag{2.1}$$

where T_x and T_y are the *covariant translation* operators defined by

$$T_x=\sum_{m,n} c_{m+1,n}^\dagger c_{m,n} e^{i\theta_{m,n}^x}, \tag{2.2}$$

$$T_y=\sum_{m,n} c_{m,n+1}^\dagger c_{m,n} e^{i\theta_{m,n}^y}, \tag{2.3}$$

where $c_{m,n}$ is the annihilation operator for an electron at site (m,n). (This Hamiltonian, is obtained from the freely hopping Hamiltonian with dispersion $E=2\cos k_x+2\cos k_y$ with the Peierls substitution, $\hbar k=p+eA$.) The phase factors $\theta_{m,n}^x$ and $\theta_{m,n}^y$ are related to the flux per plaquette $\phi_{m,n}$ by

$$\mathrm{rot}_{(m,n)}\theta=\Delta_x\theta_{m,n}^y-\Delta_y\theta_{m,n}^x=2\pi\phi_{m,n}, \tag{2.4}$$

where the difference operators Δ_x and Δ_y operate on a lattice function $f_{m,n}$ as

$$\Delta_x f_{m,n}=f_{m+1,n}-f_{m,n}, \tag{2.5}$$

$$\Delta_y f_{m,n}=f_{m,n+1}-f_{m,n}. \tag{2.6}$$

As mentioned in the Introduction, H has a local $U(1)$ gauge symmetry, i.e., it is invariant under

$$c_i\to\Omega_i c_i, \tag{2.7}$$

$$e^{i\theta_{ij}}\to\Omega_i e^{i\theta_{ij}}\Omega_j^{-1}, \tag{2.8}$$

$$|\Omega_j|=1, \quad \forall j=(m,n). \tag{2.9}$$

The covariant translations T_x and T_y do not commute. For a one-particle state $|\Psi_{m,n}\rangle=c_{m,n}^\dagger|0\rangle$, which is localized at site (m,n), their operation satisfies

$$T_yT_x|\Psi_{m,n}\rangle=e^{i2\pi\phi_{m,n}}T_xT_y|\Psi_{m,n}\rangle. \tag{2.10}$$

Now let us assume that the magnetic field is uniform, so that we can choose a gauge in which both $\Delta_x\theta_{m,n}^y$ and $\Delta_y\theta_{m,n}^x$ are, respectively, constant independent of the site (m,n). In this case, there is an important symmetry of the Hamiltonian under the *magnetic translations*:

$$[H,\hat{T}_x]=[H,\hat{T}_y]=0, \tag{2.11}$$

where the magnetic translation operators $\hat{T}_x$ and $\hat{T}_y$ are explicitly given by

$$\hat{T}_x=\sum_{m,n} c_{m+1,n}^\dagger c_{m,n} e^{i\chi_{m,n}^x}, \tag{2.12}$$

$$\hat{T}_y=\sum_{m,n} c_{m,n+1}^\dagger c_{m,n} e^{i\chi_{m,n}^y}, \tag{2.13}$$

with the phases $\chi_{m,n}^{x,y}$ satisfying

$$\Delta_x\chi_{m,n}^x=\Delta_x\theta_{m,n}^x, \tag{2.14}$$

$$\Delta_y\chi_{m,n}^x=\Delta_x\theta_{m,n}^y(=\Delta_y\theta_{m,n}^x+2\pi\phi_{m,n}), \tag{2.15}$$

$$\Delta_x\chi_{m,n}^y=\Delta_y\theta_{m,n}^x(=\Delta_x\theta_{m,n}^y-2\pi\phi_{m,n}), \tag{2.16}$$

$$\Delta_y\chi_{m,n}^y=\Delta_y\theta_{m,n}^y. \tag{2.17}$$

In the parenthesis, we have used Eq. (2.4). We can easily solve these equations to get

$$\chi_{m,n}^x=\theta_{m,n}^x+2\pi n\phi_{m,n}, \tag{2.18}$$

$$\chi_{m,n}^y=\theta_{m,n}^y-2\pi m\phi_{m,n}. \tag{2.19}$$

Note that these phases are gauge dependent.

In the following, we further assume the magnetic field is rational, that is, $\phi_{m,n}=\phi=P/Q$ with mutually prime integers P and Q. We can take a ''diagonal'' gauge[14,11] such that

$$\theta_{m,n}^x=\pi\phi(m+n), \quad \theta_{m,n}^y=-\pi\phi(m+n+1), \tag{2.20}$$

which can be viewed as the Landau gauge in (1,1) direction. Correspondingly, we define the magnetic translations, $\hat{T}_+$ and $\hat{T}_-$, along the (1,1) and $(-1,1)$ directions, respectively, as

$$\hat{T}_+ = \hat{T}_x \hat{T}_y , \tag{2.21}$$

$$\hat{T}_- = \hat{T}_x^{-1} \hat{T}_y . \tag{2.22}$$

They commute with the Hamiltonian, $[H,\hat{T}_+] = [H,\hat{T}_-] = 0$, but do not commute with each other. However, one can explicitly check that, for example, $\hat{T}_+^Q$ and $\hat{T}_-$ commute with each other. Thus, we can take simultaneous eigenstates of H, $\hat{T}_+^Q$, and $\hat{T}_-$, which are specified by the momentum in the magnetic Brillouin zone (through the Bloch theorem) as

$$\hat{T}_+^Q |\Phi(p_+,p_-)\rangle = e^{i2Qp_+} |\Phi(p_+,p_-)\rangle, \tag{2.23}$$

$$\hat{T}_- |\Phi(p_+,p_-)\rangle = e^{ip_-} |\Phi(p_+,p_-)\rangle, \tag{2.24}$$

where $|\Phi\rangle$ is a one-particle state spanned by

$$|\Phi(p_+,p_-)\rangle = \sum_{m,n} \Psi_{m,n}(p_+,p_-) c_{m,n}^\dagger |0\rangle. \tag{2.25}$$

From Eqs. (2.23) and (2.24), we have

$$\Psi_{m,n}(\mathbf{p}) = e^{ip_-(m-n)+ip_+(m+m)} \psi_{m+n}(\mathbf{p}), \tag{2.26}$$

$$\psi_{k+2Q}(\mathbf{p}) = \psi_k(\mathbf{p}), \quad k = 1, \ldots, 2Q. \tag{2.27}$$

Written in terms of $\psi_k(\mathbf{p})$, the Schrödinger equation $H|\Phi(\mathrm{p})\rangle = E(\mathbf{p})|\Phi(\mathrm{p})\rangle$ becomes

$$(e^{i(p_++p_-)}q^{-k} + e^{i(p_+-p_-)}q^{k+1})\psi_{k+1}(\mathbf{p}) + (e^{-i(p_++p_-)}q^{k-1} + e^{-i(p_+-p_-)}q^{-k})\psi_{k-1}(\mathbf{p}) = E(\mathbf{p})\psi_k(\mathbf{p}), \tag{2.28}$$

where q is defined as

$$q = e^{i\pi\phi} = e^{i\pi(P/Q)}. \tag{2.29}$$

In the component form, the wave function is a column vector $\boldsymbol{\psi}(\mathbf{p}) = [\psi_1(\mathbf{p}), \cdots, \psi_{2Q}(\mathbf{p})]^t$, with t standing for transpose. Then Eq. (2.28) can be cast in a matrix form as $H(\mathbf{p})\boldsymbol{\psi}(\mathbf{p}) = E(\mathbf{p})\boldsymbol{\psi}(\mathbf{p})$, with

$$\begin{aligned} H(\mathbf{p}) &= e^{i(p_++p_-)}Y^{-1}X^{-1} + e^{i(p_+-p_-)}X^{-1}Y \\ &\quad + e^{-i(p_++p_-)}XY + e^{-i(p_+-p_-)}Y^{-1}X \\ &= (e^{-i(p_++p_-)}X + e^{i(p_+-p_-)}X^{-1})Y \\ &\quad + Y^{-1}(e^{-i(p_+-p_-)}X + e^{i(p_++p_-)}X^{-1}), \end{aligned} \tag{2.30}$$

where the $2Q \times 2Q$ matrices X and Y are given by

$$X = \begin{bmatrix} 0 & & & & 1 \\ 1 & 0 & & & 0 \\ 0 & 1 & 0 & & 0 \\ & & \ddots & \ddots & \\ 0 & & & 1 & 0 \end{bmatrix}, \quad Y = \mathrm{diag}(q^1, q^2, \ldots, q^{2Q}). \tag{2.31}$$

They satisfy the following relations:

$$X^{2Q} = Y^{2Q} = I_{2Q}, \tag{2.32}$$

$$qXY = YX, \quad X^t = X^{-1}, \quad Y^t = Y, \tag{2.33}$$

where I_{2Q} is the unit matrix. The group generated by X and Y with the above relations is called the Heisenberg-Weyl group.

The solvability of the present problem by the Bethe ansatz equations relies on the fact that, at special points in the magnetic Brillouin zone, the Hamiltonian $H(\mathbf{p})$ has a higher symmetry, described by the quantum group $U_q(sl_2)$.[11] As we will show in the next section, at special lines $p_+ = \pi/2 (\mathrm{mod}\ \pi/Q)$, usually called as the mid-band lines, $H(\mathbf{p})$ can be written as

$$\begin{aligned} H_{\mathrm{MB}}(p_-) &\equiv H(p_+ = \pi/2, p_-) \\ &= i(q-q^{-1})[e^{-ip_-}\rho_C(B) + e^{ip_-}\rho_C(C)]. \end{aligned} \tag{2.34}$$

Later we will see that the energy is independent of the parameter p_-. Here, $\rho_c : U_q(sl_2) \to M_{2Q}$ ($2Q$-dimensional matrices) is a cyclic representation of the generators of the quantum group $U_q(sl_2) = \{A,B,C,D\}$, explicitly given by

$$\rho_c(A) = q^{-(Q-1)/2} X^{-1}, \tag{2.35}$$

$$\rho_c(B) = -(q-q^{-1})^{-1}(X - X^{-1})Y, \tag{2.36}$$

$$\rho_c(C) = -(q-q^{-1})^{-1} Y^{-1}(X - X^{-1}), \tag{2.37}$$

$$\rho_c(D) = q^{(Q-1)/2} X. \tag{2.38}$$

We will elaborate on this in the next section to reveal the quantum group structure of the problem.

III. QUANTUM GROUP AND ITS REPRESENTATIONS

Historically quantum groups arise from mathematical structures appearing in exactly solvable one-dimensional quantum many-body models. However, the simplest case, $U_q(sl_2)$, can be understood as a q deformation (i.e., an appropriate "extension" with a complex parameter q) of the ordinary Lie algebra sl_2. Conversely, sl_2 is recovered from $U_q(sl_2)$ by taking the "classical" limit $q \to 1$.

More precisely, $U_q(sl_2)$ is an algebra generated by four generators $\{A,B,C,D\}$, with multiplication and addition (with complex coefficients), subject to the following defining relations

$$AD = DA = 1, \tag{3.1}$$

$$AB=qBA, \quad BD=qDB, \tag{3.2}$$

$$DC=qCD, \quad CA=qAC, \tag{3.3}$$

$$[B,C]=\frac{A^2-D^2}{q-q^{-1}}. \tag{3.4}$$

As $q\to 1$, these relations become those satisfied by $B\to S_+$, $C\to S_-$, and $A\to 1$, $D\to 1$, and $(A-D)/(q-q^{-1})\to S_3$, where $\{S_+,S_-,S_3\}$ are generators of usual Lie algebra sl_2 satisfying the relations $[S_+,S_-]=2S_3$, $[S_3,S_\pm]=\pm S_\pm$. Thus, Eqs. (3.1–3.4) are a q deformation of sl_2 which, as we will see below, preserves many good properties of the representations of sl_2.

It is well known that sl_2 has a $(2j+1)$-dimensional (spin-j) representation, which is realized by the differential operators,

$$\rho_h(S_3)=z\partial_z-j, \tag{3.5}$$

$$\rho_h(S_+)=z(2j-z\partial_z), \tag{3.6}$$

$$\rho_h(S_-)=\partial_z \tag{3.7}$$

acting on a function $\Psi(z)$. Since $\rho_h(S_+)z^{2j}=\rho_h(S_-)z^0=0$, this finite-dimensional representation has both a highest weight and a lowest weight; thus, it is a representation with $\Psi(z)$, a polynomial of degree $2j$. Similarly, $U_q(sl_2)$ has a highest-weight representation realized by replacing the above differential operators with the corresponding q-difference operators:

$$\rho_h(A)\Psi(z)=q^{-j}\Psi(qz), \tag{3.8}$$

$$\rho_h(B)\Psi(z)=(q-q^{-1})^{-1}z[q^{2j}\Psi(q^{-1}z)-q^{-2j}\Psi(qz)], \tag{3.9}$$

$$\rho_h(C)\Psi(z)=(q-q^{-1})^{-1}z^{-1}[\Psi(q^{-1}z)-\Psi(qz)], \tag{3.10}$$

$$\rho_h(D)\Psi(z)=q^j\Psi(q^{-1}z). \tag{3.11}$$

It has also a highest weight and a lowest weight: $\rho_h(B)z^{2j}=\rho_h(C)z^0=0$. In the standard basis, $\Psi_n(z)=z^n$ $(n=0,1,\ldots,Q-1)$, they operate as $\rho_h(A)z^n=q^{n-j}z^n$, $\rho_h(D)z^n=q^{j-n}z^n$, $\rho_h(B)z^n=[2j-n]_qz^{n+1}$, and $\rho_h(C)z^n=[n]_qz^{n-1}$. Here, $[n]$ is a q integer defined as

$$[n]_q=\frac{q^n-q^{-n}}{q-q^{-1}}, \tag{3.12}$$

which approaches to n in the limit $q\to 1$. Thus, they can be represented as the following matrices:

$$\rho_h(A): \quad \mathrm{diag}(q^{-j},q^{1-j},\ldots,q^{j-1},q^j), \tag{3.13}$$

$$\rho_h(D): \quad \mathrm{diag}(q^j,q^{-1+j},\ldots,q^{-j+1},q^{-j}), \tag{3.14}$$

$$\rho_h(B):\begin{bmatrix} 0 & & & & 0 \\ [2j]_q & 0 & & & \\ 0 & [2j-1]_q & 0 & & \\ & & \ddots & \ddots & \\ 0 & & & [1]_q & 0 \end{bmatrix}, \tag{3.15}$$

$$\rho_h(C):\begin{bmatrix} 0 & [1]_q & & & 0 \\ & 0 & [2]_q & & \\ & & \ddots & \ddots & \\ & & & 0 & [2j]_q \\ 0 & & & & 0 \end{bmatrix}. \tag{3.16}$$

Notice that their dimension is $Q\equiv 2j+1$, and they generate tridiagonal matrices.

Up to this point, $U_q(sl_2)$ is quite analogous to sl_2. However, when q is a root of unity, there is another representation, called a cyclic representation, the existence of which is characteristic to $U_q(sl_2)$. (sl_2 does not have a cyclic representation.) Our case, with $q^{2Q}=1$, is such a case. Here, a $2Q$-dimensional cyclic representation is derived from the highest-weight representation of spin-$j=(Q-1)/2$ as follows. First define a cyclic basis $\{\Psi_n,\ n=1,\ldots,2Q \pmod{2Q}\}$ by

$$\Psi_n\equiv\Psi(q^n), \tag{3.17}$$

where $\Psi(z)$ is the polynomial in the spin-$j=(Q-1)/2$ highest-weight representation. Since $q^{2Q}=1$, the periodic condition $\Psi_{n+2Q}=\Psi_n$ is satisfied, justifying the name "cyclic." Taking $z=q^n$ in Eqs. (3.8)–(3.11), a representation is induced in the cyclic basis, with the operations explicitly given by

$$\rho_c(A)\Psi_n=q^{-(Q-1)/2}\Psi_{n+1}, \tag{3.18}$$

$$\rho_c(B)\Psi_n=-(q-q^{-1})^{-1}(q^{n+1}\Psi_{n+1}-q^{n-1}\Psi_{n-1}), \tag{3.19}$$

$$\rho_c(C)\Psi_n=(q-q^{-1})^{-1}(q^{-n}\Psi_{n+1}-q^{-n}\Psi_{n-1}), \tag{3.20}$$

$$\rho_c(D)\Psi_n=q^{(Q-1)/2}\Psi_{n-1}. \tag{3.21}$$

If we write them in matrix form, we get Eqs. (2.35)–(2.38). We note that $\rho_c(B)\Psi_{2Q}$ returns to a linear combination of Ψ_1 and Ψ_{2Q-1}. So there is no highest-weight state (or vector) Ψ_n that can be annihilated by $\rho_c(B)$: $\rho_c(B)\Psi_n\neq 0$. Similarly, there is no lowest-weight state annihilated by $\rho_c(C)$. This is why this representation is called cyclic.

Thus, when $q^{2Q}=1$, a highest-weight representation of dimension Q is related to a cyclic representation of dimension $2Q$ through Eq. (3.17). This establishes the following results about the spectrum and wave functions of the Bloch electron in a magnetic field at the mid-band momenta. The spectrum is given by the eigenvalues of the following $Q\times Q$ tri-diagonal Hermite matrix:

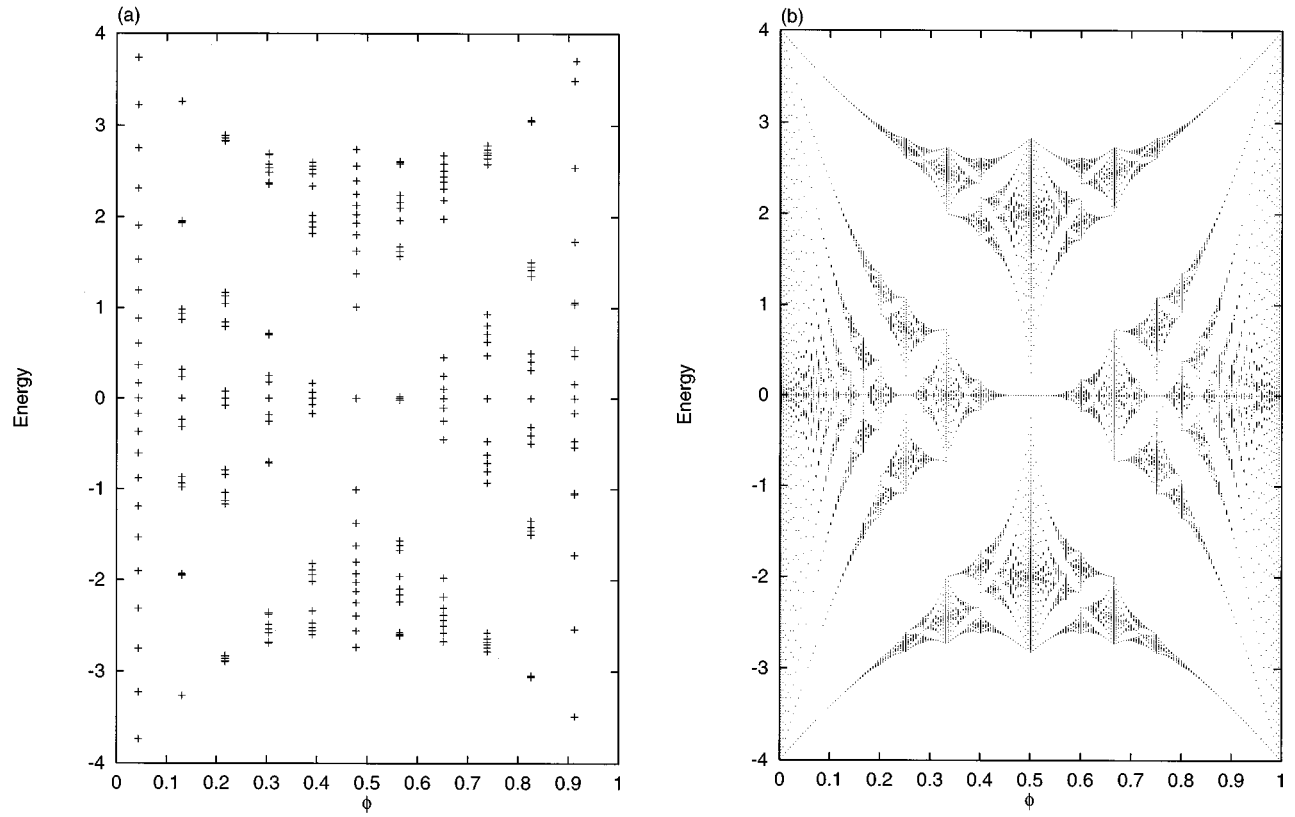

FIG. 1. Mid-band energy spectra obtained from the Bethe ansatz equations for (a) $Q=23$ and $P=1,3,5,\ldots,Q-3,Q-1$ $(+)$ and (b) $Q=401$ and $P=1,3,5,\ldots,Q-3,Q-1$ $(\cdot)$.

$$\frac{H_{\mathrm{MB}}^{\mathrm{tri}}(p_-)}{i(q-q^{-1})}=\begin{bmatrix} 0 & [1]_q e^{-ip_-} & & & \\ [1]_q e^{+ip_-} & 0 & [2]_q e^{-ip_-} & & O \\ & [2]_q e^{+ip_-} & 0 & \ddots & \\ & & & \ddots & [Q-2]_q e^{-ip_-} \\ & O & & [Q-2]_q e^{+ip_-} & 0 \\ & & & & [Q-1]_q e^{+ip_-} \end{bmatrix} \begin{matrix} \\ \\ \\ \\ [Q-1]_q e^{-ip_-} \\ 0 \end{matrix}, \tag{3.22}$$

where we have used the fact $[Q-j]_q=[j]_q$. When Q is large, the eigenvalues of the tri-diagonal matrix gives the ''backbone'' of the so-called Hofstadter's butterfly, i.e., the midpoints of the magnetic Brillouin zone. In Fig. 1, we show eigenvalues of the $H_{\mathrm{MB}}^{\mathrm{tri}}$ as a function of $\phi=-i/\pi\ln q$ for $Q=23$ and $Q=401$. Since the eigenvalues of the tridiagonal matrix does not depend on the phase of the off-diagonal matrix elements, the energy does not depend on p_-. The wave function of the Bloch electron at the mid-band momentum of the jth band is obtained from the jth eigenvector $\{v_m^{(j)}\}$, $(m=0,1,\ldots,Q-1)$, as

$$\Psi_n^{(j)}=\Psi(q^n)=\sum_{m=0}^{Q-1} v_m^{(j)} q^{nm}. \tag{3.23}$$

The dimension of the original Hamiltonian H_{MB} is $2Q$ and that of the tridiagonal functional Hamiltonian $H_{\mathrm{MB}}^{\mathrm{tri}}$ is Q. Actually, H_{MB} has doubly degenerate eigenvalues. The above state (3.23) is degenerate with another state

$$\Psi'^{(j)}_n=\Psi(q^{-n})=\sum_{m=0}^{Q-1} v_m^{(j)} q^{-nm}. \tag{3.24}$$

These two states have the same energy and they are orthogonal to each other, since

$$\sum_{n=0}^{2Q-1} \Psi_n^{(j)}\Psi'^{(j)*}_n=\sum_{0\le m,m'\le Q-1} v_m v^*_{m'} \frac{1-q^{2Q(m+m')}}{1-q^{(m+m')}}=0\,. \tag{3.25}$$

In this way, we obtain all solutions of the original problem.

IV. FUNCTIONAL BETHE ANSATZ EQUATIONS

In this section, we derive the functional Bethe ansatz equations following Wiegmann and Zabrodin.[11] We denote

p_- by p. The Schrödinger equation at a mid-band point is

$$H_{\mathrm{MB}}(p)\Psi_n = i(q-q^{-1})[e^{-ip}\rho_c(B)+e^{ip}\rho_c(C)]\Psi_n = E\Psi_n . \tag{4.1}$$

Recall that the cyclic representation is derivable from a highest weight representation. Changing ρ_c to ρ_h, and using Eqs. (3.9) and (3.10), we can write down the equivalent functional equation,

$$i(e^{-ip}qz+e^{ip}z^{-1})\Psi(qz)-i(q^{-1}e^{-ip}z+e^{ip}z^{-1})\Psi(q^{-1}z) = E\Psi(z). \tag{4.2}$$

Once we have a solution to this functional equation, we can recover all the solutions of Eq. (4.1) using Eq. (3.23) and Eq. (3.24).

Central to what follows is the nontrivial property that we learn from the above long discussion of the representation theory of $U_q(sl_2)$, namely, being a highest weight representation, $\Psi(z)$ is *a polynomial of finite degree* of the auxiliary variable z. Thus, it can be factorized:

$$\Psi(z)=\prod_{m=1}^{Q-1}[z-z_m(p)]. \tag{4.3}$$

First, let us absorb the p dependence of Eq. (4.2) by the gauge transformation,

$$\bar{z}=ze^{-ip}, \tag{4.4}$$

$$\bar{z}_m=z_m(p)e^{-ip}, \tag{4.5}$$

$$\overline{\Psi}(\bar{z})=\prod_{m=1}^{Q-1}[z-z_m(p)]=e^{i(Q-1)p}\prod_{m=1}^{Q-1}(\bar{z}-\bar{z}_m). \tag{4.6}$$

Suppressing the bar over z, we have

$$i(qz+z^{-1})\overline{\Psi}(qz)-i(q^{-1}z+z^{-1})\overline{\Psi}(q^{-1}z)=E\overline{\Psi}(z). \tag{4.7}$$

From Eq. (4.7) and Eq. (4.3), we get

$$i(qz+z^{-1})\prod_{m=1}^{Q-1}\frac{(qz-z_m)}{(z-z_m)}-i(q^{-1}z+z^{-1})\prod_{m=1}^{Q-1}\frac{(q^{-1}z-z_m)}{(z-z_m)} = E. \tag{4.8}$$

Comparing both sides and using the pole-free condition, one obtains the Bethe ansatz equations,

$$\frac{z_l^2+q}{qz_l^2+1}=-\prod_{m=1}^{Q-1}\frac{qz_l-z_m}{z_l-qz_m} \quad (l=1,\ldots,Q-1), \tag{4.9}$$

$$E=-i(q-q^{-1})\sum_{m=1}^{Q-1}z_m . \tag{4.10}$$

From these equations, one can easily see that the energy E is independent of p_-. Note that the functional solution, which corresponds to Ψ_n', is $\Psi(z^{-1})$; it is not a polynomial and has poles.

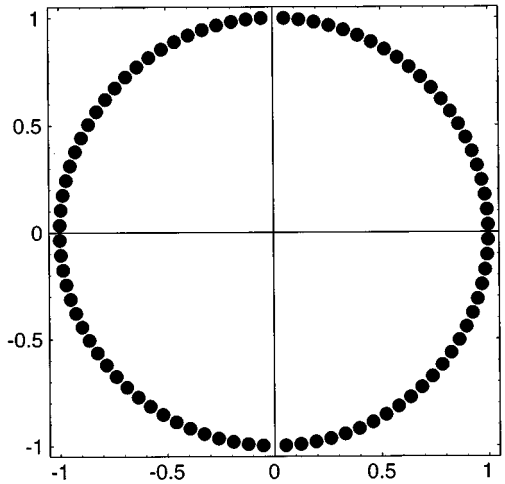

FIG. 2. The roots of the Bethe ansatz equations in the complex plane for $E=0$, in the case with $P=1$ and $Q=89$.

V. DISTRIBUTION OF ROOTS FOR $E=0$

Let us consider first Eq. (4.8). For the zero energy $E=0$, Wiegmann and Zabrodin showed that $\Psi(z)$ is given by the so-called continuous q-ultraspherical polynomial[15] as $\Psi(z)=(q^2;q^2)_n/(q;q^2)_n(-iz)^nP_n(-iz)$, where $P_n(z)=\Sigma_{k=0}^{n}(q;q^2)_k(q;q^2)_{n-k}/(q^2;q^2)_k(q^2;q^2)_{n-k}z^{n-2k}$, $n=(Q-1)/2$, and $(a;q)_k=\Pi_{m=0}^{k-1}(1-aq^m)$.[16] The q-ultraspherical polynomials have several interesting properties. However, we treat the function $\Psi(z)$ directly.

In order to understand the properties of the wave function at $E=0$, the center of the spectrum, we solve the Bethe ansatz equation (4.9) explicitly. For $E=0$, Eq. (4.8) is written as

$$q(z^2+q^{-1})\Psi(qz)=q^{-1}(z^2+q)\Psi(q^{-1}z). \tag{5.1}$$

First put $z=iq^{+1/2}$, then one obtains $q(-q+q^{-1})\Psi(iq^{3/2})=0$ and $\Psi(iq^{3/2})=0$. So $iq^{3/2}$ is one root. Next put $z=iq^{+5/2}$, then one obtains $q(-q^5+q^{-1})\Psi(iq^{7/2})\propto\Psi(iq^{3/2})=0$. Thus, $iq^{7/2}$ is also a root. We can repeat this procedure and get a series of roots: $iq^{2m-1/2}$ with $m=1,\ldots,(Q-1)/2$. The restriction of m arises due to the fact that the prefactor of Ψ vanishes at the last step.

Similarly, setting $z=iq^{-1/2}$, one obtains another sequence of roots: $iq^{-2m+1/2}$, with $m=1,\ldots,(Q-1)/2$. The number of independent roots should equal the degree of the polynomial, $Q-1$; so we have obtained a complete set of roots as

$$z_m=\begin{cases} iq^{2m-1/2}=ie^{i2\pi\phi(m-1/4)} \\ iq^{-2m+1/2}=ie^{i2\pi\phi(m-1/4)} \end{cases} \quad [m=1,\ldots,(Q-1)/2]. \tag{5.2}$$

They are all on the unit circle. Let us write them as $z_m=e^{i\theta_m}$, and consider the distribution of θ_m by $\rho(\theta)=\lim_{Q\to\infty}Q\Delta\theta$, where $\Delta\theta$ is the difference between adjacent θ_m. [Note that the gauge transformation (4.5) only leads to a simple shift of all θ_m by a constant p.]

The restriction on m $[1\leqslant m\leqslant(Q-1)/2]$ gives rise to a nontrivial distribution of the roots. The phase factor θ_m looks

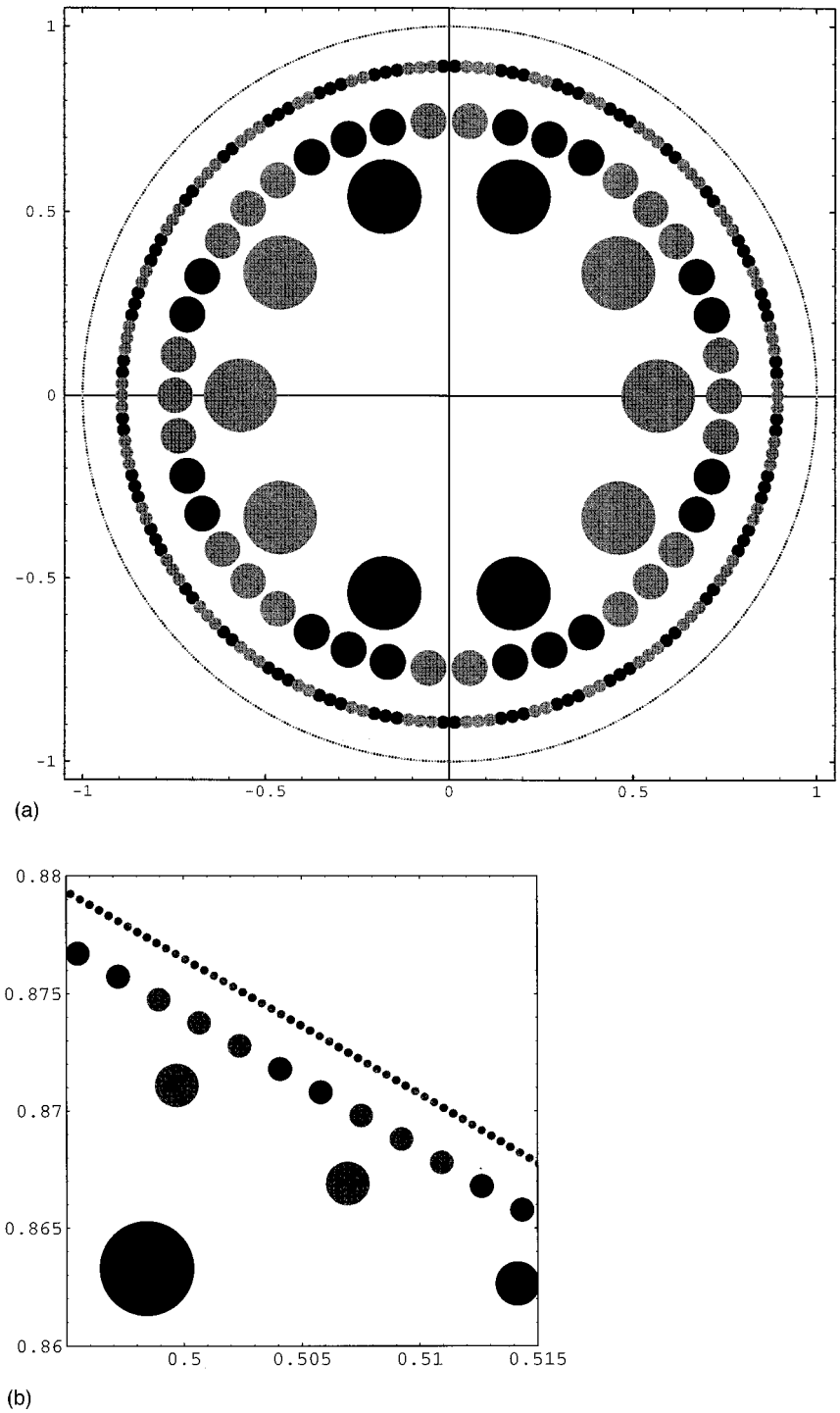

FIG. 3. The roots and pseudoroots of the Bethe ansatz equations for $E=0$ with the rational fluxes, which converge to $1/\tau$: $\phi_k=P_k/Q_k=3/5$, 13/21, 55/89, 233/377, 987/1597, 4181/6765 (a) in the whole complex plane and (b) in an enlarged figure. In each case, the roots and the pseudoroots are always on the unit circle. The radii for ϕ_k have been scaled to show the branching rule.

like a pseudomomentum in the Bethe ansatz solvable models in which the system size is Q. In the following, we consider the limit $Q\to\infty$, which corresponds to the thermodynamic limit in exactly solvable models.

Consider first the case $P=1$, i.e., $\phi=1/Q$. The roots $\{z_m\}$ are distributed uniformly on the unit circle, except near $z=\pm i$. The roots for $Q=89$ are shown in Fig. 2. In the semiclassical limit $Q\to\infty$, that is, $q\to 1$, the distribution function $\rho(\theta)=\lim_{Q\to\infty}Q\Delta\theta$ is smooth (actually constant) everywhere, except at $z=\pm i$. A continuous behavior of $\rho(\theta)$ is usually obtained in the Bethe ansatz solvable models in which $\rho(\theta)$ is determined by an integral equation.

When the flux is irrational, the situation is quite different. First, let us take the flux $\phi=1/\tau=(\sqrt{5}-1)/2$, where τ is the golden mean $(\sqrt{5}+1)/2$. To reach this flux, we consider a sequence of rational fluxes $\phi_k=P_k/Q_k$, where $Q_k=F_{3k+1}$, $P_k=F_{3k}$ and F_k is a Fibonacci number defined by $F_{k+1}=F_k+F_{k-1}$, $F_1=1$, $F_0=1$. In this sequence, P_k and Q_k are all odd. The two types of roots in Eq. (5.2) are nested and $\Delta\theta$ has a complex distribution. To gain an insight into the distribution of roots, it is helpful to consider the pseudoroots, which are defined also by Eq. (5.2), but with the range of m modified to $m=-(Q-1)/2,\ldots,-1,0$. In Fig. 3 we show the distributions of the roots (black points)

and pseudoroots (gray points) for several ϕ_k ($k=1,2,3,4,5,6$). Here, the radius of the unit circle has been scaled so as to show all the cases at once. These figures clearly show that there is a branching rule for the true roots (denoted by A) and the pseudoroots (by B) as follows:

$$A^3 \to A^3B^2A^3B^2A^3,$$

$$A^2 \to A^3B^2A^3,$$

$$B^3 \to B^3A^2B^3A^2B^3,$$

$$B^2 \to B^3A^2B^3. \tag{5.3}$$

The initial condition is $B^3A^2B^3A^2$ (cyclic). At the kth stage, the number of clusters of the true roots A^3 and A^2 is $Q_{k-1}-1$ and $P_{k-1}+1$, respectively. This branching rule gives rise to a self-similar structure for the distribution $\rho(\theta)$ in the limit $k\to\infty$. To characterize the distribution, let us define the generation of the roots. According to the branching rule (5.3), each true (pseudo-) root branches into a cluster of three new true (pseudo-) roots, each of which in a sense has a parent. At the same time, between these clusters of new pseudo- (true) roots, there is a pair of new-born true (pseudo-) roots, which have no parent. We assign the generation number to a root, so that it is 1 when the root does not have a parent, otherwise, it is one plus the generation number of its parent. Let us denote the number of true (pseudo-) roots at the kth stage with generation g by $n_A(g,k)(n_B(g,k))(g=1,\ldots,k)$. [Then in the special case $P=1$, $n_A(g,k)=n_B(g,k)=\delta_{kg}$.] In the present case, there is a recursion formula from the branching rule: i.e., $n_A(g,k)=3n_A(g-1,k-1)$, with $g=2,\ldots,k, n_A(1,k)=2(P_{k-1}+1)$. Thus, $n_A(g,k)=2\times 3^{g-1}(P_{k-g}+1)$ and $n_B(g,k)=2\times 3^{g-1}(P_{k-g}-1)$.

Next, let us consider the flux $\phi=1/\sigma=\sqrt{2}-1$, where σ is the so-called silver mean $\sqrt{2}+1$. To realize this flux in the large Q limit, we consider a sequence of rational fluxes $\phi_k=P_k/Q_k$, where $Q_k=G_{k+1}$, $P_k=G_k$, and G_k is defined by $G_{k+1}=2G_k+G_{k-1}$, $G_2=1$, $G_1=1$. In this case, we can apply our technique of Bethe ansatz equations to each step of the sequence, since all Q_k and P_k are odd. In Fig. 4, we show the distributions of the roots (black points) and pseudoroots (gray points) for several $\phi_k(k=1,2,3,4,5,6)$. The radius of the unit circle is also scaled as before. One can observe a clear branching rule. It is a little different from the golden-mean case and is given by

$$A \to A^3 \to BA^2BA^2B,$$

$$A^2 \to A^4 \to BA^2BA^2BA^2B,$$

$$B \to B^3 \to AB^2AB^2A,$$

$$B^2 \to B^4 \to AB^2AB^2AB^2A. \tag{5.4}$$

The initial condition is AB^2AB^2 (cyclic). From this branching rule, we have a self-similar behavior at every two stages. At the odd stage, all the roots and the pseudoroots appear as A, A^2, B and B^2 and at the even stage, they appear as A^3, A^4, B^3, and B^4. For both cases, the distribution function of the roots has a self-similarity and reflects the self-similar structure of the original problem.

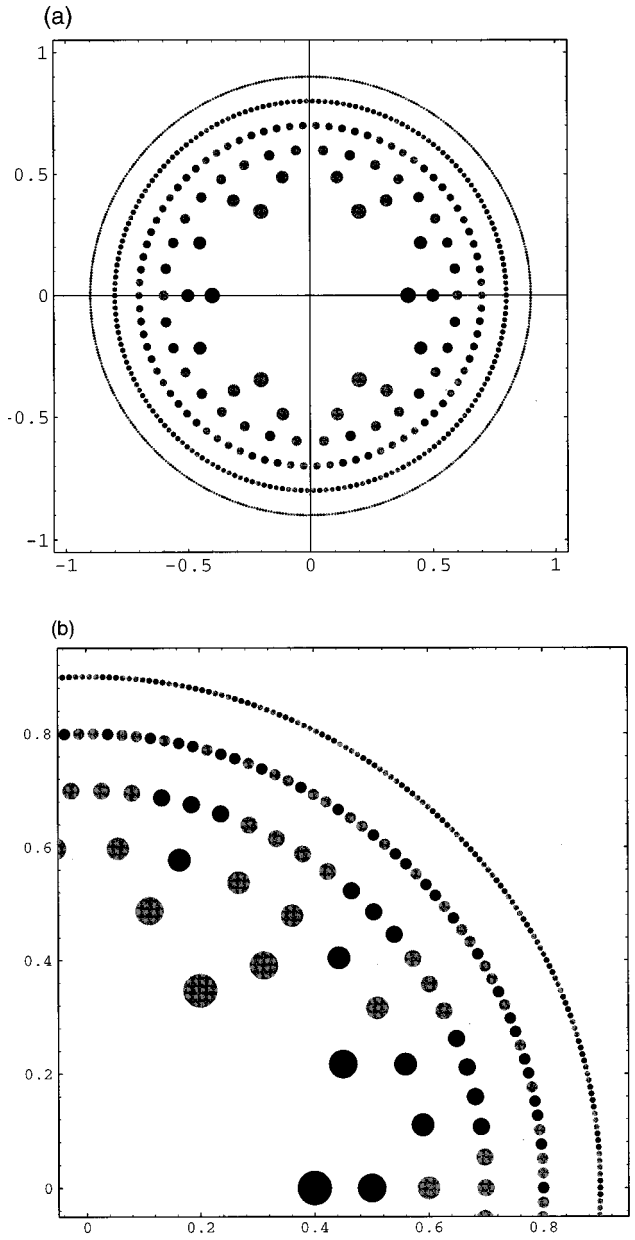

FIG. 4. The roots of the Bethe ansatz equations for $E=0$ with the rational fluxes, which converge to $1/\sigma$, $\phi_k=P_k/Q_k=3/7$, 7/17, 17/41, 41/99, 99/239, 239/577, and 577/1393 (a) in the whole complex plane and (b) in an enlarged figure. The radii are scaled.

The above considerations clearly exemplify the difference in the root distribution between the semiclassical limit and the incommensurate cases. While the root distribution for the semiclassical limit is smooth, it has a self-similar structure and is nowhere differentiable in the incommensurate limit. We believe these features are characteristic to the incommensurate case.

Another way to characterize the distribution is to map it to the dual (reciprocal) space. This can be done for arbitrary P and Q analytically. We lift θ_m ($z_m=e^{i\theta_m}$, z_m's are the roots) to the real axis periodically. On the real axis, the true

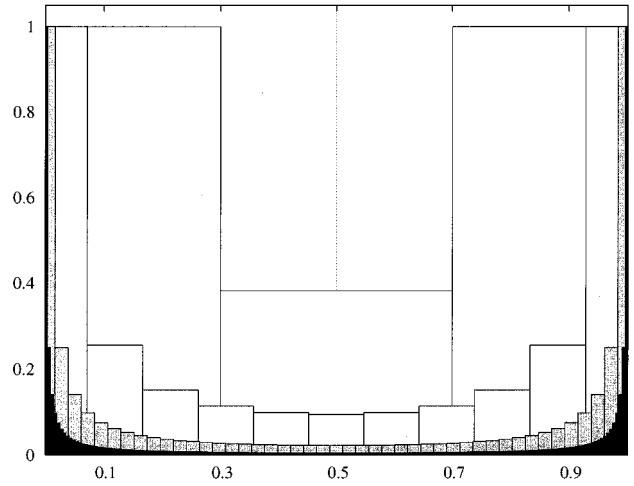

FIG. 5. Squared amplitudes of the wave functions at $E=0$ for $P=1$ and $Q=5,21,89,377$. The wave functions are normalized by their peak heights.

and pseudoroots occupy lattice points $\{j/2Q|j:\text{integer}\}$ with spacing $1/2Q$. Thus, we define the Fourier transform of the so-called defining function by

$$S_Q(k)=\sum_{j=-\infty}^{\infty} e^{ikj}\tilde{S}_Q(j), \tag{5.5}$$

where $\tilde{S}_Q(j)=1$, if there is a true root at $j/2Q$, and $\tilde{S}_Q(j)=0$ otherwise. As shown in Appendix A,

$$S_Q(k)=\frac{\pi}{Q}\sum_{r=0}^{Q-1} s_Q^r\delta(k-k_r) \quad (0\leqslant k<2\pi), \tag{5.6}$$

where

$$s_Q^r=\begin{cases} (Q-1), & r=0, \\ \dfrac{(-)^{r+1}}{\cos(Pk_r/2)}, & r=1,\ldots Q-1; \end{cases} \tag{5.7}$$

and $k_r=2\pi(r/Q)$, which converges to usual momentum continuum $k\in[0,2\pi]$ in the $Q\to\infty$ limit. In this limit, we have

$$|S_Q(k)|^2=\frac{1}{4\cos(Pk/2)} \quad (0<k<2\pi). \tag{5.8}$$

In the semiclassical limit $P=1$ and $Q\to\infty$, $|S_Q(k)|^2$ is well defined and behaves smoothly. In the incommensurate limit, $|S_Q(k)|^2$ is a singular function and is not even differentiable since P tends to infinity. Also, it can be shown that the original defining function is given by

$$\tilde{S}_Q(j)=\frac{1}{2Q}\sum_{n=1}^{Q-1}\left[1-(-)^n\frac{\cos(2\pi jn/Q)}{\cos(\pi Pn/Q)}\right]. \tag{5.9}$$

VI. EXPLICIT WAVE FUNCTIONS AT $E=0$

Now let us consider the wave functions obtained from the above explicit solutions of the Bethe ansatz equations. The wave function at site j is given by $\Psi_j=\Psi(q^j)$ and it is

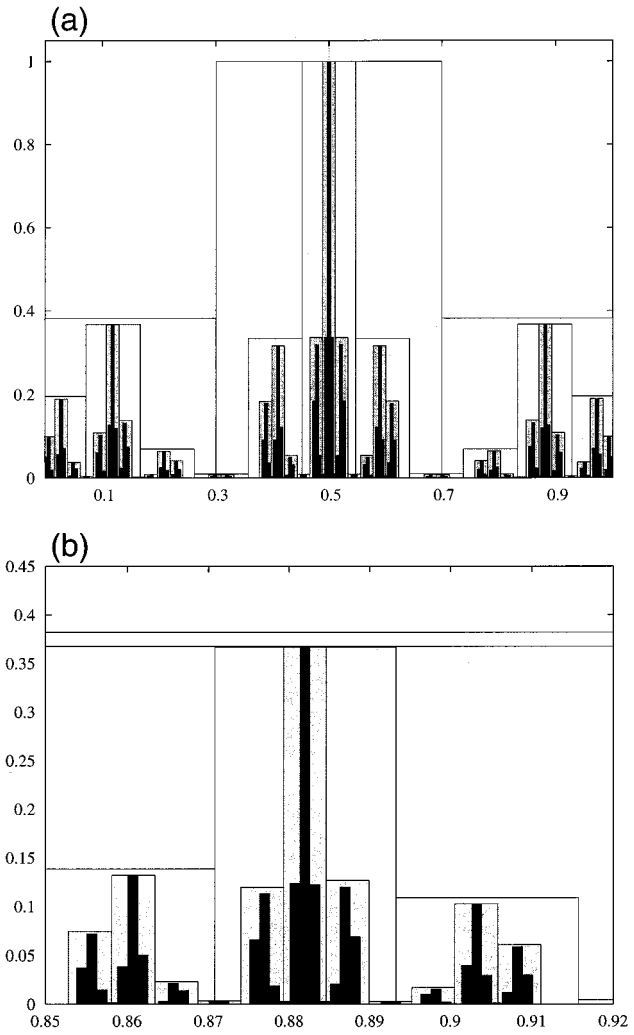

FIG. 6. Squared amplitudes of the wave functions at $E=0$ for the ratios of successive Fibonacci numbers: (a) in the whole region ($\phi=3/5$, 13/21, 55/89) and (b) in an enlarged plot in the region near 0.89 ($\phi=P_k/Q_k=3/5$, 13/21, 55/89, 233/377, 987/1597, 4181/6765). The wave function with a larger value of Q is shaded darker. The wave functions are normalized by their peak heights.

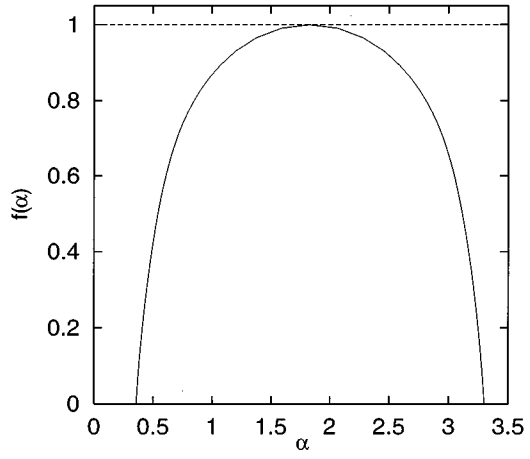

FIG. 7. The function $f(\alpha)$ for the wave function at $E=0$ for the golden-mean flux.

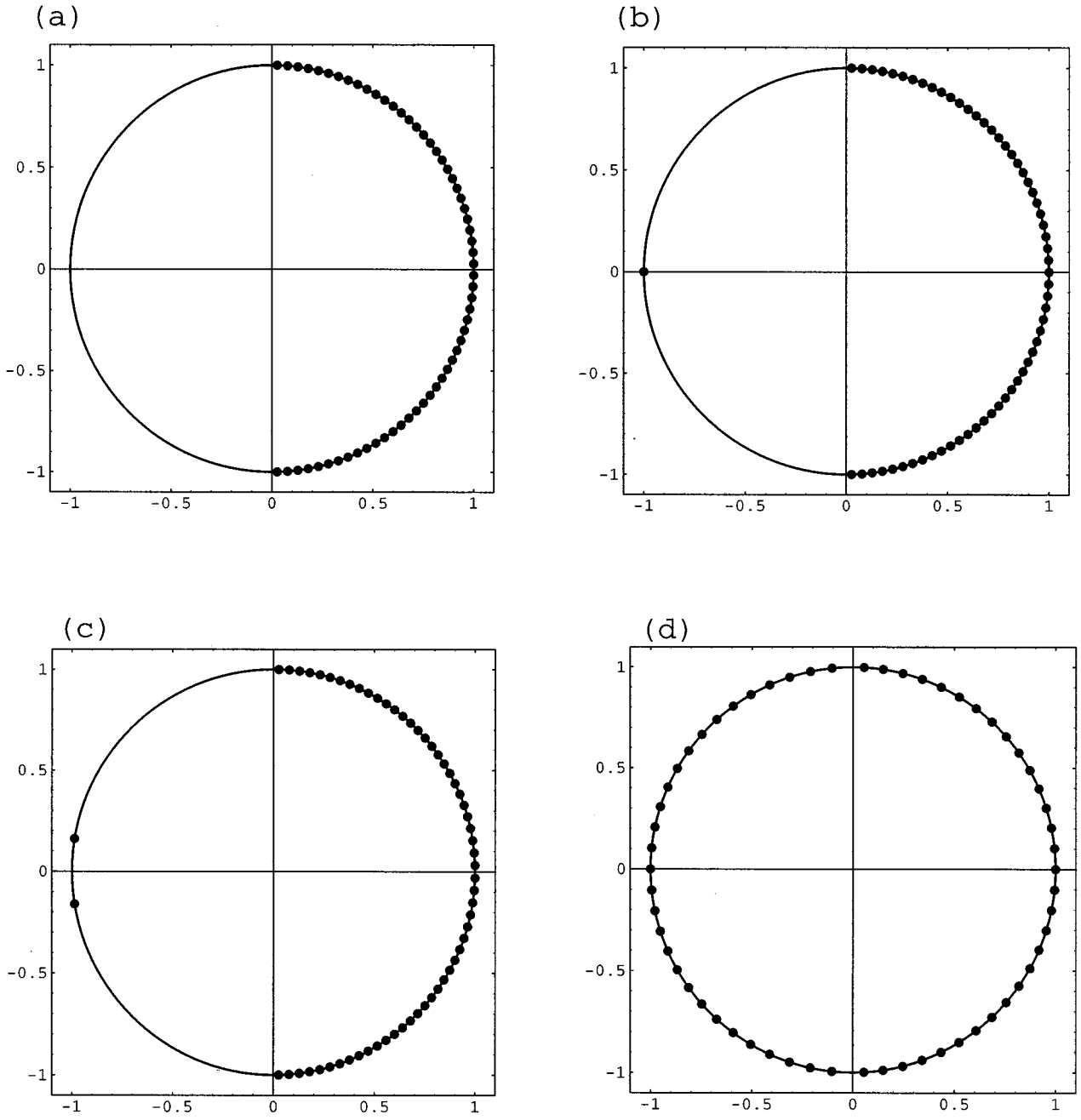

FIG. 8. The roots of the Bethe ansatz equations with $P=1$ and $Q=61$ in the complex plane for mid-band energies ($E\neq 0$) in (a) the highest band; (b) the second highest; (c) the third highest; and (d) in the 30th band, the one just above the central band, the mid-band energy of which is $E=0$.

written in a compact factorized form as

$$\Psi_j=\prod_{m=1}^{(Q-1)/2}(q^j-iq^{2m-1/2})(q^j-iq^{-2m+1/2})$$
$$=(-q)^{-j}(iq^{-j+3/2};q^2)_{(Q-1)/2}(iq^{-j-3/2};q^{-2})_{(Q-1)/2}, \tag{6.1}$$

$$=\sum_{\nu=0}^{(Q-1)/2}\sum_{\nu'=0}^{(Q-1)/2}\begin{bmatrix}(Q-1)/2\\ \nu\end{bmatrix}_q\begin{bmatrix}(Q-1)/2\\ \nu'\end{bmatrix}_q$$
$$\times(-iq^{-j})^{\nu+\nu'}q^{Q(\nu-\nu')/2}, \tag{6.2}$$

where we have used the q-binomial theorem

$$\prod_{n=1}^{N-1}(1+q^{N-1-2n}z)=\sum_{\nu=0}^{N}\begin{bmatrix}N\\ \nu\end{bmatrix}_q z^\nu, \tag{6.3}$$

$$\begin{bmatrix}\mu\\ \nu\end{bmatrix}_q=\frac{[\mu]_q!}{[\nu]_q![\mu-\nu]_q!}. \tag{6.4}$$

It is convenient to shift the site, $j=\bar{j}+J_1$, by the amount J_1, which is determined by $PJ_1=(Q-P)/2(\mathrm{mod}\,2Q)$. Then, $\Psi_{\bar{j}}=0$ at $\bar{j}=2m$, $-2m+1$ $[m=1,\ldots,(Q-1)/2]$, and

the wave function is nonzero only at $\bar{j}=1,3,\ldots,Q$ and $Q+1,Q+3,\ldots 2Q$, as shown in Appendix B. Thus, one has

$$|\Psi_{\bar{1}}|^2=Q, \tag{6.5}$$

$$|\Psi_{\bar{j}}|^2=Q\frac{[\bar{j}-2]_q!!}{[\bar{j}-1]_q!!}, \tag{6.6}$$

for $\bar{j}=3,5,\ldots,Q$ and $|\Psi_{\bar{j}+Q}|^2=|\Psi_{\bar{j}}|^2$, where $[j]_q!!=[j]_q[j-2]_q\cdots[2]_q$ (or $[1]_q$).

When $P=1$, we can proceed further analytically. As shown in Appendix B, we get a compact form for the wave function in the semiclassical limit $P=1$ and $Q\to\infty$,

$$|\psi(x)|^2=\frac{2}{\sin(\pi x)} \quad (0<x<1), \tag{6.7}$$

up to a constant factor, where $\psi(x)=\Psi_{\bar{j}}$, $x=(2j-1)/2Q$. The squared amplitude of the wave function is given by the inverse chord distance. The recursion relation $|\Psi_{\bar{j}}|^2=|\Psi_{\bar{j}-2}|^2\sin^2[\pi(j-2)/Q]/\sin^2[\pi(j-1)/Q]$ obtained from Eq. (6.6) has played a key role in the derivation.[17]

We can also calculate the finite-size correction near the edge $x\approx 0$ or $x\approx 1$ for $Q\to\infty$. As shown in Appendix C, it is given by

$$|\psi(x_{2l+1})|^2=C(l)\frac{2}{\sin\pi x_{2l+1}}, \tag{6.8}$$

$$C(l)=\pi(l+\tfrac{1}{4})\prod_{k=1}^{l}\left(1-\frac{1}{2k}\right)^2=\frac{\pi}{2}\frac{1+1/4l}{1+1/2l}\prod_{k=1}^{l}\left(1-\frac{1}{4k^2}\right), \tag{6.9}$$

where $C(0)=\pi/4=0.785\,39\ldots$, $C(1)=5\pi/16=0.981\,74\ldots$, $C(2)=81\pi/256=0.994\,02\ldots$, ..., since $\sin z=z\Pi_{k=1}^{\infty}(1-z^2/k^2\pi^2)$, $\Pi_{k=1}^{\infty}(1-1/4k^2)=2/\pi$. So the finite-size correction factor $C(l)$ converges to unity very rapidly and Eq. (6.7) is quite accurate even for small Q. The norm of the wave function is $\ln Q+$ const and unnormalizable, characteristic to a critical wave function. In Fig. 5, the amplitudes of the analytic wave functions, normalized by the peak heights, are shown for several values of Q.

Next let us discuss the case with golden-mean flux. We plot the analytic results (6.6) in Fig. 6 for a sequence of rational fluxes converging to $1/\tau$. One can easily recognize the self-similar behavior of the wave function. Each peak branches into three peaks in the next stage. Presumably, these are the reflection of the self-similar distribution of the roots, Eq. (5.3).

We have performed a multifractal analysis[8,18,19] to investigate the nature of the wave functions. This is useful to distinguish critical wave functions from extended wave functions. The results for critical wave functions reveal multifractal properties. Let us consider a kth generation $Q=Q_k=F_{3k+1}$. First, we define a probability measure p_j ($\Sigma_j p_j=1$) of the wave function as

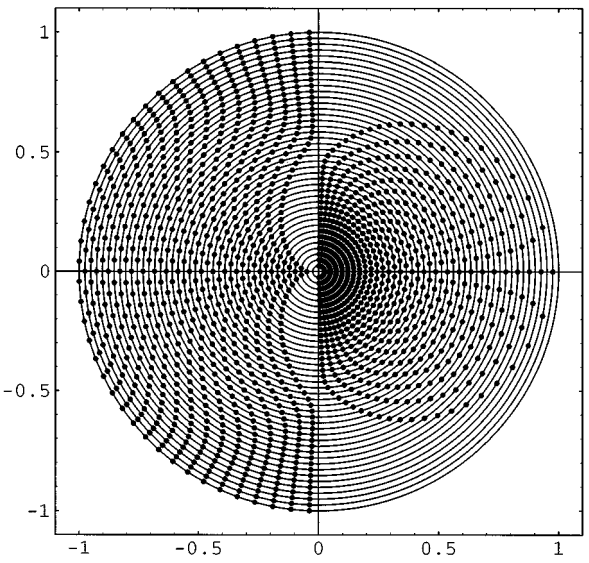

FIG. 9. The roots of the Bethe ansatz equations in the complex plane for each mid-band energy with $P=1$ and $Q=41$. All the roots are on the unit circle. The radii are scaled to show them in a diagram. The unit circle for a higher state is scaled to have a smaller radius.

$$p_j=\frac{1}{N_k}|\Psi_j|^2, \tag{6.10}$$

$$N_k=\sum_{j=1}^{Q_k}|\Psi_j|^2. \tag{6.11}$$

Next, we define a Lebesgue measure l_k of each site as

$$l_k=\frac{1}{Q_k}. \tag{6.12}$$

From p_j and l_k, the singularity of the probability measure is represented by an exponent α_j as

$$p_j=l_k^{\alpha_j}. \tag{6.13}$$

Let $\Omega_k(\alpha)d\alpha$ be the number of sites, the value of α_j of which lies in the interval $[\alpha,\alpha+d\alpha]$. One exploits the distribution $\Omega_k(\alpha)$ to characterize the nature of the wave function. Since Q_k increases exponentially as k increases, so does Ω_k. It is natural to introduce the entropy function $S(\alpha)$,[19] defined by

$$S(\alpha)=\lim_{k\to\infty}S_k(\alpha)=\lim_{k\to\infty}\frac{1}{k}\ln\Omega_k(\alpha), \tag{6.14}$$

or the function $f(\alpha)$ to characterize the wave function:

$$f(\alpha)=\frac{S(\alpha)}{\epsilon}, \tag{6.15}$$

where

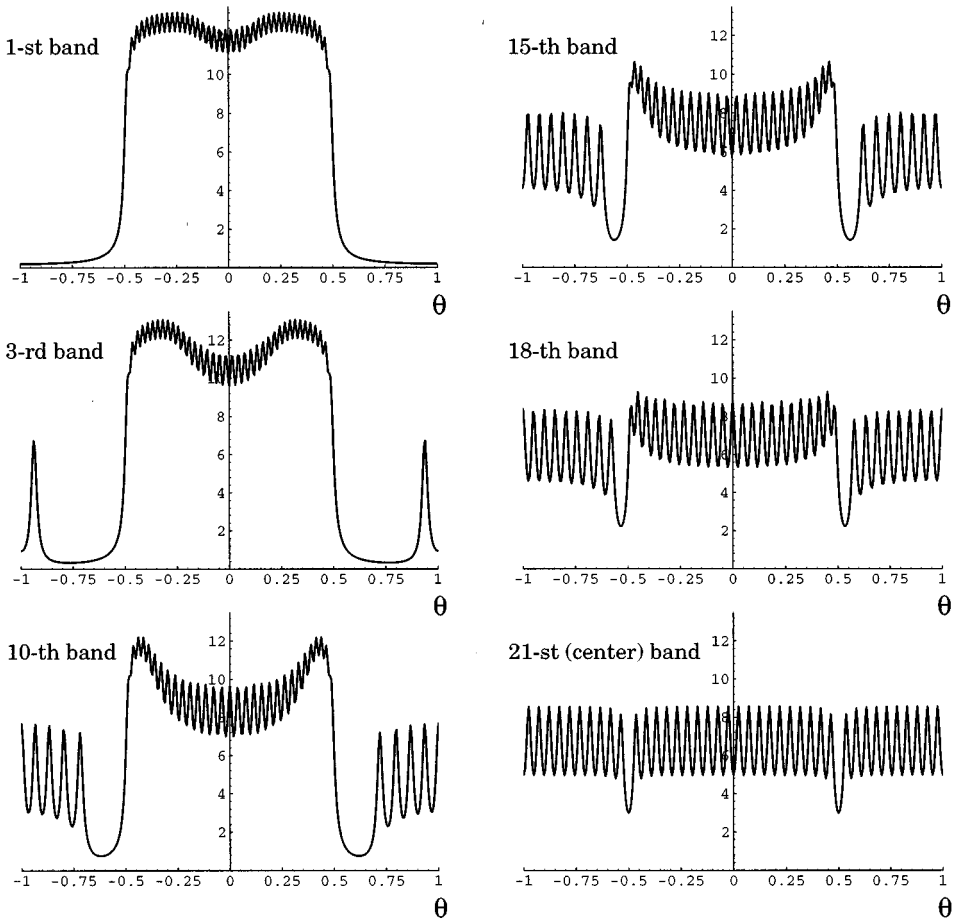

FIG. 10. Approximated distribution functions of the roots of the Bethe ansatz equations with $P=1$ and $Q=41$ for mid-band points in (a) the highest band; (b) the 3rd; (c) the 10th; (d) the 15th; (e) the 18th; and (f) the 21st band, which is the central band (with the mid-band point $E=0$).

$$\epsilon = \lim_{k\to\infty} \epsilon_k = -\lim_{k\to\infty} \frac{1}{k} \ln l_k . \tag{6.16}$$

Thus, we have $\Omega_k(\alpha) \approx l_k^{-f(\alpha)}$.

To calculate $f(\alpha)$, let us define the partition function

$$Z_k(r) = \sum_{j=1}^{Q_k} p_j^r = \sum_{j=1}^{Q_k} l_k^{-rk\alpha_j \epsilon_k} . \tag{6.17}$$

Then one obtains $f(\alpha) = \lim_{k\to\infty} f_k(\alpha)$ from

$$G_k(r) = \frac{1}{k} \ln Z_k(r), \tag{6.18}$$

$$\alpha = -\frac{1}{\epsilon_k} \frac{d}{dr} G_k(r), \tag{6.19}$$

$$f_k(\alpha) = \frac{1}{\epsilon_k} G_k(r) + r\alpha. \tag{6.20}$$

The maximum of $f(\alpha)$ gives the Hausdorff dimension of the Lebegue measure, so it is always one in the present cases. The nature of the wave functions is characterized by the function $f(\alpha)$ as follows: When the wave function is extended, sites with $p_j \approx l_k$ dominate, so $f(\alpha=1)$ at $\alpha=1$. On the other hand, for a localized wave function, $f(\alpha)$ consists of two points at $\alpha=0$ and $\alpha=\infty$. When $f(\alpha)$ is a continuous function, the maximum $f(\alpha)=1$ of which is at $\alpha=\alpha_0\neq 1$, the wave function is critical in between the extended and localized cases. We have performed the above multifractal analysis.[20] Here, we stress the importance of the finite-size effects. In order to have a reliable $f(\alpha)$, one must perform extrapolations of the finite-size data. [$f(\alpha)$ obtained from a finite system is different from the true $f(\alpha)$ (Refs. 8 and 21).] We have done such calculations. The results are shown in Fig. 7 for the golden-mean flux. It gives a smooth $f(\alpha)$. This clearly shows that this wave function is multifractal and critical. We note the striking resemblance of these wave functions to that of the $1d$ quasicrystal Fibonacci lattice at the center of the spectrum.[22] The latter was obtained exactly by a different method and $f(\alpha)$ is obtained analytically.[21]

VII. NUMERICAL SOLUTIONS FOR MID-BAND POINTS $E\neq 0$

For mid-band points other than $E=0$, we have not been able to obtain analytic results. A natural thought is to try a numerical approach. However, the Bethe ansatz equations are high-degree algebraic equations of many variables and a direct attack would be extremely difficult even numerically. Again, we have found that information on the quantum group can be exploited to reduce the difficulty. Instead of solving the Bethe ansatz equations (4.9) directly, we construct a polynomial $\Psi^{(j)}(z)$ for a mid-band energy $E^{(j)}$ by

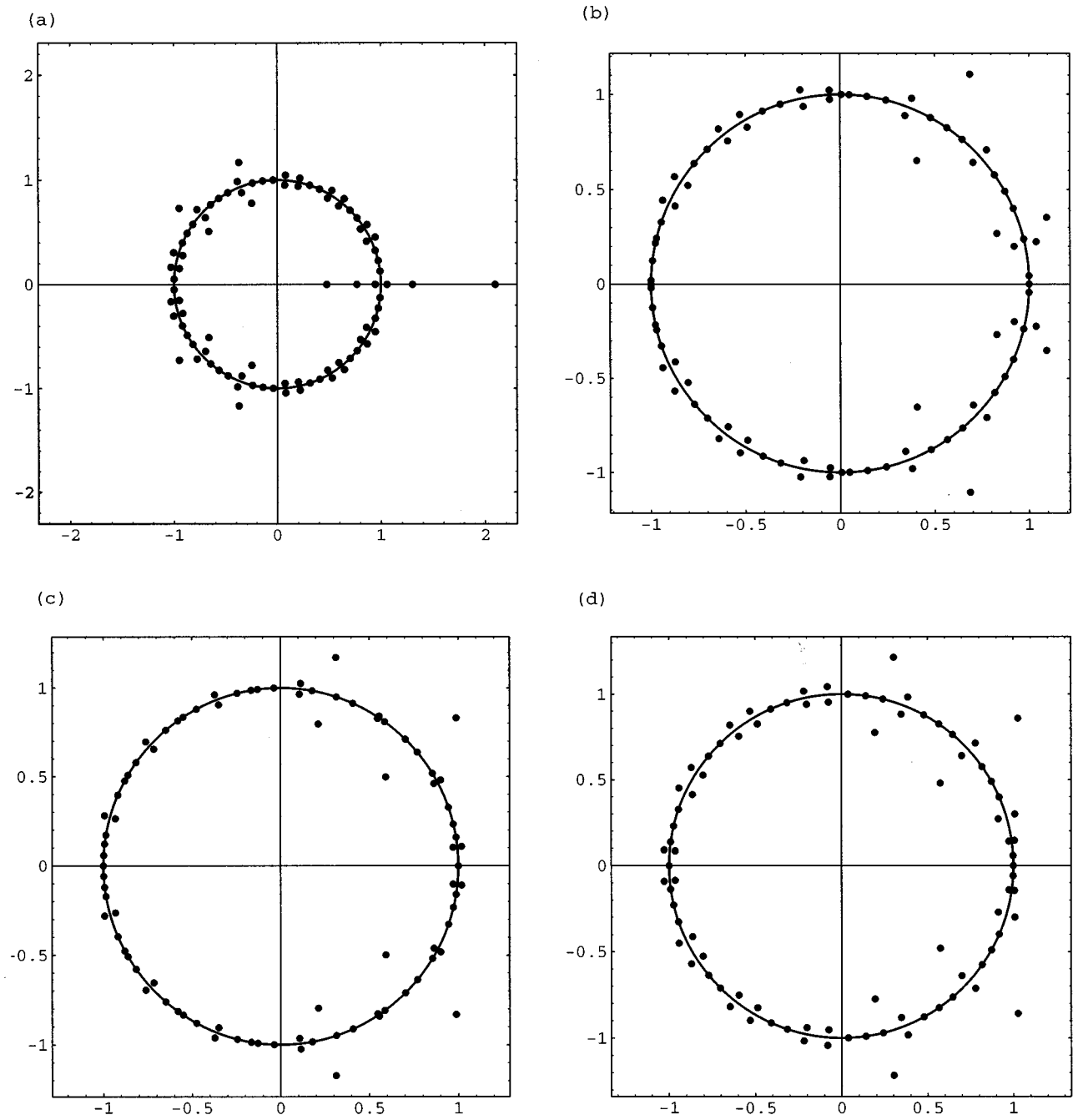

FIG. 11. The roots of the Bethe ansatz equations with $P=55$ and $Q=89$ for mid-band states in (a) the highest band, (b) the 22nd, (c) the 30th, and (d) the 34th band.

$$\Psi^{(j)}(z)=\sum_{m=0}^{Q-1} v_m^{(j)} z^m, \tag{7.1}$$

where $[v_0^{(j)},\ldots,v_{Q-1}^{(j)}]^t$ is given by the jth eigenvector of the $Q\times Q$ tridiagonal matrix $H_{\mathrm{MB}}^{\mathrm{tri}}$. The roots of the Bethe ansatz equations are given by the roots of this one-variable polynomial. Since the $H_{\mathrm{MB}}^{\mathrm{tri}}$ is real symmetric, $\Psi^j(z)$ is a real polynomial. Thus, the roots can be obtained by the traditional numerical techniques.

Let us consider the cases with $P=1$ first. We have calculated the roots of the Bethe ansatz equations for a large number of different odd-number values of Q. All roots are on the unit circle. We conjecture that all roots of the Bethe Ansatz equations are on the unit circle when $\phi=1/Q$ with odd Q. In Fig. 8, we present results for several mid-band points with $Q=61$. In Fig. 9, all roots for $Q=41$ are shown. We notice that for the highest energy band, the roots are on the right half of the unit circle and it is almost uniformly distributed (though not exactly for finite Q). For the second highest band, one root appears on the left semicircle (at $z=-1$). For the third highest band, one more root appears on the left, and so on. In this way, each time as the ordinal number of the band (from the top) decreases by one, one of the roots of the Bethe ansatz equations for the mid-band point moves from the right semicircle to the left. We have also calculated the

distribution function of roots. For a finite Q, we define an approximate distribution function $\rho_\epsilon^Q(\theta)$ as

$$\rho_\epsilon^Q(\theta)=\sum_{m=1}^{Q-1}\delta_\epsilon(\theta-\theta_m), \tag{7.2}$$

$$\delta_\epsilon(x)=-\frac{1}{\pi}\mathrm{Im}\left(\frac{1}{2\sin\left(\frac{x}{2}\right)-i\epsilon}\right), \tag{7.3}$$

where ϵ is the width of an approximate δ function and we take it to be $\epsilon=2/Q$. The distance is measured by the chord distance. In Fig. 10, we plot the results for several mid-band energies with $Q=41$. From these results for finite Q, one may speculate that the distributions for mid-band points $E\neq 0$ in the $Q\to\infty$ limit are likely to have singularities at $z=\pm i$, which we have seen is true for the state at $E=0$.

To get an impression for what happens in the cases with $P\neq 1$, we choose $P=55$ and $Q=89$ as an example and present the results in Fig. 11. The roots of the Bethe ansatz equations are no longer on the unit circle.

ACKNOWLEDGMENTS

This work was supported in part by U.S. NSF Grant No. PHYS-9309458 and Grant-in-Aid from the Ministry of Education, Science and Culture of Japan. Y.-S.W. thanks the Institute for Solid State Physics, University of Tokyo for support and warm hospitality during his visit, when this collaboration began. He also thanks R. Musto and H.C. Fu for helpful discussions, and, in particular, D. Mattis for useful comments on an early version of the manuscript. Y.H. and M.K. are grateful to P. Wiegmann for valuable discussions.

APPENDIX A: THE STRUCTURE FACTOR OF THE BETHE ANSATZ ROOTS

Here we derive Eq. (5.6) and Eq. (5.7). Using Eq. (5.2), one obtains

$$\begin{aligned}S_Q(k)&=\sum_{j=-\infty}^{\infty}e^{ikj}\tilde{S}_Q(j)=\sum_{l=-\infty}^{\infty}\sum_{m=1}^{Q-1}\left[e^{ik\{2Ql+2Pm+(1/2)(Q-1P)\}}+e^{ik\{2Ql-2Pm+(Q+P)/2\}}\right]\\&=\sum_{l=-\infty}^{\infty}e^{i2kQl}\left\{\frac{1-e^{iPk(Q-1)}}{1-e^{i2Pk}}e^{i(Q+3P)k/2}+\frac{1-e^{-iPk(Q-1)}}{1-e^{-i2Pk}}e^{i(Q-3P)k/2}\right\}\\&=\sum_{r'=0}^{2Q-1}\frac{2\pi}{2Q}\delta\left(k-\frac{2\pi r'}{2Q}\right)\frac{\sin P(Q-1)k/2}{\sin Pk}e^{iQk/2}2\cos\frac{1}{2}PQk.\end{aligned} \tag{A1}$$

The structure factor is a sum of δ functions, but the amplitude of the peaks has a nontrivial feature. At $k=2\pi r'/2Q$, we have

$$\frac{\sin P(Q-1)k/2}{\sin Pk}e^{iQk/2}2\cos\frac{1}{2}PQk=\begin{cases}Q-1, & r'=0\\ \dfrac{(-)^r}{\cos\pi\dfrac{P}{Q}r}, & r'=2r\neq 0\\ 0, & \text{otherwise.}\end{cases} \tag{A2}$$

This leads to Eq. (5.6) and Eq. (5.7). In the latter, $\cos Pk/2$ is a smooth function of $k\in(-\pi,\pi]$, when P is finite. It, however, oscillates wildly when $P\to\infty$, corresponding to an irrational flux.

APPENDIX B: EXPLICIT WAVE FUNCTIONS IN CLOSED FORM

From the explicit solution to the Bethe ansatz equations for $\phi=1/Q$, the wave function at site j can be written as

$$\begin{aligned}\psi_j&=\prod_{m=1}^{(Q-1)/2}(q^j-iq^{2m-1/2})(q^j-iq^{-2m+1/2})=q^{(Q-1)j}\prod_{m=1}^{(Q-1)/2}(1-iq^{2m-1/2-j})(q^j-iq^{-2m+1/2-j})\\&=(-q)^{-j}(iq^{3/2-j};q^2)_{(Q-1)/2}(iq^{-3/2-j};q^{-2})_{(Q-1)/2}.\end{aligned} \tag{B1}$$

The two sequences of roots that we have discussed before immediately lead to

$$\psi_j=q^{(Q-1)j}\prod_{m=1}^{(Q-1)/2}\left(1-\exp\left[i2\pi\frac{1}{2Q}\{P(2m-j)+(Q-P)/2\}\right]\right)\left(1-\exp\left[i2\pi\frac{1}{2Q}\{P(-2m-j)+(Q+P)/2\}\right]\right). \tag{B2}$$

Let us define J_1 by

$$PJ_1 \equiv \frac{1}{2}(Q-P) \pmod{2Q}, \tag{B3}$$

then $P(J_1+1)\equiv\frac{1}{2}(Q+P) \pmod{2Q}$ and $\psi_j=0$, for $j\equiv 2m+J_1$, $-2m+J_1+1$, $[m=1,\ldots(Q-1)/2]$. Thus, we have

$$\psi_{\bar{j}}=0 \quad \text{for } \bar{j}=2m \quad \text{or } -2m+1\,[m=1,\ldots(Q-1)/2], \tag{B4}$$

where $\bar{j}=j-J_1$.

Now let us consider the amplitude of the wave function. Apparently, $|\psi_{\bar{j}}|=|\psi_{\bar{j}+2Q}|$. However, $2Q$ is not the smallest period. Actually, the smallest period is Q, because $\{z_m\}=\{z_m^*\}$. Thus, we define l by $j=2l-1$ and consider $\psi(l)=\psi_{\bar{j}=2l-1}$ $[j=J_1+2l-1,\ l=1,2,\ldots,(Q+1)/2]$. Then, the squared amplitude is

$$|\psi(l)|^2=\prod_{m=1}^{(Q-1)/2}\left|1-\exp\left[i2\pi\left\{\frac{P}{2Q}(2m-1)-\frac{P}{Q}(l-1)\right\}\right]\right|^2\left|1-\exp\left[i2\pi\left\{\frac{P}{2Q}(-2m)-\frac{P}{Q}(l-1)\right\}\right]\right|^2 \tag{B5}$$

$$=\prod_{m=1}^{(Q-1)/2}(1-q^{-2(l-1)}q^{2m-1})(1-q^{2(l-1)}q^{-(2m-1)})(1-q^{-2(l-1)}q^{-2m})(1-q^{2(l-1)}q^{2m}). \tag{B6}$$

Let us first consider $|\psi(l)|$ for $l=1$. Denoting

$$S_1=\{2m-1|m=1,\ldots,(Q-1)/2\}=\{1,3,\ldots,Q-4,Q-2\},$$

$$S_2=\{-(2m-1)|m=1,\ldots,(Q-1)/2\}=\{-1,-3,\ldots,-(Q-4),-(Q-2)\}\equiv\{Q+2,Q+4\ldots,2Q-3,2Q-1\} \pmod{2Q},$$

$$S_3=\{-2m|m=1,\ldots,(Q-1)/2\}=\{-2,-4,\ldots,-(Q-3),-(Q-1)\}\equiv\{Q+1,Q+3\ldots,2Q-4,2Q-2\} \pmod{2Q},$$

$$S_4=\{2m|m=1,\ldots,(Q-1)/2\}=\{2,4,\ldots,(Q-3),(Q-1)\},$$

we have

$$S=S_1\cup S_2\cup S_3\cup S_4\equiv\{m|m=1,\ldots,2Q\}\backslash\{0,Q\} \pmod{2Q}. \tag{B7}$$

Since P and $2Q$ are mutually prime, S is invariant under a multiplication by $P \pmod{2Q}$. This leads to

$$\prod_{m=1}^{(Q-1)/2}(z-q^{2m-1})(z-q^{-(2m-1)})(z-q^{-2m})(z-q^{2m})=\prod_{m\in\{1,\ldots,2Q\}\backslash\{0,Q\}}(z-e^{i2\pi(m/2Q)})=\frac{z^{2Q}-1}{(z-1)(z+1)}. \tag{B8}$$

In a limit $z\to 1$, we have

$$|\psi(1)|^2=Q, \tag{B9}$$

which is independent of P.

In order to obtain the other amplitudes, we use the recursion relation, which is obtained from Eq. (B6):

$$|\psi(l+1)|^2=|\psi(l)|^2\frac{1-q^{-2(l-1)-1}}{1-q^{-2(l-1)+Q-2}}\frac{1-q^{2(l-1)+1}}{1-q^{2(l-1)-Q+2}}\frac{1-q^{-2(l-1)-(Q+1)}}{1-q^{-2(l-1)-2}}\frac{1-q^{2(l-1)+(Q+1)}}{1-q^{2(l-1)+2}}$$

$$=|\psi(l)|^2\left[\frac{\sin\pi\frac{P}{Q}(2l-1)}{\sin\pi\frac{P}{Q}2l}\right]^2. \tag{B10}$$

From this we have

$$|\psi(l+1)|^2=Q\left(\frac{[2l-1]_q!!}{[2l]_q!!}\right)^2, \tag{B11}$$

where $[2n]_q!!=[2]_q[4]_q\cdots[2n]_q$ and $[2n-1]_q!!=[1]_q[3]_q\cdots[2n-1]_q$.

Now fix $P=1$ and take the large Q limit. The continuum coordinate x_l and the square amplitudes are defined by

$$x_{2l-1}=(l-\tfrac{3}{4})\Delta x \quad [l=1,\ldots,(Q+1)/2], \tag{B12}$$

$$\Delta x=\frac{2}{Q}, \tag{B13}$$

$$n(x_{2l-1}) \equiv |\psi(l)|^2. \tag{B14}$$

In the large Q limit, x is in the interval (0,1). Taking Eq. (B10) up to the first order in Δx, we get

$$\frac{d}{dx}\ln n(x) = -\frac{d}{dx}\ln \sin \pi x. \tag{B15}$$

Thus,

$$n(x) = \frac{2xC}{\sin \pi x}, \tag{B16}$$

where C is a constant. The coefficient C is determined by considering $l = O(1) \ll Q$:

$$n(x_{2l+1}) = Q\prod_{j=1}^{l}\left[\frac{\sin 2\pi \frac{P}{Q}\left(j-\frac{1}{2}\right)}{\sin 2\pi \frac{P}{Q} j}\right] = Q\prod_{j=1}^{l}\left(\frac{[2j-1]_q}{[2j]_q}\right)^2. \tag{B17}$$

APPENDIX C: THE FINITE-SIZE CORRECTIONS FOR LARGE Q

We derive the finite size corrections near the edges of the wave function for $P=1$ and $E=0$. We write the amplitude as $|\psi(x_{2l+1})|^2 = C(l)2/\sin \pi x_{2l+1}$, where $x_{2l+1} = (l+1/4)2/Q$. The finite-size correction $C(l)$ near the edge is given by

$$C(l) = \frac{1}{2}|\psi(x_{2l+1})|^2 \sin \pi x_{2l+1} = \begin{cases} \frac{1}{2}Q\left(\frac{[2l-1]_q!!}{[2l]_q!!}\right)^2 \sin\frac{2\pi}{Q}(l+\frac{1}{4}) \to Q\prod_{k=1}^{l}\left(\frac{2k-1}{2k}\right)^2 \frac{\pi}{2Q}(2l+\frac{1}{2}) & (l/Q \to 0) \\ \frac{\pi}{2}(2l+\frac{1}{2})\prod_{k=1}^{l}\frac{2k-1}{2k+1}\frac{(2k-1)(2k+1)}{4k^2} = \frac{\pi}{2}\frac{2l+1/2}{2l+1}\prod_{k=1}^{l}\left(1-\frac{1}{4k^2}\right), \end{cases} \tag{C1}$$

where l is large but finite, and $Q \to \infty$, so that $l/Q \to 0$.

*Electronic address: hatsugai@coral.t.u-tokyo.ac.jp

[1] *Geometric Phases in Physics*, edited by A. Shapere and F. Wilczek (World Scientific, Singapore, 1989).

[2] R. B. Laughlin, Phys. Rev. B **23**, 5632 (1981).

[3] B. I. Halperin, Phys. Rev. B **25**, 2185 (1982).

[4] *The Quantum Hall Effect*, edited by R. E. Prange and S. M. Girvin (Springer-Verlag, Berlin, 1987); *Quantum Hall Effect*, edited by M. Stone (World Scientific, Singapore, 1992).

[5] D. J. Thouless, M. Kohmoto, P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982); Q. Niu, D. J. Thouless, and Y. S. Wu, Phys. Rev. B **31**, 3372 (1985).

[6] Y. Hatsugai, Phys. Rev. B **48**, 11 851 (1993); Phys. Rev. Lett. **71**, 3697 (1993).

[7] P. G. Harper, Proc. Phys. Soc. London Sect. A **68**, 874 (1955); G. H. Wannier, J. Math. Phys. **21**, 2844 (1980).

[8] H. Hiramoto and M. Kohmoto, Phys. Rev. B **40**, 8225 (1989); Int. J. Mod. Phys. B **6**, 281 (1992).

[9] D. R. Hofstadter, Phys. Rev. B **14**, 2239 (1976).

[10] M. Wilkinson, Proc. Phys. Soc. London Sect. A **391**, 305 (1984).

[11] P. B. Wiegmann and A. V. Zabrodin, Phys. Rev. Lett. **72**, 1890 (1994); see also Ref. 12.

[12] L. D. Fadeev and R. M. Kashaev (unpublished).

[13] Y. Hatsugai, M. Kohmoto, and Y. S. Wu, Phys. Rev. Lett. **73**, 1134 (1994).

[14] M. Kohmoto and Y. Hatsugai, Phys. Rev. B **41**, 9527 (1990).

[15] This $P_n(z)$ can be expressed in terms of the continuous q-ultraspherical polynomials as $P_n(z) = C_n[(z+z^{-1})/2); q|q^2]$. See G. Gasper and M. Rahman, *Basic Hypergeometric Series* (Cambridge University Press, Cambridge, England, 1990), p. 169.

[16] Our 1d difference equation (in the diagonal gauge) at $E=0$ has another solution given by $\tilde{\Psi}_l = \tilde{\Psi}(q^l)$, with $\tilde{\Psi}(z) = (q^2;q^2)_n/(q;q^2)_n(-iz^{-1})^{n+1}P_n(iz^{-1})$, which has a pole at $z=0$. This solution does not satisfy the pole-free condition, and separate consideration is needed. However, since $\{z_m\} = \{z_m^*\}$ as shown later, we have $|\Psi_l| = |\tilde{\Psi}_l|$. Thus, we only consider Ψ_l from now on.

[17] $|\Psi_{j=1}^{-}|^2 = Q$ for arbitrary $\phi = P/Q$.

[18] T. C. Halsey, M. H. Jensen, L. Kadanoff, I. Procaccia, and B. Shraiman, Phys. Rev. A **33**, 1141 (1986).

[19] M. Kohmoto, Phys. Rev. A **37**, 1345 (1988).

[20] When the flux is $1/Q$, we can calculate $f(\alpha)$ explicitly using Eq. (6.7). It consists of two points $f(\alpha=0)=0$ and $f(\alpha=1)=1$. Singular $1/x$ behavior near $x \approx 0$ gives $f(\alpha=0)=0$ and the rest gives $f(\alpha=1)=1$.

[21] T. Fujiwara, M. Kohmoto, and T. Tokihiro, Phys. Rev. B **40**, 7413 (1989).

[22] M. Kohmoto and J. R. Banavar, Phys. Rev. B **34**, 563 (1986).