

Critical wave functions and a Cantor-set spectrum of a one-dimensional quasicrystal model

Mahito Kohmoto and Bill Sutherland

Department of Physics, University of Utah, Salt Lake City, Utah 84112

Chao Tang*

*Department of Physics, University of Utah, Salt Lake City, Utah 84112
and The James Franck Institute, The University of Chicago, Chicago, Illinois 60637*

(Received 31 July 1986)

The electronic properties of a tight-binding model which possesses two types of hopping matrix element (or on-site energy) arranged in a Fibonacci sequence are studied. The wave functions are either self-similar (fractal) or chaotic and show "critical" (or "exotic") behavior. Scaling analysis for the self-similar wave functions at the center of the band and also at the edge of the band is performed. The energy spectrum is a Cantor set with zero Lebesgue measure. The density of states is singularly concentrated with an index α_E which takes a value in the range $[\alpha_E^{\min}, \alpha_E^{\max}]$. The fractal dimensions $f(\alpha_E)$ of these singularities in the Cantor set are calculated. This function $f(\alpha_E)$ represents the global scaling properties of the Cantor-set spectrum.

I. INTRODUCTION

There is much current interest in quasiperiodic structures. These systems are intermediate between the completely periodic perfect crystals and the random or disordered amorphous solids. Undoubtedly, a major push toward understanding them was given by the experiments of Shechtman *et al.*,¹ which seem to show some evidence for a quasicrystal in the material $\text{Al}_{0.86}\text{Mn}_{0.14}$.² The theoretical understanding of these structures is based on the non-periodic tiling of the two-dimensional plane first introduced by Penrose³ and described by Gardner.⁴ The papers of de Bruijn are probably the most complete investigation in print.^{5,6} The first suggestion that a Penrose tiling might serve as a model for a physical system was made by MacKay.^{7,8} A particularly simple quasicrystal structure is obtained by projecting a higher-dimensional lattice.⁹⁻¹³

Electronic properties of the two-dimensional Penrose lattice were studied by several groups.¹⁴⁻¹⁷ Kohmoto and Sutherland¹⁶ found localized states which are infinitely degenerate at the center of the spectrum. These states are localized due to the lattice topology and form rings of various dimensions (ring states).

A one-dimensional version of quasicrystals can be obtained by projecting a two-dimensional square lattice onto a line. For a choice of angle between the line and an edge of the square lattice whose tangent is an inverse of the golden mean $\phi^{-1} = (\sqrt{5} - 1)/2$, we obtain a Fibonacci lattice which is quasiperiodic or, more precisely, almost periodic. The separation of successive lattice points in the Fibonacci lattice takes one of the two values A and B . The sequence of A 's and B 's is the Fibonacci sequence S_∞ , which is constructed recursively as follows: $S_{j+1} = \{S_{j-1}, S_j\}$ for $j \geq 1$ with $S_0 = \{B\}$ and $S_1 = \{A\}$. An alternative method for constructing the Fibonacci sequence is to use the "inflation" transformation $A \rightarrow B'A'$, $B \rightarrow A'$. It is easy to check that S_j with A and B is transformed to S_{j+1} with A' and B' ; hence, the two

methods above produce the same Fibonacci sequence.

A natural, and the simplest, model for electronic properties of the one-dimensional quasicrystal is a tight-binding equation

$$t_{n+1}\psi_{n+1} + t_n\psi_{n-1} = E\psi_n, \quad (1.1)$$

where ψ_n denotes the wave function at the n th site and $\{t_i\}$ is the Fibonacci sequence with two kinds of hopping matrix elements t_A and t_B . This is an off-diagonal version of an almost periodic Schrödinger equation,

$$\psi_{n+1} + \psi_{n-1} + V_n\psi_n = E\psi_n, \quad (1.2)$$

where the potential V_n takes two values arranged in the Fibonacci way. This equation was actually studied by two groups¹⁸⁻²⁰ before the discovery of the tenfold Bragg pattern by Shechtman *et al.*

Although it seems that there is no simple transformation between (1.1) and (1.2), Kohmoto and Banavar²¹ obtained a renormalization-group (RG) equation for (1.1) which is the same as the one originally obtained by Kohmoto, Kadanoff, and Tang¹⁸ (KKT) and Ostlund, Pandit, Rand, Schellnhuber, and Siggia²⁰ for (1.2). Therefore, we would expect the same properties in the localization problem (energy spectrum, wave function, etc.) for the two models. The tight-binding model (1.1) was also studied numerically by Lu, Odagaki, and Birman,²² Nori and Rodriguez,²³ and Fujita and Machida.²⁴ The vibrational problem for the one-dimensional quasicrystal was studied using the KKT renormalization-group method in Ref. 21 and by Luck and Petritis.²⁵ These models are examples of the almost periodic Schrödinger equation, which has been of considerable interest to mathematicians²⁶ as well as to physicists.²⁷

The Cantor-set spectrum for the quasiperiodic model has a rich structure. In fact, it has scaling properties which come from criticality of the system. For the localization problem, there are three types of wave functions:

localized (normalizable), extended (unnormalizable), and critical. These three types of wave functions uniquely correspond to point, absolutely continuous, and singular continuous spectra, respectively.²⁸ Therefore, the spectral property determines the type of wave function.

Kohmoto²⁹ and Ostlund and Pandit³⁰ studied local scaling of the spectrum of a quasiperiodic model which is known to have a transition between localized and extended states.³¹ (This model is called Harper's equation or almost Mathieu equation.) These works show that existence of scaling for a spectrum intrinsically comes from a criticality, i.e., a singular continuous spectrum. Recently, Tang and Kohmoto³² investigated the global scaling properties of the spectrum for the above model. They used the method of Halsey, Jensen, Kadanoff, Procaccia, and Shraiman³³ which was originally applied to the analysis of fractal sets appearing at the onset of chaos.

In this paper we investigate the one-dimensional quasi-crystal model. Scaling aspects of the energy spectrum and the electronic states are emphasized. Strong evidence is presented for the spectrum being purely singular continuous. For this type of a spectrum, wave functions are "critical" or "exotic" and are neither localized nor extended in a standard fashion.

In Sec. II the KKT method to obtain the Cantor-set spectrum using a dynamical map is explained. In order to obtain the wave function, we need further analysis of the renormalization-group equation. This is discussed in Sec. III. The critical wave functions are analyzed in Sec. IV, and Sec. V contains the global scaling analysis of the Cantor-set spectrum. Section VI is a summary.

II. RENORMALIZATION-GROUP METHOD OF KOHMOTO, KADANOFF, AND TANG

In this section we review the KKT method¹⁸ which is applied to the diagonal model (1.2) and the off-diagonal model (1.1). (See also a parallel work of Ostlund *et al.*²⁰)

A. Diagonal model

The (discrete) almost periodic Schrödinger equation (1.2) can be written as

$$\begin{pmatrix} \psi_{n+1} \\ \psi_n \end{pmatrix} = \underline{M}(n) \begin{pmatrix} \psi_n \\ \psi_{n-1} \end{pmatrix}, \quad (2.1)$$

where $\underline{M}(n)$ is a transfer matrix,

$$\underline{M}(n) = \begin{pmatrix} E - V_n & -1 \\ 1 & 0 \end{pmatrix}. \quad (2.2)$$

Successive applications of the transfer matrices give values of the wave function at arbitrary sites as

$$\begin{pmatrix} \psi_{N+1} \\ \psi_N \end{pmatrix} = \underline{M}(N)\underline{M}(N-1) \cdots \underline{M}(2)\underline{M}(1) \begin{pmatrix} \psi_1 \\ \psi_0 \end{pmatrix}. \quad (2.3)$$

Therefore, solving the Schrödinger equation is completely equivalent to calculating products of transfer matrices.

For the Fibonacci lattice, the potential V_n take two values V_A and V_B , and the sequence of $\underline{M}(n)$'s is a Fibonacci sequence with two matrices

$$\underline{M}_A = \begin{pmatrix} E - V_A & -1 \\ 1 & 0 \end{pmatrix} \quad (2.4)$$

and

$$\underline{M}_B = \begin{pmatrix} E - V_B & -1 \\ 1 & 0 \end{pmatrix}. \quad (2.5)$$

Let us define \underline{M}_j by

$$\underline{M}_j = \underline{M}(F_j)\underline{M}(F_j-1) \cdots \underline{M}(2)\underline{M}(1), \quad (2.6)$$

where F_j is a Fibonacci number given by a recursion relation $F_{j+1} = F_j + F_{j-1}$ for $j \geq 1$ with $F_0 = F_1 = 1$. The matrix \underline{M}_j generates a wave function at the Fibonacci number site. Since a Fibonacci sequence is constructed as $S_{j+1} = \{S_{j-1}, S_j\}$ for $j \geq 1$ with $S_0 = \{B\}$ and $S_1 = \{A\}$, it can be shown that \underline{M}_j obeys

$$\underline{M}_{j+1} = \underline{M}_{j-1}\underline{M}_j \quad \text{for } j \geq 1, \quad (2.7)$$

with $\underline{M}_0 = \underline{M}_B$ and $\underline{M}_1 = \underline{M}_A$.

The matrix recursion relation (2.2) can be regarded as a kind of renormalization-group equation. Recall the second and equivalent way to construct a Fibonacci sequence by the inflation $A \rightarrow B'A'$, $B \rightarrow A'$. The transfer matrix is transformed as $\underline{M}_j(A, B) \rightarrow \underline{M}_{j+1}(A', B')$, where $\underline{M}_j(A, B)$ is a product of the F_j transfer matrices of \underline{M}_A 's and \underline{M}_B 's being arranged in a Fibonacci sequence, and $\underline{M}_{j+1}(A', B')$ is defined similarly.

Once \underline{M}_j is calculated, a wave function at a Fibonacci number site is obtained as

$$\begin{pmatrix} \psi_{F_{j+1}} \\ \psi_{F_j} \end{pmatrix} = \underline{M}_j \begin{pmatrix} \psi_1 \\ \psi_0 \end{pmatrix}. \quad (2.8)$$

In order to obtain an energy spectrum, we look for energies where corresponding solutions ψ_n do not grow exponentially. This leads to a condition that the modulus of the eigenvalue of \underline{M}_j should be unity as $j \rightarrow \infty$.³⁴ Note that \underline{M}_j depends on an energy E . Since the determinant of \underline{M}_j is unity (unimodular), the condition for E to be in the spectrum is that $|\text{Tr}\underline{M}_j| \leq 2$ or, in a relaxed version, that $\text{Tr}\underline{M}_j$ is bounded as $j \rightarrow \infty$, where Tr is a trace of a matrix.

The matrix recursion relation (2.7) can be considered to be a discrete dynamical system (mapping). Since \underline{M}_j is a real 2×2 matrix with unit determinant ($\text{SL}(2, \mathbf{R})$), three real numbers are needed to specify \underline{M}_j . Hence (2.7) is a six-dimensional mapping. As we will see later, a constant of motion exists for this map and the dimensionality is actually reduced to 5. A five-dimensional mapping is a very complicated problem to study. As explained above, the energy spectrum is determined from the behavior of a projection of an orbit in five dimensions to a trace of \underline{M}_j . Once a five-dimensional orbit giving a bounded $\text{Tr}\underline{M}_j$ is found, then the corresponding energy is in the spectrum and the wave function is obtained from the five-dimensional orbit.

Here we state a key theorem which enables us to merely study a two-dimensional dynamical system to determine the spectrum. Moreover, as will be seen in the next sec-

tion, the reduced dynamical system dictates the full dynamical system.

*Theorem.*¹⁸ Consider a set of matrices $\underline{M}_j \in \text{SL}(2, \mathbb{R})$ or $\text{SL}(2, \mathbb{C})$, $j \in \mathbb{Z}$, which satisfies $\underline{M}_{j+1} = \underline{M}_{j-1} \underline{M}_j$, then $\text{Tr} \underline{M}_{j+1} = \text{Tr} \underline{M}_j \text{Tr} \underline{M}_{j-1} - \text{Tr} \underline{M}_{j-2}$.

A proof is shown in the Appendix. To our knowledge this rather surprising and simple property of a set of 2×2 unimodular matrices was not known previously.

Define

$$x_j = \frac{1}{2} \text{Tr} \underline{M}_j; \quad (2.9)$$

then the theorem implies

$$x_{j+1} = 2x_j x_{j-1} - x_{j-2}, \quad (2.10)$$

and the initial condition for this reduced subdynamical system can be taken as

$$x_{-1} = 1, \quad x_0 = (E - V_B)/2,$$

and

$$x_1 = (E - V_A)/2. \quad (2.11)$$

By defining a three-dimensional vector $\mathbf{r}_j = (x_j, y_j, z_j) = (x_j, x_{j+1}, x_{j+2})$, (2.10) and (2.11) are alternatively written as

$$\mathbf{r}_{j+1} = \mathbf{F}(\mathbf{r}_j), \quad (2.12)$$

with an initial condition

$$\mathbf{r}_{-1} = (x_{-1}, y_{-1}, z_{-1}) = \left(1, \frac{E - V_B}{2}, \frac{E - V_A}{2} \right), \quad (2.13)$$

where \mathbf{F} is a nonlinear map in three dimension explicitly given by

$$x_{j+1} = y_j, \quad y_{j+1} = z_j,$$

and

$$z_{j+1} = 2y_j z_j - x_j. \quad (2.14)$$

The mapping (2.10) or (2.11) has a constant of motion

$$I = x_j^2 + y_j^2 + z_j^2 - 2x_j y_j z_j - 1. \quad (2.15)$$

By a direct substitution using (2.3), it can be shown that I is indeed independent of j . The initial condition (2.11) gives the value of I to be

$$I = \frac{1}{4} (V_A - V_B)^2. \quad (2.16)$$

This tells us that the difference of the two values of the potential is a key parameter of this problem as it should be, and it remains invariant under the RG transformation. The constant of motion I determines a two-dimensional manifold on which an orbit remains. The manifold is noncompact for $I > 0$. An example is shown in Fig. 1. At $I = 0$, the middle part of the manifold becomes compact and it touches the four noncompact manifolds at a single point, respectively. This case corresponds to $V_A = V_B$ and the model is simply a periodic (with period 1) one-band model and the complete solution is available rather trivially. For $I < 0$, the middle part is detached

from the four noncompact manifolds. This is an unphysical region as seen from (2.13) as far as the quasiperiodic electronic problem concerned. However, recently Sutherland³⁶ pointed out that dynamics of a spin under a quasiperiodically pulsed magnetic field is described by the mapping with $-1 < I < 0$.

B. Off-diagonal model

The tight-binding equation (1.1) is written as

$$\begin{bmatrix} \psi_{n+1} \\ \psi_n \end{bmatrix} = \underline{M}(n+1, n) \begin{bmatrix} \psi_n \\ \psi_{n-1} \end{bmatrix}, \quad (2.17)$$

with

$$\underline{M}(n+1, n) = \begin{bmatrix} E/t_{n+1} & -t_n/t_{n+1} \\ 1 & 0 \end{bmatrix}. \quad (2.18)$$

This problem is a little bit more complicated than the previous diagonal model since the transfer matrix (2.18) depends on two bonds n and $n+1$. [Compare with (2.2).] However, it is possible to obtain the same type of RG equation.²¹

First note that there are three types of transfer matrices,

$$\underline{M}_{AA} = \begin{bmatrix} E/t_A & -1 \\ 1 & 0 \end{bmatrix},$$

$$\underline{M}_{AB} = \begin{bmatrix} E/t_A & -t_B/t_A \\ 1 & 0 \end{bmatrix},$$

and

$$\underline{M}_{BA} = \begin{bmatrix} E/t_B & -t_A/t_B \\ 1 & 0 \end{bmatrix}. \quad (2.19)$$

On the other hand, we would like to have two types of

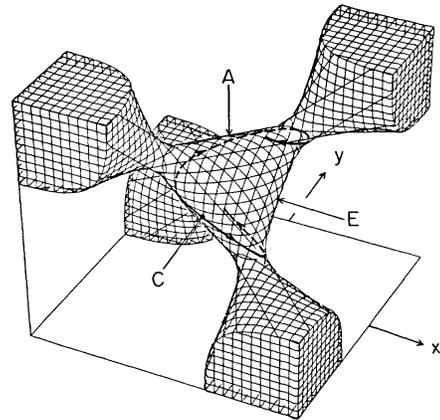


FIG. 1. Example of the manifolds for $I = 0.2$. There are six saddle points, A, B, C, D, E , and F . The points B, D , and F are located at the antipodal positions of E, A , and C , respectively. These points are the six-cycle of the trace map.

matrices to have a simple recursion relation. The Fibonacci sequence does not allow two consecutive bonds of type B , and so \underline{M}_{BA} is always followed by \underline{M}_{AB} . Instead of bonds B and A , we choose a bond A (which follows B) and two consecutive bonds BA to be basic units for the off-diagonal Fibonacci lattice. Since (A, BA) and (B, A) are related by the inflation transformation, one may write symbolically

$$\overline{B} = A$$

and (2.20)

$$\overline{A} = BA .$$

Two important observations here are the following: (i) Both of the two basic bond units \overline{B} and \overline{A} begin with a bond A and (ii) a single bond is used twice in a string of transfer matrices consisting of \underline{M}_{AB} , \underline{M}_{AB} , and \underline{M}_{BA} . Guided by these, we define

$$\underline{M}_{\overline{B}} = \underline{M}_{AA}$$

and (2.21)

$$\underline{M}_{\overline{A}} = \underline{M}_{AB}\underline{M}_{BA} .$$

Now, we have only two types of transfer matrices and it is evident from the way they are defined above that a transfer matrix \underline{M}_j representing a wave function at a Fibonacci number site F_j is given recursively as

$$\underline{M}_{j+1} = \underline{M}_{j-1}\underline{M}_j , \quad (2.22)$$

with an initial condition

$$\underline{M}_1 = \underline{M}_{\overline{B}} = \underline{M}_{AA}$$

and (2.23)

$$\underline{M}_2 = \underline{M}_{\overline{A}} = \underline{M}_{AB}\underline{M}_{BA} .$$

The initial condition for the trace map $x_{j+1} = 2x_j x_{j-1} - x_{j-2}$ is given by

$$x_{-1} = \frac{1}{2} \left(\frac{t_B}{t_A} + \frac{t_A}{t_B} \right), \quad x_0 = \frac{1}{2} \frac{E}{t_B} ,$$

and

$$x_1 = \frac{1}{2} \frac{E}{t_A} . \quad (2.24)$$

The constant of motion (2.15) is

$$I = \frac{1}{4} \left(\frac{t_B}{t_A} - \frac{t_A}{t_B} \right)^2 . \quad (2.25)$$

III. STRUCTURE OF THE TRANSFER MATRICES

It was shown in Sec. II that the trace map is sufficient to obtain the energy spectrum. In order to obtain the wave functions, however, we need solutions of the full matrix map. In this section, the structure of the transfer matrices is investigated. We emphasize the ‘‘Lorentz’’ transformation of the transfer matrices which leaves a trace unchanged. Different matrices which have a com-

mon value of trace are related to each other by such a Lorentz transformation. Once an orbit of the trace map is known, a corresponding orbit of the full matrix map is constructed by Lorentz transformations. This observation is extremely useful when we analyze wave functions which correspond to cycles of the trace map. It turns out that we need only one additional matrix which represents Lorentz transformation to produce a full orbit of the matrix map from a cycle of the trace map.

The transfer matrices are elements of $SL(2, \mathbb{R})$, the set of all real, 2×2 matrices with determinant 1, and hence invertible. If $\underline{A}, \underline{B}$ are two matrices in $SL(2, \mathbb{R})$, then when we multiply on the left by a third matrix in $SL(2, \mathbb{R})$, we have

$$\underline{A}, \underline{B} \rightarrow \underline{A}', \underline{B}' = \underline{C} \underline{A}, \underline{C} \underline{B} .$$

Considering the Killing form $(\underline{B}, \underline{A}) = \text{Tr}(\underline{B}^{-1}\underline{A})/2$ as a scalar product, we see that it is invariant under both left and right multiplication.

Let us introduce a basis for the 2×2 matrices as

$$\tau_0 = 1, \quad \tau_1 = i\sigma_y, \quad \tau_2 = \sigma_z, \quad \tau_3 = \sigma_x , \quad (3.1)$$

where σ_j are the Pauli spin matrices. Thus the τ_j are real. We note

$$\text{Tr}(\tau_0)/2 = 1, \quad \text{while } \text{Tr}(\tau_j)/2 = 0, \quad j = 1, 2, 3 . \quad (3.2)$$

Also,

$$(\tau_0)^2 = -(\tau_1)^2 = (\tau_2)^2 = (\tau_3)^2 = 1 . \quad (3.3)$$

The multiplication rules are

$$\tau_1\tau_2 = -\tau_3 = -\tau_2\tau_1, \quad \tau_2\tau_3 = \tau_1 = -\tau_3\tau_2 , \quad (3.4)$$

$$\tau_3\tau_1 = -\tau_2 = -\tau_1\tau_3 .$$

Then we can write a general matrix \underline{A} in $SL(2, \mathbb{R})$ as

$$\underline{A} = y\tau_0 + a_1\tau_1 + a_2\tau_2 + a_3\tau_3 . \quad (3.5)$$

The coefficients are obtained by projecting out on the τ 's using the scalar product. The condition that the determinant is 1 translates into the following condition on the coefficients:

$$y^2 + a_1a_1 - a_2a_2 - a_3a_3 = 1 . \quad (3.6)$$

This suggests that we parametrize the coefficients as

$$y = \cosh(\theta) \cos(\varphi), \quad a_1 = \cosh(\theta) \sin(\varphi) , \quad (3.7)$$

$$a_2 = \sinh(\theta) \cos(\psi), \quad a_3 = \sinh(\theta) \sin(\psi) .$$

The range of parameters is $-\pi < \varphi \leq \pi$, $-\pi < \psi \leq \pi$, $0 \leq \theta$. Thus, the parameter space of φ, ψ has the topology of a torus. When $\theta \rightarrow 0$, then the torus degenerates into a circle with the single coordinate φ ; this corresponds to the invariant I of the trace map equal to zero.

The scalar product becomes

$$(\underline{B}, \underline{A}) = xy + b_1a_1 - b_2a_2 - b_3a_3 = b_0a_0 + \mathbf{b} \cdot \mathbf{a} . \quad (3.8)$$

We have introduced a notation where the bold letter \mathbf{a} represents a Lorentz three-vector with timelike component a_1 and spacelike components a_2, a_3 . Then the Lorentz invariant scalar product $\mathbf{b} \cdot \mathbf{a}$ between Lorentz

three-vectors is given by

$$\mathbf{b} \cdot \mathbf{a} = b_1 a_1 - b_2 a_2 - b_3 a_3. \quad (3.9)$$

Note that the “length” $|\mathbf{a}|$ of \mathbf{a} is variable, since

$$\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2 = a_1 a_1 - a_2 a_2 - a_3 a_3 = 1 - y^2. \quad (3.10)$$

In particular, \mathbf{a} can be either timelike or spacelike, as $|y| < 1$ or > 1 . Given the length $|\mathbf{a}| \leq 1$, we can determine y up to a sign by $y = \pm(1 - |\mathbf{a}|^2)^{1/2}$. Therefore, up to the sign of y , we may parametrize the element \underline{A} by \mathbf{a} , writing $\underline{A}(\mathbf{a})$. The sign of y we call the parity $P(\underline{A})$ of \underline{A} .

This emphasis on Lorentz invariance is important, and the previous consideration is not simply a parametrization for convenience, for if we consider the Lorentz transformation of $\text{SL}(2, \mathbb{R})$, they are given as transformations of the form

$$\underline{A}(\mathbf{a}) \rightarrow \underline{A}' = \underline{C}(\mathbf{c}) \underline{A}(\mathbf{a}) \underline{C}^{-1}(\mathbf{c}) = \underline{A}(\mathbf{a}') = \underline{A}[\underline{L}(\mathbf{c})\mathbf{a}], \quad (3.11)$$

where $\underline{L}(\mathbf{c})$ is the three-dimensional Lorentz transformation with parameter \mathbf{c} . These Lorentz transformations preserve the Lorentz scalar product $\mathbf{b} \cdot \mathbf{a}$, and thus the three-lengths $\mathbf{a} \cdot \mathbf{a} = |\mathbf{a}|^2$. In particular, the ambiguity in the sign of c_0 is not important, for either choice leads to the same Lorentz transformation $\underline{L}(\mathbf{c})$, much in the same way as the relation between $\text{SU}(2)$ and the rotations $\text{O}(3)$ is two-to-one. Furthermore, the Lorentz transformations preserve the sign or parity of a_0 . Thus $P(\underline{A})$ and $|\mathbf{a}|$, or equivalently a_0 , constitute a complete set of invariants of an element \underline{A} under Lorentz transformations. Finally, we emphasize that the Lorentz transformations include one-dimensional rotations about the a_1 axis, as well as boosts along a two-dimensional velocity vector in the (a_2, a_3) plane.

The traces of matrices previously introduced in this paper have a simple relationship to these Lorentz invariants, for

$$\text{Tr}(\underline{A})/2 = y, \quad \text{Tr}(\underline{B})/2 = x,$$

and thus are preserved under a Lorentz transformation. Furthermore,

$$\text{Tr}(\underline{B} \underline{A})/2 = yx - \mathbf{a} \cdot \mathbf{b} = z.$$

Since x, y and $\mathbf{a} \cdot \mathbf{b}$ are each Lorentz invariants, the trace of the product is a Lorentz invariant. (This could also have been shown directly.)

Furthermore, these three invariants are a complete set for the matrix pair $(\underline{B}, \underline{A})$. By a complete set of invariants, we mean that if we have another matrix pair, $(\underline{B}', \underline{A}')$ with the same invariants, then there exists a unique Lorentz transformation \underline{L} (up to parity) which will bring the two pairs into alignment:

$$(\underline{B}', \underline{A}') = \underline{L}(\underline{B}, \underline{A}) \underline{L}^{-1}. \quad (3.12)$$

We make use of this observation in the following way.

Suppose we find a Q -cycle of the trace map. Since the traces of three consecutive matrices repeat, this means that the Lorentz invariants of two consecutive matrices, say, $\underline{B}_{k+Q} = \underline{M}_{k+Q-1}$ and $\underline{A}_{k+Q} = \underline{M}_{k+Q}$, repeat as well. Thus, the Q -times-iterated pair $(\underline{B}_{k+Q}, \underline{A}_{k+Q})$ is

equivalent by a Lorentz transformation to the original pair $(\underline{B}_k, \underline{A}_k)$, or

$$(\underline{B}_{k+Q}, \underline{A}_{k+Q}) = \underline{T}(\underline{B}_k, \underline{A}_k) \underline{T}^{-1}. \quad (3.13)$$

Since the matrix dynamics is invariant under Lorentz transformations, if two points on an orbit of the matrix map a distance Q apart are equivalent by a Lorentz transformation \underline{T} , then all pairs of points on the orbit a distance Q apart are equivalent by the same Lorentz transformation \underline{T} . Therefore, we emphasize the important point that this Lorentz transformation \underline{T} is the same for all k .

Therefore, an invariant set of the matrix map corresponding to the Q -cycle of the trace map is the set of all distinct pairs of matrices $\underline{T}^n(\underline{B}_k, \underline{A}_k) \underline{T}^{-n}$, for all integer n , and $k = 1, 2, \dots, Q$. These sets lie within the Q one-parameter submanifolds $\underline{T}^n(\underline{B}_k, \underline{A}_k) \underline{T}^{-n}$, where the parameter n is now any real number. More complicated invariant sets of the trace map than Q -cycles can likewise be treated.

We now show how the Lorentz transformation is determined. Suppose we have a matrix pair $(\underline{B}, \underline{A})$. Apart from an arbitrary Lorentz frame, it is specified by the set of three Lorentz invariants, which we take to be $x = \text{Tr}(\underline{B})/2$, $y = \text{Tr}(\underline{A})/2$, and $z = \text{Tr}(\underline{B} \underline{A})/2$ and denote collectively as the parameter $\mathbf{r} = (x, y, z)$. We choose our Lorentz frame to bring the matrix pair into a standard alignment, which we take as

$$\begin{aligned} \underline{B} &= x\tau_0 + b_1\tau_1 + b_2\tau_2, \quad b_1, b_2 > 0, \\ \underline{A} &= \begin{cases} y\tau_0 + a_1\tau_1, & \text{if } |y| < 1, \\ y\tau_0 + a_2\tau_2, & \text{if } |y| > 1. \end{cases} \end{aligned} \quad (3.14)$$

We write this initial pair, in standard alignment, with parameter \mathbf{r} as $(\underline{B}[\mathbf{r}], \underline{A}[\mathbf{r}])$. The notation is as follows: A pair $(\underline{B}[\mathbf{r}], \underline{A}[\mathbf{r}])$ will always denote a matrix pair, in standard alignment, with parameter \mathbf{r} . The brackets are to distinguish the pair with parameter \mathbf{r} from a matrix $\underline{A}(\mathbf{a})$ labeled by the Lorentz three-vector which is not in standard alignment in general. This notation will be followed in the rest of the paper.

We now iterate the matrix map once to give the new matrix pair $(\underline{B}', \underline{A}') = (\underline{A}, \underline{B} \underline{A})$ with parameter \mathbf{r}' , which is the old parameter \mathbf{r} iterated once by the trace map. This new matrix pair is not in standard alignment. It can, however, be brought into standard alignment by a Lorentz transformation \underline{S} , so that

$$(\underline{B}', \underline{A}') = \underline{S}[\mathbf{r}](\underline{B}[\mathbf{r}'], \underline{A}[\mathbf{r}']) \underline{S}^{-1}[\mathbf{r}]. \quad (3.15)$$

In words, we start with a pair $(\underline{B}[\mathbf{r}], \underline{A}[\mathbf{r}])$ in standard alignment, so that after iteration the pair $(\underline{B}', \underline{A}')$ is completely determined by the parameter \mathbf{r} ; in particular, the parameter \mathbf{r}' is determined by \mathbf{r} through the trace map. But furthermore, the orientation of the pair $(\underline{B}', \underline{A}')$ is also determined uniquely by \mathbf{r} , so the Lorentz transformation \underline{S} needed to bring the pair $(\underline{B}', \underline{A}')$ into the standard alignment $(\underline{B}(\mathbf{r}'), \underline{A}(\mathbf{r}'))$ is uniquely determined by \mathbf{r} . This dependence of \underline{S} on \mathbf{r} we write as $\underline{S}[\mathbf{r}]$. It is *not* a Lorentz transformation $\underline{L}(\mathbf{r})$ with parameter \mathbf{r} ; that is why we denote it by the symbol $\underline{S}[\]$ instead of $\underline{L}(\)$.

Now, by repeating this procedure each time, as we iterate the matrix map k times, we obtain

$$(\underline{M}_{k+1}, \underline{M}_k) = \underline{S}[\mathbf{r}_1] \cdots \underline{S}[\mathbf{r}_k] (\underline{B}[\mathbf{r}_{k+1}], \underline{A}[\mathbf{r}_{k+1}]) \\ \times \underline{S}^{-1}[\mathbf{r}_k]^{-1} \cdots \underline{S}^{-1}[\mathbf{r}_1]. \quad (3.16)$$

The orbit of the trace map is $(\dots, \mathbf{r}_k, \mathbf{r}_{k-1}, \dots, \mathbf{r}_2, \mathbf{r}_1)$. Thus if we have a Q -cycle of the trace map, so that $\mathbf{r}_{Q+1} = \mathbf{r}_1$, then we can make the identification

$$\underline{T} = \underline{S}[\mathbf{r}_1] \cdots \underline{S}[\mathbf{r}_Q]. \quad (3.17)$$

This expression must be invariant if we translate along the orbit of the trace map.

To summarize then, we have demonstrated the following structure and dynamics for the transfer matrices: The time evolution of the matrices consists of the evolution of the traces by the trace map, followed by a Lorentz transformation. Furthermore, the Lorentz transformation is completely determined by the trace map alone. Thus, the matrix dynamics is integrable, and we need only concentrate our attention on the trace dynamics.

IV. SELF-SIMILAR WAVE FUNCTIONS

Equipped with the (RG) analysis of the transfer matrices described in Secs. II and III, we are ready to understand the critical wave functions. We analyze two important wave functions, at the center of the band and the edge of the bands, respectively. They correspond to the six-cycle and two-cycle of the trace map, respectively. It is shown that wave functions corresponding to a cycle of the trace map indeed have a self-similar structure, and consequently, they are fractal.

We define two exponents β and $\alpha(\lambda)$ to characterize these wave functions. An exponent β describes a power-law behavior of an envelope of a wave function as

$$|\psi_n| \sim n^\beta. \quad (4.1)$$

This relation is for the sites with a wave function having increasing peak values which are encountered sequentially as n is increased. (See Fig. 4.) On the other hand, $\alpha(\lambda)$ describes how a measure (defined below) is concentrated spatially:

$$S(n) = \sum_{m \leq |n-n_0|} |\psi_m|^\lambda \sim n^{\alpha(\lambda)}, \quad (4.2)$$

where n_0 is a site where a wave function has the maximum value in the neighborhood. When $\lambda=2$, the measure $|\psi_m|^2$ represents a probability for an electron to be at a site m . We consider general values for λ in a later analysis.

A. Wave function at the center of the band

The six-cycle $(0, y, 0, 0, -y, 0)$ of the trace map dominates the behavior of the wave function at the center of the band.^{19,20} See Fig. 2 for an example of the spectrum. Actually, the full matrix map also has a six-cycle.²⁰ Therefore, we have a special case where the Lorentz transformation described in Sec. IV is unity. Moreover, it is pointed out in Ref. 21 that the six-cycle represents an

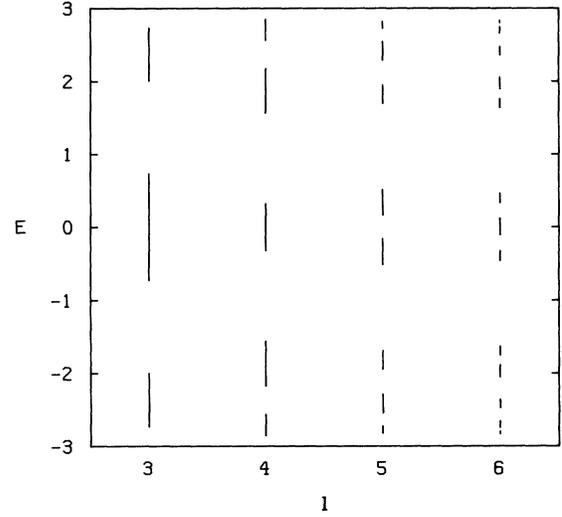


FIG. 2. Band structures of (1.1) with $t_A=1$ and $t_B=2$. These are periodic approximations and become a better approximation to the Cantor-set spectrum as one goes to the right.

exact solution of the off-diagonal model (1.1) at $E=0$. For the diagonal model (1.2), we need to find a stable manifold of the six-cycle of the dynamical map. A particular value of E which is close to zero puts an initial point on the stable manifold. This initial point flows into the six-cycle. Thus the wave function corresponding to this orbit is governed by the six-cycle at length scales larger than some distance which is determined by the flow rate of the orbit into the six-cycle.

The six-cycle governs the wave function of not only the center of the band but also each center of subcluster of any energy scale in the Cantor-set spectrum. This properly comes from the Smale horseshoe structure in the trace map.¹⁹

The matrix six-cycle (1.1) is explicitly given by

$$\underline{M}_1 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \underline{M}_2 = \begin{bmatrix} -e^\theta & 0 \\ 0 & -e^{-\theta} \end{bmatrix}, \\ \underline{M}_3 = \begin{bmatrix} 0 & e^{-\theta} \\ -e^\theta & 0 \end{bmatrix}, \quad \underline{M}_4 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad (4.3) \\ \underline{M}_5 = \begin{bmatrix} e^{-\theta} & 0 \\ 0 & e^\theta \end{bmatrix}, \quad \underline{M}_6 = \begin{bmatrix} 0 & -e^\theta \\ e^{-\theta} & 0 \end{bmatrix}.$$

It is easy to check using (2.19) and (2.23) that this six-cycle represents an exact solution of the off-diagonal model (1.1) at $E=0$ with $e^\theta = t_B/t_A$.

An exponent β has been calculated in Ref. 21 and is given by

$$\beta = |\ln(t_B/t_A)| / \ln \phi^3 = |\theta| / \ln \phi^3, \quad (4.4)$$

where $\phi = (\sqrt{5} + 1)/2$ is the golden mean. The reason for a power-law behavior of the wave function is basically that

$$\underline{M}_3 \underline{M}_6 = \begin{pmatrix} e^{-2\theta} & 0 \\ 0 & e^{2\theta} \end{pmatrix}$$

has an eigenvalue which is greater than unity. Hence, even if the wave function does not grow and only takes the values $\pm 1, \pm e^{\pm\theta}$ at Fibonacci sites, it grows as a power law at non-Fibonacci sites.

Let us calculate $\alpha(\lambda)$. For the sake of simplicity, we denote

$$\begin{aligned} B \leftarrow M_1 &= M_{\bar{B}} \\ \text{and} & \\ A \leftarrow M_2 &= M_{\bar{A}}. \end{aligned} \tag{4.5}$$

We note the relations $B^2 = (BA)^2 = -1$. These are the defining relations of a discrete group generated by B and A . Any element g of the group can be written as

$$g(q,r) = B^q A^r; \quad q = 0, 1, 2, 3$$

and

$$r = \dots, -2, -1, 0, 1, 2, 3, \dots \tag{4.6}$$

Therefore, after we have moved down the lattice n steps, we need only give the coordinates (q,r) to specify the element g of the group, and hence the transfer matrix $M(n) = g$.

The quantity we wish to calculate is $P(g | n)$ defined as the number of times $M(m)$ is equal to $g(q,r)$ as $m = 1, 2, \dots, n$. We shall calculate this quantity by inflating n through the Fibonacci numbers, so that starting from a single matrix A , n takes the values $F_{3k+1}, k = 0, 1, 2, \dots$. The inflation scheme we use is a symmetric, two-step similarity transformation of the infinite lattice, given by

$$B \rightarrow B' = ABA, \quad A \rightarrow A' = ABABA. \tag{4.7}$$

However, using the defining relations of the group, we can rewrite these as

$$B \rightarrow B' = ABA = B, \quad A \rightarrow A' = ABABA = -A. \tag{4.8}$$

Thus, after an inflation, the point n on the lattice has moved considerably further down the lattice, to a point n' , but the matrix $M(n)$ is equal to $M(n')$, up to a possible sign change. The sign change depends on whether the original expression for $M(n)$ contained an even or an odd number of A 's. But since both the defining relations of the group, and the inflation transformation only change the number of A 's by an even number, we can write the following: If $M(n) = g(q,r)$, then $M(n') = g(q+2r,r)$. (Addition for q is modulo 4.)

Clearly it is easier to not even keep track of n and n' at all, but instead to simply keep track of the number of $M(n)$ that takes the value $g(q,r)$, since these values are invariant under inflation. However, inflation does introduce many additional matrices; that is the advantage of the inflation transformation. Thus, suppose we have $M(n)$ and this is followed by B to give $M(n+1) = BM(n)$. Then after inflation, we have $M(n') = \pm M(n)$, but this is followed by A , then B , and finally by A before we arrive at $\pm M(n+1) = \pm BM(n)$, the image of $M(n)$ under inflation.

The point is, we need to keep track not only of which

element of the group we visit, but also of the matrix by which we will multiply this element next. Thus, we are finally led to define the quantity $P(g,a;k)$ as the number of times $M(n)$ takes the value $g = g(q,r)$, with the next matrix being $a = A$ or B , after having inflated k times. Of course, we must also specify the initial conditions—what the lattice was before inflation. We expect, however, that P approaches a limiting form independent of initial conditions. As we shall see, this is the case. The connection with the previous distribution is that

$$P(g | F_{3k+1}) = \sum_a P(g,a;k), \tag{4.9}$$

assuming the inflation starts with the single matrix A .

Next, we write a linear, Markov equation for the evolution of this distribution under inflation; i.e., we determine the coefficients in the equation

$$P(g,a;k+1) = \sum_{g'} \sum_{a'} dT(g,a | g',a') P(g',a';k). \tag{4.10}$$

The equations are considerably simplified by our previous observation that $g(q,r)$ can only change by at most a sign, reflected in the parity of q .

The result of multiplying $g(q,r)$ by either B or A is found from the defining relations to be

$$Bg(q,r) = g(q+1,r), \quad Ag(q,r) = g(q,r + (-1)^q). \tag{4.11}$$

We are now ready to determine the coefficients in the evolution equation by answering the following question: Suppose we have a transfer matrix $g(q,r)$ and the next matrix is B , what happens after inflation? We symbolize this situation as a "step" (q,r,B) . In place of B , after inflation we have ABA , so we have the steps

$$\begin{aligned} &(q+2r,r,A), \quad (q+2r,r+(-1)^q,B), \\ &(q+2r+1,r+(-1)^q,A). \end{aligned}$$

These all make a contribution to the evolution equation.

On the other hand, if we have a transfer matrix $g(q,r)$ and the next matrix is A , the step is (q,r,A) . Then after inflation, the single step is replaced by five steps:

$$\begin{aligned} &(q+2r,r,A), \quad (q+2r,r+(-1)^q,B), \\ &(q+2r+1,r+(-q)^q,A), \quad (q+2r+1,r,B), \\ &(q+2r+2,r,A). \end{aligned}$$

These then are all the contributions to the evolution equation.

Collecting together all of these contributions to a given step $(g,a) = (q,r,a)$ after inflation, we find for the evolution equation

$$\begin{aligned} P(q,r,B;k+1) &= P(q-2R-2,r-(-1)^q,B;k) \\ &\quad + P(q-2r-2,r-(-1)^q,A;k) \\ &\quad + P(q-2r-1,r,A;k), \\ P(q,r,A;k+1) &= P(q-2r,r,A;k) \\ &\quad + P(q-2r+1,r-(-1)^q,A;k) \\ &\quad + P(q-2r-2,r,A;k) + P(q-2r,r,B;k) \\ &\quad + P(q-2r+1,r-(-1)^q,A;k). \end{aligned} \tag{4.12}$$

These equations are very easy to iterate. We find that the results are not sensitive to initial conditions. Most importantly, all distribution functions are found to iterate to a scaling form

$$P(q, n, a; k) \rightarrow P_{q,a}(z) \exp[kf(z)], \quad (4.13)$$

with the scaled variable z being given by

$$-1 \leq z = r/k \leq 1. \quad (4.14)$$

The scaling function $f(z)$ is well fit by the expression

$$f(z) = 3 \ln(\phi)(1 - z^2), \quad (4.15)$$

where ϕ is the golden mean, given by $\phi = (\sqrt{5} + 1)/2$. (However, see the *Note added in proof* at the end of this paper.)

If we take as our initial wave function for the transfer matrices either of the two choices

$$\psi_+ = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \psi_- = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad (4.16)$$

then

$$M(n)\psi_{\pm} = \psi_{\pm}(n) = g(q, r)\psi_{\pm} = B^q A^r \psi_{\pm} = e^{\pm r\theta} B^q \psi_{\pm} \quad (4.17)$$

and thus

$$|\psi_{\pm}(n)| = e^{\pm r\theta}. \quad (4.18)$$

To summarize, the distribution of transfer matrices $P(g | n)$ gives us the distribution of wave functions as well, and this distribution function approaches the same universal scaling limit.

After inflation k times, we have a lattice of length $n = F_{3k+1} \rightarrow \phi^{3k}$. The maximum value of the wave functions is $e^{k\theta}$; let us normalize the wave function to this value, so that $|\psi_{\pm}(n)| = e^{-(k \pm r)\theta}$.

We now calculate a quantity which we expect to scale as a power of the length of the lattice,

$$\sum_{m \leq n} |\psi(m)|^{\lambda} \rightarrow n^{\alpha(\lambda)}. \quad (4.19)$$

This quantity can be expressed in terms of our distribution functions as

$$\sum_g P(g | n) e^{-(k \pm r)\lambda\theta} \rightarrow e^{-k\lambda\theta} \times \int_{-1}^1 dz P(z) \exp[kf(z) + kz\lambda\theta]. \quad (4.20)$$

Since k is large, this integral can be evaluated by saddle-point methods, which essentially consist in finding the maximum z_0 of the exponential in the interval -1 to 1 . Then we have

$$n^{\alpha(\lambda)} = \exp\{k[z_0\lambda\theta + f(z_0) - \lambda\theta]\}, \quad (4.21)$$

where $n \rightarrow \phi^{3k}$. Substituting the scaling expression for $f(z)$, we find

$$z_0 = \begin{cases} \lambda\theta / \ln(\phi^6), & |\lambda\theta| < \ln(\phi^6), \\ 1, & \lambda\theta > \ln(\phi^6), \\ -1, & \lambda\theta < -\ln(\phi^6). \end{cases} \quad (4.22)$$

Finally, we have as our expression for $\alpha(\lambda)$,

$$\alpha(\lambda) = \begin{cases} [1 - \lambda\theta / \ln(\phi^6)]^2, & |\lambda\theta| < \ln(\phi^6), \\ 0, & \lambda\theta > \ln(\phi^6), \\ -4\lambda\theta / \ln(\phi^6), & \lambda\theta < -\ln(\phi^6). \end{cases} \quad (4.23)$$

[However, the singularities are due to the approximation used for $f(x)$, and will not be present in an exact calculation, as explained in the *Note added in proof*.]

B. Wave function at the edge of the band

Consider a two-cycle of the trace map

$$\begin{aligned} x &\rightarrow y \rightarrow x \rightarrow y \rightarrow \dots, \\ x &= J + (J^2 - J)^{1/2}, \\ y &= J - (J^2 - J)^{1/2}, \end{aligned} \quad (4.24)$$

with

$$J = [3 + (25 + 16I)^{1/2}] / 8, \quad (4.25)$$

where I is the constant of motion given by (2.16) or (2.25). We note that $x > 1$ and $y < 1$, and the energy which gives an orbit on the stable manifold of this two-cycle is the largest in the spectrum. Thus the wave function corresponding to the two-cycle represents the state at the edge of the spectrum.

Let \underline{M}_j ($j=0, 1, 2, \dots$) be the orbit of the full dynamical map corresponding to the trace map. Then,

$$\frac{1}{2} \text{Tr} \underline{M}_j = \begin{cases} x & \text{for } j=0, 2, 4, \dots, \\ y & \text{for } j=1, 3, 5, \dots. \end{cases} \quad (4.26)$$

As explained in Sec. III, these matrices are related by a Lorentz transformation

$$\underline{M}_{j+2} = \underline{T} \underline{M}_j \underline{T}^{-1}. \quad (4.27)$$

Consider a series of matrices

$$\underline{M}_0 \underline{M}_2 \underline{M}_4 \dots \underline{M}_{2m} = (\underline{M}_0 \underline{T})^m \underline{T}^{-m}, \quad (4.28)$$

$m=1, 2, 3, \dots$. The larger eigenvalue of these matrices grows algebraically and the envelope of the wave function is determined at the sites $F_2, F_2 + F_4, F_2 + F_4 + F_6, \dots$. Thus an exponent β is calculated once \underline{T} is determined. Unfortunately a calculation of \underline{T} is tedious and it does not have a compact form. Therefore, we present data for β numerically calculated from the formula.

The results of a numerical evaluation of the explicit formula are shown in Fig. 3. The growth of the wave function at the selected sites $F_2 + F_4 + \dots + F_{2m} \rightarrow \phi^{2m}$, where $m=1, 2, 3, \dots$, is the peak number and is determined by the maximum eigenvalue of the corresponding transfer matrix $\underline{M}_0 \underline{M}_2 \dots \underline{M}_{2m}$ of (4.28). This maximum eigenvalue grows exponentially with the peak number and depends upon the value of the invariant I . In Fig. 3 we show the dependence by plotting the logarithm of the maximum eigenvalue of the transfer matrix of (4.28) divided by the peak number m , as a function of the invariant I , for values of $m=1, 2, \dots, 10$; these are the lower curves in Fig. 3. The upper curve is the exact calculation of the limiting behavior, as calculated by the formalism of Sec. III. Obviously the fit is excellent and the approach to the limiting curve at $I=0$ is nonuniform. The connection with the exponent β is $\beta \ln(\phi^2) = \ln(\text{maximum eigen-})$

value)/(peak height). Thus, Fig. 3 can represent β by simply changing the vertical scale by a factor of $1/\ln(\phi^2)$.

C. Numerical results for the wave functions

In order to calculate a wave function numerically, we first need to specify an energy in the spectrum. Since the energy spectrum is a Cantor set with zero Lebesgue measure,^{18,19} it is impossible, in principle, to specify an energy in the spectrum numerically. And also we do not expect a smooth change of behavior of a wave function as the energy is varied in the spectrum. Equivalently, two wave functions look different at a sufficiently long distance length scale no matter how small the energy difference. (This property may have important relevance to quantum chaos.) Therefore, we need a careful treatment of numerical calculations based on the knowledge of the Cantor-set spectrum which comes from the KKT renormalization-group method.

The energy spectrum is divided into three subclusters at each hierarchical step of partition. (See Fig. 2.) Therefore, a point in the spectrum is specified by an infinite sequence of symbols 1, 0, and $\bar{1}$, where 1 represents an upper subcluster; 0, a middle subcluster; and $\bar{1}$, a lower subcluster. The sequence $\{C_n\}$ with $C_n=0$ for all the positive integers n represents the center of the band and we have the six-cycle wave function discussed in Sec. IV A. Also the sequence $\{C_n\}$ with $C_n=1$ (or $\bar{1}$) for all the positive integers n represents the edge of the band and we have the two-cycle wave function of Sec. IV B.

In Fig. 4 the wave function at the middle of the band $E=0$ is shown. This wave function corresponds to the six-cycle and has a code $\{0,0,0,\dots\}$. In the successive figures, the sites are rescaled around the peak. The self-similarity of the wave function is clearly seen.

The wave function at the edge of the band is shown in Fig. 5. This wave function corresponds to the two-cycle and has a code $\{1,1,1,\dots\}$. The self-similarity is clearly seen in the figures. The quantity $S(n)$ which represents the concentration of the measure [see (4.2)] with $\lambda=1$ for this wave function is plotted in Fig. 6. The scaling form of (4.2) is supported by the linear behavior in this $\ln S(n)$

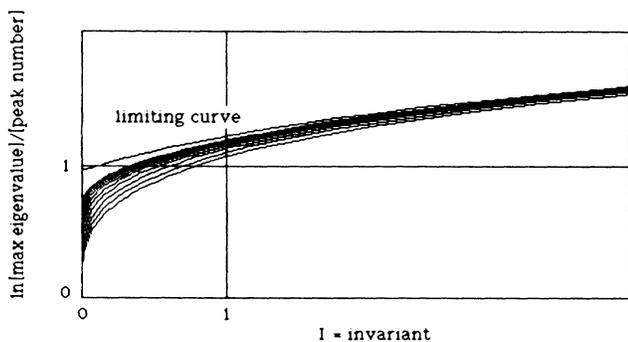


FIG. 3. Logarithm of the maximum eigenvalue of the transfer matrix (4.28) by the peak number m versus the invariant I , for values $m=1$ to 10 and the exact limiting curve. The exponent β is obtained from this curve by dividing by $\ln\phi^2$.

versus $\ln n$ plot. An analytical study of the index $\alpha(\lambda)$ for this wave function has not yet been performed.

An interesting wave function which has a code $\{1,1,1,1,0,0,0,0,\dots\}$ is shown in Fig. 7. The trace map corresponding to this wave function starts near the stable manifold of the two-cycle and then makes a crossover to the stable manifold of the six-cycle. Therefore, it looks similar to the six-cycle wave function at a large length scale [comparing Fig. 7(a) with Fig. 4], but it looks similar to the two-cycle wave function at a small length scale [compare Fig. 7(c) with Fig. 5]. Note that this wave function is at an energy which is only about 10^{-4} away from the edge.

Finally, we present an example of wave functions which have a code of random sequence. In Fig. 8, a wave function with a code $\{0,1,\bar{1},\bar{1},0,\bar{1},1,0\}$ is shown from 0 to $F_{21}=17711$. At this range, the rest of the code does not influence the shape of the wave function. The concentration of the measure $S(n)$ is plotted in Fig. 9. Although there is no apparent scaling, the wave function does not seem to be totally random. The understanding of this type of wave function in relation to quantum chaos still remains as an important unsolved problem.

V. GLOBAL SCALING PROPERTIES OF THE ENERGY SPECTRUM

Local scaling of a spectrum can be defined as follows. Let E and $E+\Delta E$ both be in the spectrum. If the integrated density of states behaves as

$$D(E+\Delta E)-D(E)\sim(\Delta E)^{\alpha_E} \text{ as } \Delta E\rightarrow 0, \quad (5.1)$$

we say that the spectrum has scaling at E with a scaling index α_E .

In Ref. 19, α_E is calculated exactly at the center of the band from the fixed-point analysis of the six-cycle and is given by

$$\alpha_E(\text{center})=\ln\phi^6/\ln\varepsilon_6, \quad (5.2)$$

where ϕ is the golden mean $(\sqrt{5}+1)/2$, and

$$\varepsilon_6=\{[1+4(1+I)^2]^{1/2}+2(1+I)\}^2 \quad (5.3)$$

is the eigenvalue of the linearized equation around the six-cycle. The constant of motion I is given in (2.16) or (2.25).

Since the edge of the spectrum is governed by the two-cycle, α_E is obtained exactly there and given by

$$\alpha_E(\text{edge})=\ln\phi^2/\ln\varepsilon_2, \quad (5.4)$$

where

$$\varepsilon_2=\{8J-1+[(8J-1)^2-4]^{1/2}\}/2 \quad (5.5)$$

with

$$J=[3+(25+16I)^{1/2}]/8$$

is the eigenvalue of the linearized equation around the two-cycle.

The center of each subcluster also has scaling with the index $\alpha_E(\text{center})$. A point of this type is coded by $\{C_n\}$ ($n=1,2,3,\dots$) with $C_i=0$ for $i>N$, where N is an in-

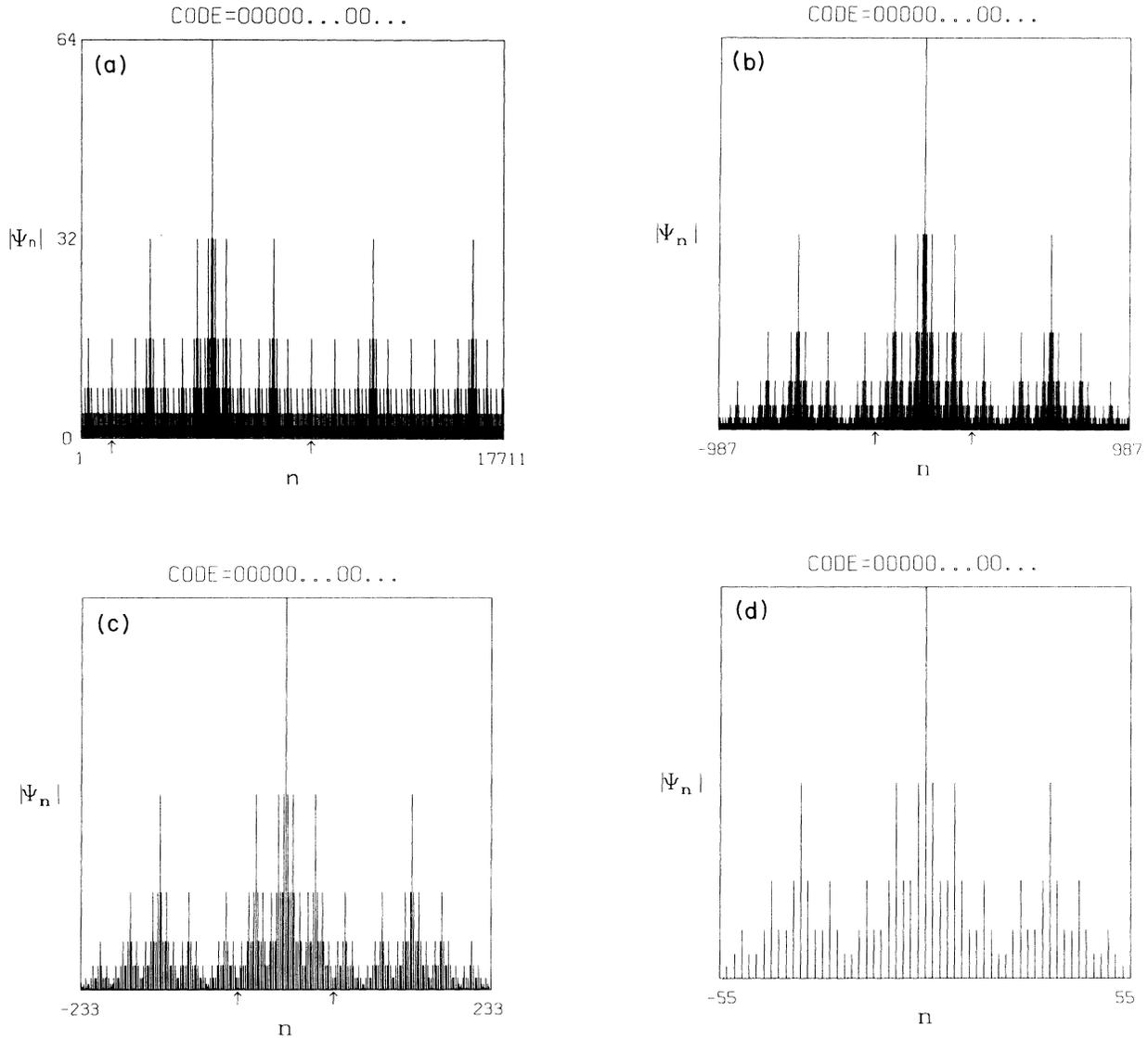


FIG. 4. Wave function at the center of the spectrum $E=0$ of (1.1) with $t_A=1$ and $t_B=2$. The portion around the maximum limited by the two arrows in (a) is rescaled and shown in (b). (c) and (d) are rescaled in the same procedure.

teger. (See Sec. IV C.) Also an edge of a subcluster is coded by a $\{C_n\}$ with $C_i=1$ (or $\bar{1}$) for $i > N$. The number of these points is infinite, but is only countably infinite. Since the number of elements in the Cantor-set spectrum is uncountably infinite, we expect much richer scaling than that being represented by the indices $\alpha_E(\text{center})$ and $\alpha_E(\text{edge})$.

It turns out that every point in the spectrum has scaling and a scaling index takes a value in a range $[\alpha_E^{\min}, \alpha_E^{\max}]$, where the maximum value is $\alpha_E(\text{center})$ and the minimum value is $\alpha_E(\text{edge})$. The centers are the most rarefied regions and the edges are the most dense regions. There are infinitely many points in the spectrum which have a given scaling index α_E . A set of these points is fractal and

given a fractal dimension $f(\alpha_E)$.³³ The fractal dimension $f(\alpha_E)$ is a kind of measure of density of points with an index α_E . As a function of α_E , $f(\alpha_E)$ is smooth and represents global scaling properties of the spectrum. Since the number of points with $\alpha_E^{\min}=\alpha_E(\text{edge})$ and $\alpha_E^{\max}=\alpha_E(\text{center})$ is only countably infinite, the fractal dimension $f(\alpha_E)$ vanishes there and is positive in between.

Let us calculate $f(\alpha_E)$ following Ref. 33. First we make a partition of the spectrum. In Sec. II it is shown that the spectrum for the quasiperiodic model is given by $|x_j(E)| \leq 1$ as $j \rightarrow \infty$. [See (2.9), (2.10), and (2.11).] If we apply this condition for a finite j , the resulting set of E represents a spectrum of a periodic system with period F_j . As j is increased, this spectrum becomes a better ap-

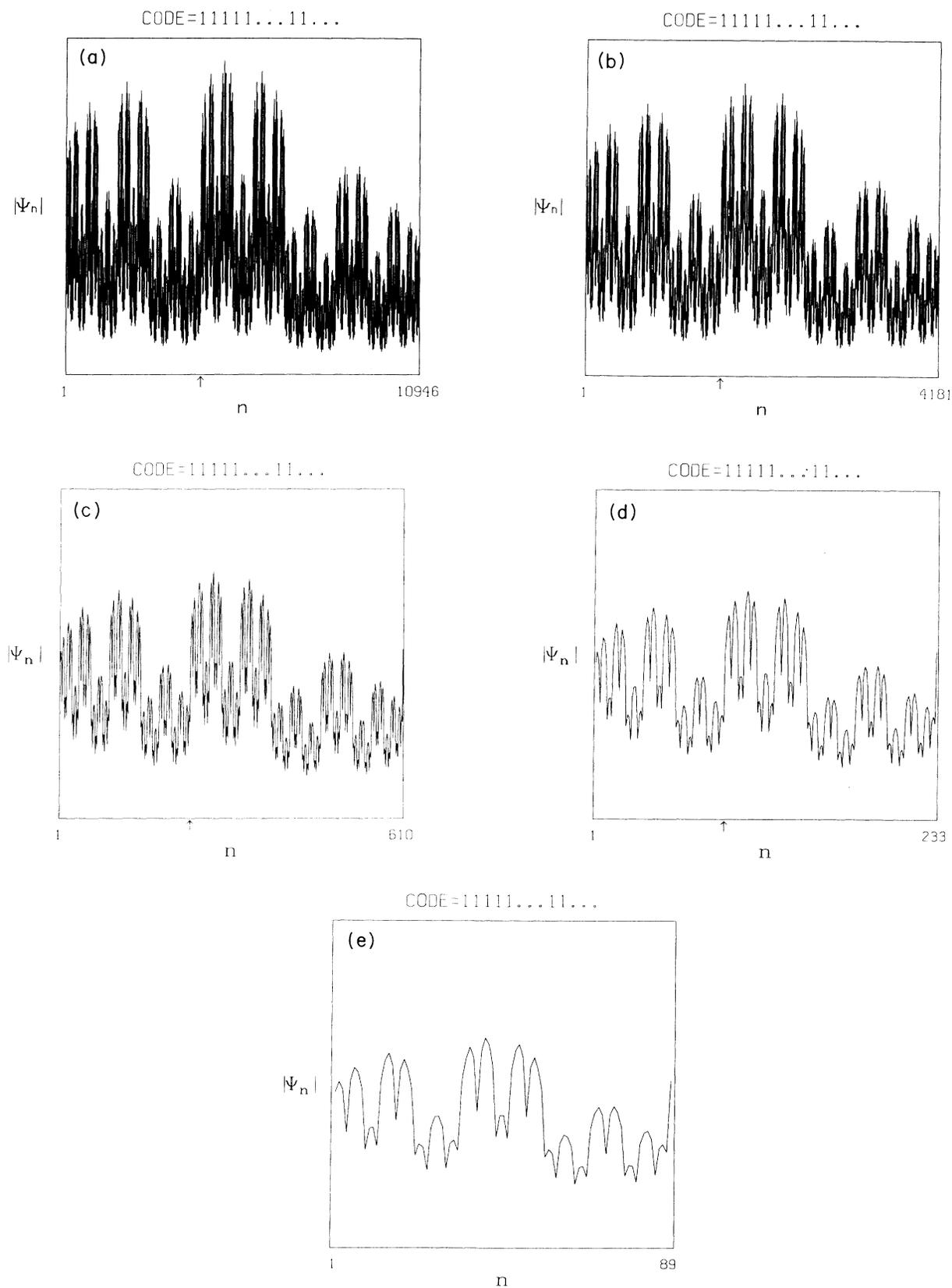


FIG. 5. Same as Fig. 4 for the wave function at the edge of the spectrum $E = 2.833\,956\,5$.

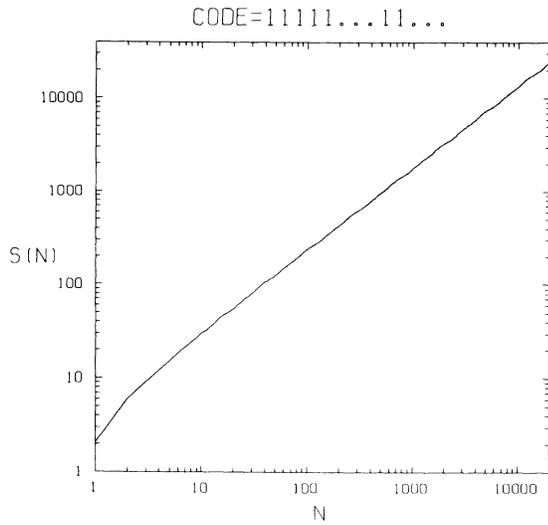


FIG. 6. Function $S(N)$ [see (6.2)] for the wave function at the edge of the spectrum of (1.1) with $t_A=1$ and $t_B=2$.

proximation of the Cantor-set spectrum. We use a spectrum of the periodic system as a j th partition of the Cantor-set spectrum.

For a fixed j , the spectrum consists of F_j bands. Denote a bandwidth of the i th band by $(\Delta E)_i$. Each band contains the same number of density of states (measure) $P_i=1/F_j$. Define a partition function by

$$\Gamma_j(q, \tau) = \sum_{i=1}^{F_j} \frac{(P_i)^q}{(\Delta E_i)^\tau} \quad (5.6)$$

and also α_i by

$$P_i = (\Delta E_i)^{\alpha_i} . \quad (5.7)$$

The partition function has a limit

$$\Gamma(q, \tau) = \lim_{j \rightarrow \infty} \Gamma_j(q, \tau) , \quad (5.8)$$

which is either zero or infinity unless τ is chosen appropriately for a given value of q . Therefore, the condition $\Gamma(q, \tau)=1$, for example, specifies a function $\tau(q)$. It is argued in Ref. 33 that α_i has a definite limit α_E as $j \rightarrow \infty$ and the fractal dimension for a set of points having a given value of α_E is related to $\tau(q)$ by a Legendre transformation

$$f(\alpha_E) = -\tau(q) + q\alpha_E . \quad (5.9)$$

In addition, it is conjectured in Ref. 32 that the existence of $f(\alpha_E)$ is a condition of a singular continuous spectrum. For an absolutely continuous spectrum which corresponds to extended states, the scaling is "trivial," namely, the f - α_E curve consists of a point $f=1, \alpha_E=1$ and a point $f=0, \alpha_E=\frac{1}{2}$. The point $f(1)=1$ represents a measure (density of states) that does not have singular concentration. The point $f(\frac{1}{2})=0$ represents remnants of Van Hove singularities whose number is at most count-

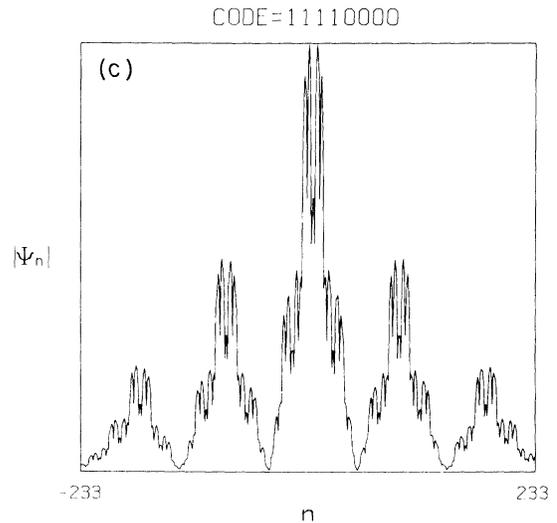
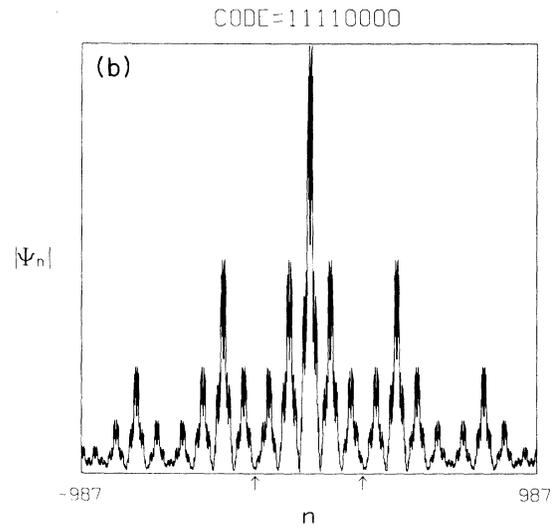
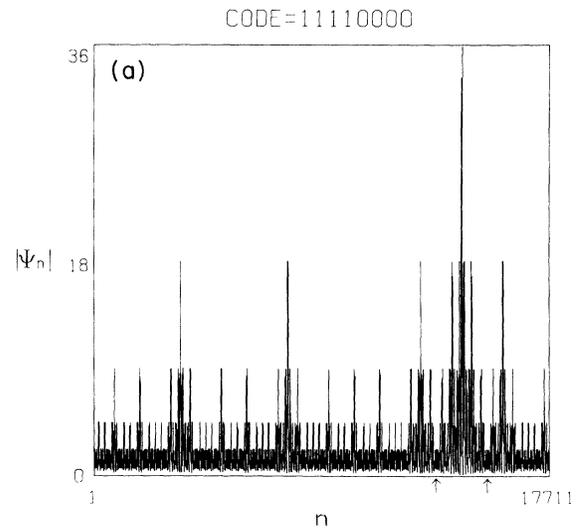


FIG. 7. Same as Fig. 4 for the wave function at $E=2.833\,023\,89$ (code = {1,1,1,1,0,0,0,0,...}).

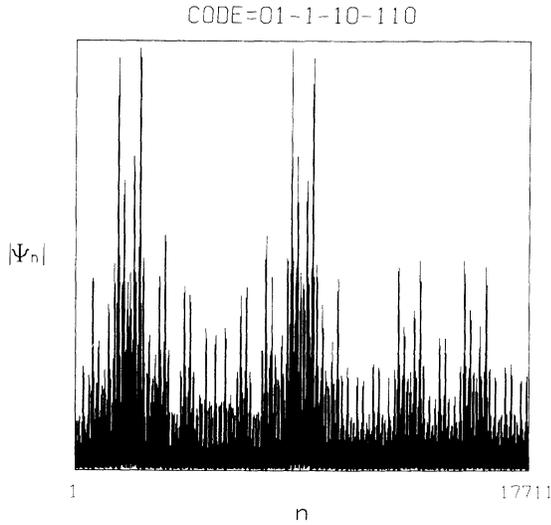


FIG. 8. Non-self-similar wave function at $E=0.267958$ (code = {0, 1, $\bar{1}$, $\bar{1}$, 0, $\bar{1}$, 1, 0, ...}).

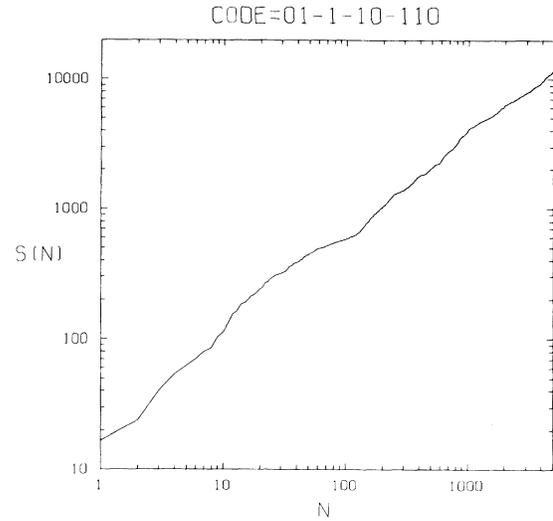


FIG. 9. Function $S(n)$ for the non-self-similar wave function of Fig. 8.

ably infinite. For a point spectrum, a measure is more singularly concentrated than those represented by (5.7).

The f - α_E curve for the present model (1.1) is calculated using a condition

$$\Gamma_j(q, \tau) / \Gamma_{j'}(q, \tau) = 1$$

for sufficiently large values of j and j' . This condition is numerically more efficient than the condition

$$\lim_{j \rightarrow \infty} \Gamma_j(q, \tau) = 1.$$

In Fig. 10 an example of $f(\alpha_E)$ computed from a condition $\Gamma_{12}/\Gamma_{15}=1$ is shown together with the exact values of α_E^{\min} and α_E^{\max} from (5.2) and (5.4). The maximum of $f(\alpha_E)$ is a Hausdorff dimension of the Cantor-set spectrum. For all the values of coupling ($t_B/t_A = e^\theta$) treated, we obtained a similar $f(\alpha_E)$ curve. However, exact shapes of the curves depend on the coupling. This numerical work gives a strong support to the conjecture that the spectrum of the one-dimensional quasicrystal is singular continuous and the scaling properties depend on the coupling. This is an analogous situation to the low-temperature phase (Kosterlitz-Thouless phase) of the two-dimension XY model.

VI. SUMMARY

The electronic properties of a one-dimensional quasicrystal are studied. The wave functions at the center of the band and the edge of the band are investigated by the scaling analysis based on the KKT renormalization-group method. The fractal nature (or self-similarity) and other critical properties of these wave functions are well understood. Also, the existence of non-self-similar wave functions is shown. An analysis of this type of wave function

in relation to quantum chaos remains an important unsolved problem.

The energy spectrum is shown to be a Cantor set with infinitely many types of scaling. There are infinitely many points in the spectrum for a given type of scaling. The fractal dimension for those points is numerically determined. As a by-product of this analysis, the Hausdorff dimension of the spectral set is obtained.

The existence of scaling properties of the wave functions and the energy spectrum shown in this work is con-

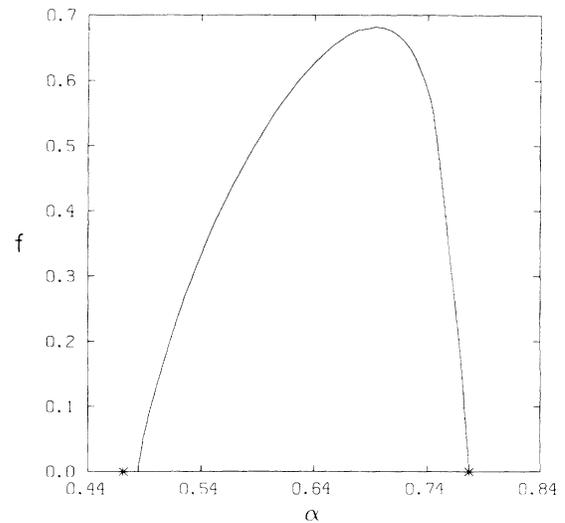


FIG. 10. f - α_E curve for the Cantor-set spectrum of (1.1) with $t_A=1$ and $t_B=2$. * denotes the exact values of α_E^{\min} and α_E^{\max} .

sistent with the conjecture that the spectrum of the quasi-periodic model is purely singular continuous, irrespective of the coupling constant. The scaling indices change continuously as the coupling constant is varied.

Note added in proof. After this paper was accepted for publication, one of the authors (B.S.) succeeded in calculating exactly the scaling function, as well as the exponent $\alpha(\lambda)$, for the six cycle. The results agree quantitatively very well with the numerical results in this paper. However, qualitatively they differ in that the exact result has no singularities at the point $|\lambda\theta| = \ln(\phi^6)$. This difference arises because of the difficulty in determining from the numerical results whether the scaling function in fact vanishes with a large but finite slope, or whether it vanishes faster than linearly. These new results are sufficiently involved that they will be published as a separate paper.

ACKNOWLEDGMENT

We are grateful to L. P. Kadanoff for very useful conversations.

APPENDIX: DERIVATION OF THE TRACE MAP

Let us regard the recursion $\underline{M}_{j+1} = \underline{M}_{j-1}\underline{M}_j$ as the transformation $(\underline{A}, \underline{B}) \rightarrow (\underline{A}', \underline{B}')$ with

$$\underline{A}' = \underline{B}\underline{A}, \quad \underline{B}' = \underline{A}, \quad (\text{A1})$$

where \underline{A} , \underline{B} , \underline{A}' , and \underline{B}' are 2×2 matrices with unit determinant. We represent them as

$$\underline{A} = y\tau_0 + a_1\tau_1 + a_2\tau_1 + a_2\tau_2 + a_3\tau_3, \quad (\text{A2})$$

$$\underline{B} = x\tau_0 + b_1\tau_1 + b_2\tau_2 + b_3\tau_3,$$

where τ_0 , τ_1 , τ_2 , and τ_3 are related to the Pauli matrices. [See (3.1)–(3.4).] The space of \underline{A} and \underline{B} are parametrized by eight numbers $(x, y, a_1, b_1, a_2, b_2, a_3, b_3)$, but we have two constraints

$$y^2 = \mathbf{a} \cdot \mathbf{a} = 1, \quad x^2 + \mathbf{b} \cdot \mathbf{b} = 1, \quad (\text{A3})$$

which make the determinants of \underline{A} and \underline{B} unity. Here the scalar product of three-vectors is defined as

$$\mathbf{a} \cdot \mathbf{a} = a_1 a_1 - a_2 a_2 - a_3 a_3, \quad (\text{A4})$$

$$\mathbf{b} \cdot \mathbf{b} = b_1 b_1 - b_2 b_2 - b_3 b_3.$$

Thus the transformation (A1) represents a map in a six-dimensional space. Let us pick three coordinates x , y , and $w = \mathbf{a} \cdot \mathbf{b} = a_1 b_1 - a_2 b_2 - a_3 b_3$ out of six. It turns out that (w, x, y) uniquely determines (w', x', y') . Straightforward matrix multiplication gives

$$w' = yw + x(1 - y^2), \quad x' = y, \quad y' = yx - w. \quad (\text{A5})$$

Since $\underline{M}_{j-1} = \underline{B}$, $\underline{M}_j = \underline{A}$, and $\underline{M}_{j+1} = \underline{M}_{j-1}\underline{M}_j$, we have $y' = \frac{1}{2} \text{Tr} \underline{M}_{j+1}$, $x' = y = \text{Tr} \underline{M}_j$, and $x = \frac{1}{2} \text{Tr} \underline{M}_{j-1}$. From (A5) we have

$$\begin{aligned} \text{Tr} \underline{M}_{j+2} &= 2y'' = 2(y'x' - w') \\ &= 2(2yy' - x) \\ &= \text{Tr} \underline{M}_{j+1} \text{Tr} \underline{M}_j - \text{Tr} \underline{M}_{j-1}, \quad \text{Q.E.D.} \end{aligned}$$

*Present address: Department of Physics, Brookhaven National Laboratory, Upton, NY 11973-5000.

¹D. Schechtman, I. Blech, D. Gratias, and J. W. Cahn, Phys. Rev. Lett. **53**, 1951 (1984).

²D. Levine and P. J. Steinhardt, Phys. Rev. Lett. **53**, 2477 (1984).

³R. Penrose, Bull. Inst. Math. Appl. **10**, 266 (1974).

⁴M. Gardner, Sci. Am. **236**, 110 (1977).

⁵N. G. de Bruijn, Ned. Akad. Wet. Proc. Ser. A **84**, 39 (1981).

⁶N. G. de Bruijn, Ned. Akad. Wet. Proc. Ser. A **84**, 53 (1981).

⁷A. L. MacKay, Physica **114A**, 609 (1982).

⁸A. L. MacKay, Kristallografiya **26**, 910 (1981) [Sov. Phys.—Crystallogr. **26**, 517 (1981)].

⁹P. Kramer and R. Neri, Acta Crystallogr. Sect. A **40**, 580 (1984).

¹⁰P. A. Kalugin, A. Yu Kitaev, and L. C. Levitov, Pis'ma Zh. Eksp. Teor. Fiz. **41**, 119 (1985) [JETP Lett. **41**, 119 (1985)].

¹¹V. Elster, Phys. Rev. B **32**, 4892 (1985).

¹²M. Duneau and A. Katz, Phys. Rev. Lett. **54**, 2688 (1985).

¹³R. K. P. Zia and W. J. Dallas, J. Phys. A **18**, L341 (1985).

¹⁴T. Odagaki and D. Nguyen, Phys. Rev. B **33**, 2184 (1986).

¹⁵T. C. Choy, Phys. Rev. Lett. **55**, 2915 (1985).

¹⁶M. Kohmoto and B. Sutherland, Phys. Rev. Lett. **56**, 2740 (1986); Phys. Rev. B **34**, 3849 (1986).

¹⁷H. Tsunetsugu, T. Fujiwara, K. Ueda, and T. Tokihiro, J. Phys. Soc. Jpn. **55**, 1420 (1986).

¹⁸M. Kohmoto, L. P. Kadanoff, and C. Tang, Phys. Rev. Lett. **50**, 1870 (1983).

¹⁹M. Kohmoto and Y. Oono, Phys. Lett. **102A**, 145 (1984).

²⁰S. Ostlund, R. Pandit, D. Rand, H. J. Schellnhuber, and E. Siggia, Phys. Rev. Lett. **50**, 1873 (1983).

²¹M. Kohmoto and J. R. Banavar, Phys. Rev. B **34**, 563 (1986).

²²J. P. Lu, T. Odagaki, and J. L. Birman, Phys. Rev. B **33**, 4809 (1986).

²³F. Nori and J. P. Rodriguez, Phys. Rev. B **34**, 2207 (1986).

²⁴M. Fujita and K. Machida, Solid State Commun. **59**, 61 (1986).

²⁵J. M. Luck and D. Petritis, J. Stat. Phys. **42**, 289 (1986).

²⁶B. Simon, Adv. Appl. Math. **3**, 463 (1982).

²⁷J. B. Sokoloff, Phys. Rep. **126**, 189 (1985).

²⁸See, e.g., M. Reed and B. Simon, *Methods of Modern Mathematical Physics, IV, Analysis of Operators* (Academic, New York, 1978).

²⁹M. Kohmoto, Phys. Rev. Lett. **51**, 1198 (1983).

³⁰S. Ostlund and R. Pandit, Phys. Rev. B **29**, 1394 (1984).

³¹G. Andre and S. Aubry, Ann. Israel Phys. Soc. **3**, 133 (1980).

³²C. Tang and M. Kohmoto, Phys. Rev. B **34**, 2041 (1986).

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

³⁴This implies that the Lyapunov exponent (=inverse of localization length) defined by $\gamma = \lim_{N \rightarrow \infty} [N^{-1} \ln |\underline{M}(N)\underline{M}(N-1) \cdots \underline{M}(2)\underline{M}(1)|]$ is zero in the spectrum for this model. A localized state if it exists has a positive Lyapunov exponent. However, the absence of localization for this model is proven rigorously in Ref. 35.

³⁵F. Delyon and D. Petritis, Commun. Math. Phys. **103**, 441 (1986).

³⁶B. Sutherland, Phys. Rev. Lett. **57**, 770 (1986).